

1-1-2021

Evaluation of psychometric properties of the Thai SF-12v2 in the general Thai population

Krittaphas Kangwanrattanakul

Follow this and additional works at: <https://digital.car.chula.ac.th/tjps>



Part of the [Pharmacology Commons](#)

Recommended Citation

Kangwanrattanakul, Krittaphas (2021) "Evaluation of psychometric properties of the Thai SF-12v2 in the general Thai population," *The Thai Journal of Pharmaceutical Sciences*: Vol. 45: Iss. 6, Article 14.
Available at: <https://digital.car.chula.ac.th/tjps/vol45/iss6/14>

This Article is brought to you for free and open access by the Chulalongkorn Journal Online (CUJO) at Chula Digital Collections. It has been accepted for inclusion in The Thai Journal of Pharmaceutical Sciences by an authorized editor of Chula Digital Collections. For more information, please contact ChulaDC@car.chula.ac.th.



Evaluation of psychometric properties of the Thai SF-12v2 in the general Thai population

Krittaphas Kangwanrattanakul

Department of Social and Administrative Pharmacy, Faculty of Pharmaceutical Sciences, Burapha University, Chonburi, Thailand

Corresponding

Author: Krittaphas Kangwanrattanakul,
Department of Social and Administrative Pharmacy,
Faculty of Pharmaceutical Sciences, Burapha University,
Chonburi, Thailand.
Phone: +663-839-0400-5. Fax: +663-839-0401.
E-mail: krittaphas@buu.ac.th

Received: Feb 26, 2021

Accepted: Apr 14, 2021

Published: Dec 04, 2021

ABSTRACT

Background: This study investigated the psychometric properties of the Thai Short Form-12 Health Survey (SF-12v2) and reported SF-12v2 scale scores based on the general Thai population's age and gender. **Methods:** Multistage sampling was used to recruit 1200 general Thai participants. For all SF-12v2 scales, item description, ceiling/floor effects, convergent validity for item-scale and SF-12v2 and EQ-5D correlations, known-group validity, exploratory factor analysis (EFA), test-retest reliability, and responsiveness were studied. **Results:** Ceiling effect was observed for all scales (5.8–72.6%). Floor effect was detected for all scales except for social functioning, role emotion, and mental health. Item-scale correlation coefficients ranged from 0.65 to 1.00. Moderate correlations between SF-12v2 and EQ-5D were observed. EFA yielded a two-factor structure accounting for 64% of the total variance. Scale scores were lower among female, older, and poorly educated individuals and those with lower incomes and chronic diseases. Test-retest reliability demonstrated fair-to-good reproducibility for all scales (Intraclass correlation coefficients: 0.49–0.64). Moreover, the SF-12v2 was more sensitive to health deterioration than to health improvement. **Conclusion:** These preliminary results confirmed that the Thai SF-12v2 is valid and reliable. Therefore, the Thai SF-12v2 is an appropriate health-related quality of life instrument for the Thai population.

Keywords: Evaluation, general population, health-related quality of life, psychometric properties, Thai SF-12v2

INTRODUCTION

In recent years, attention to patient-reported outcomes (PRO) has increased because they are directly reported by patients simply to convey their health status, quality of life, or functional status related to medical treatment.^[1,2] Health-related quality of life (HRQoL) is a preferred PRO because it can complement safety data, survival rates, effect of drug treatment, policy decision-making, evaluation of health care service, medical intervention, and population health surveys.^[3-9] Because HRQoL is an important PRO in medical intervention, one important concern is selection of the most appropriate instrument to elicit HRQoL levels for individual patients.

Two common instruments used to measure HRQoL levels include generic and disease-specific instruments. The disease-specific instrument is designed to measure HRQoL levels in a disease-specific population;^[10,11] however, there are limitations to comparing HRQoL levels across different populations.

The generic instrument is developed to measure HRQoL levels in both general populations and the therapeutic area. Nevertheless, the sensitivity and relevance to some disease-specific populations are limited.^[12] Evidence has revealed that the Short Form-36 Health Survey (SF-36) is one of the most common, generic instruments capable of evaluating HRQoL levels in both patient groups and population surveys.^[13-15] Therefore, the SF-36 can be administered to measure and compare HRQoL levels in a variety of types and severities of disease, treatments, or interventions, and across varied demographic and cultural subgroups.^[6,16] The SF-36 has 36 questions covering eight dimensions:^[17,18] (1) Physical functioning (PF: 10 items), (2) role limitations due to physical problems (RP: 4 items), (3) bodily pain (BP: 2 items), (4) general health perception (GH: 5 items), (5) vitality (VT: 4 items), (6) social functioning (SF: 2 items), (7) role limitations due to emotional problems (RE: 3 items), and (8) mental health (MH: 5 items), along with a single item on health transition.^[19,20] Furthermore, these eight scales form two

summary component scores related to physical and mental health, the physical component score (PCS) and the mental component score (MCS).^[20] Several studies have reported the SF-36's validity and reliability in both the therapeutic area and the general population.^[21-29]

The SF-36 is too lengthy for some applications and might burdensome respondents, particularly in a large population survey; further, some psychometric weaknesses were found, including poor reliability due to inconsistent rating scales across all eight dimensions and several negatively worded items.^[30] Therefore, the SF-12 was developed as an abridged SF-36 version to measure HRQoL levels. Moreover, the SF-12 is expected to retain strong measurement properties and to reproduce PCS and MCS with fewer items.^[31] Conveniently, it can be administered as a one-page health survey in an average of <2 min.^[32]

The SF-12 can be used as an alternative instrument to measure HRQoL levels in both large surveys and patient groups, focusing on overall physical and mental health outcomes.^[31,33] The SF-12 also showed substantial correlations of the two summary components, PCS and MCS, estimated using the SF-12 and SF-36 across all nine European countries and China,^[31,34] further suggesting the SF-12's use as an alternative to the SF-36 in HRQoL measurement.

To date, the Short Form-12 version 2 (SF-12v2) is available with several modifications from the SF-12v1, for example, simplified wording, revised layout to improve comprehension and reduce missing responses, and revised response options from a dichotomous (yes/no) to a 5-point Likert Scale in RP and RE dimensions, while reducing response options from a 6- to a 5-point Likert Scale for items in MH and VT dimensions.^[32,35] Several previous studies have supported the SF-12v2 instrument's use in both general population^[36-39] and patient groups.^[40-47]

Most studies have investigated the Thai SF-36's psychometric properties in the general Thai population^[28,29] and in clinical areas, including cardiac patients^[48] and those with low back pain.^[49] Although the Thai SF-12v2 is available, no previous studies have investigated its psychometric properties in the general Thai population. Therefore, the present study evaluated the Thai SF-12v2's psychometric properties and reported the instrument's scores based on the general Thai population.

METHODS

Study Design

This study's dataset was applied from "Psychometric properties comparison between EQ-5D-5L and EQ-5D-3L in the general Thai population," a cross-sectional survey conducted with a convenience sample ($n = 1200$) living in five provinces across Thailand: Nakhon-Srithammarat, Khon-Kaen, Chonburi, Chaing-Mai, and Bangkok (the capital city). A three-stage stratified random sampling was employed to select provinces, districts, and villages. However, the local village leaders approached all convenient samples and asked them to participate in the study and complete the questionnaire. Convenient samples from each selected household were

selected based on the proportions of age and gender to ensure that samples were representative of the general Thai population. Study participants were selected if (1) they were from 18 to 70 years old and (2) they could read and understand the data collection process, as evaluated by the interviewers or researcher. Potential participants were excluded if they had (1) acute or life threatening disease, (2) cognitive impairment, or (3) disability.

Data Collection

For the first of two data collection phases, all convenient participants were asked to complete the self-administered EQ-5D-3L (3L), EQ-5D-5L (5L), and SF-12v2 at their residences with interviewers who were allowed to read questions and response options without explaining or elaborating new meanings for those who were elderly or had eyesight problems. Moreover, before the study commenced, all participants received an information sheet written in plain language to explain study purposes and the overall research process. Of course, they could withdraw from the study at any time if they felt uncomfortable. This study was approved by the Burapha University Institutional Review Board (BUU-IRB): 108/2562. All participants provided written informed consent.

In the second data collection phase, 400 participants were randomly selected to complete the questionnaire again at home to assess whether their health status had changed at 2 weeks after their first self-assessment, using a five-point Likert scale: (1) Much better, (2) somewhat better, (3) same as the previous assessment, (4) somewhat worse, and (5) much worse. Participants returned the questionnaire by mail, and any questionnaires reaching the researcher after 21 days were excluded from the analysis.

Instrument

This study used the Thai SF-12v2, an abridged version of the SF-36v2, with 12 items grouped into eight scales/dimensions including physical functioning (PF: 2 items), role physical (RP: 2 items), bodily pain (BP: 1 item), general health (GH: 1 item), vitality (VT: 1 item), social functioning (SF: 1 item), role emotion (RE: 2 items), and mental health (MH: 2 items). Each scale score ranges from 0 to 100, with higher scores indicating a better health condition. These 12 items also produce two summary component scores, the PCS and the MCS,^[35,50] which can be transformed into norm-based scores of the United States general population survey (mean = 50, standard deviation [SD] = 10).^[31,51]

The EQ-5D is a brief, self-administered questionnaire for measuring current health status. It has five dimensions: Mobility (MO), self-care (SC), usual activities (UA), pain/discomfort (PD), and anxiety/depression (AD). The EQ-5D has two versions—the 3L and the 5L—which have three and five response items, respectively, in each dimension. The respondents were asked to complete their current health status on the day the questionnaire was administered.

Data Analyses

Participants' general characteristics and item descriptions were reported through descriptive statistics. For general

characteristics, frequencies and percentage explained gender, marital status, educational level, family monthly income, health status, smoking, and alcohol consumption, while mean and SD were reported for participants' ages. Responses to each item and the SF-12v2 scale score were reported using percentage and mean \pm SD, respectively.

Ceiling and floor effects were computed as percentage of participants rating the best and worst health status for each SF-12v2 scale divided by the total number of participants, respectively. The acceptable percentage of these two effects was set at $<15\%$.^[52]

Convergent validity was assessed to determine whether items correlated with their hypothesized scale through Spearman's rho correlation. Colton's rule was employed to determine the strength of correlations as follows: weak or none ($r < 0.25$), moderate ($0.25 \leq r < 0.50$), moderate to strong ($0.50 \leq r < 0.75$), and strong ($r \geq 0.75$).^[53] The convergent validity was also investigated between SF-12v2 items and EQ-5D dimensions using Spearman's rho correlation. High correlations were expected between similar dimensions of both SF-12 items and EQ-5D dimensions. Strong correlations were hypothesized between PF/MO, BP/PD, and MH/AD.

Known-group validity was performed to determine whether SF-12v2 scale scores changed among participant subgroups as defined by sociodemographic and health-related factors. SF-12v2 scale scores were hypothesized to be lower among women, smokers/ex-smokers, drinkers/ex-drinkers, older samples (≥ 60 years), and those with poor education levels (no schooling or primary school), lower incomes ($\leq 30,000$ THB or 990 USD), and suffering from disease. Among participant characteristics, the student's *t*-test or analysis of variance determined differences in SF-12v2 scale scores.

Exploratory factor analysis (EFA) was employed to investigate the structural validity of the SF-12v2 using principal component analysis with varimax rotation. The underlying structure of the questionnaire was identified if in the case of Eigenvalue >1 and factor loadings ≥ 0.4 . It was hypothesized that EFA yielded a two-factor structure (physical and mental health components) in which PF, RP, BP, and GH items were loaded on physical health component, and the RE, SF, MH, and VT items were loaded on mental health component.

Reliability was assessed by test-retest reliability for participants indicating no health status change at 2 weeks. Intra-class correlation coefficients (ICCs) determined reliability for each SF-12v2 scale. Three levels of agreement for ICCs were determined based on Rosner's guidelines: Poor (<0.4), fair-to-good ($0.4-0.75$) and excellent (≥ 0.75).^[54]

Responsiveness was assessed among participants reporting either improved or worsened health perception after the 2-week interval. Standardized effect sizes (SES) and the standardized response mean (SRM) were computed as mean changes of each SF-12 scale score divided by SD at baseline and SD of mean changes, respectively. Degree of responsiveness was defined as follows: trivial (<0.2), small (≥ 0.2 and <0.5), moderate (≥ 0.5 and <0.8), and large (≥ 0.8).^[55,56]

All statistical analyses were performed using SPSS IBM version 23. A *p*-value < 0.05 was considered statistically significant.

RESULTS

Participants' General Characteristics

Of the 1200 participants, the mean age was 42.7 years ($SD = 13.7$). Most were women (53.3%), married (63.7%), and had finished high school education (42.9%). Of the participants, 356 reported underlying diseases. Moreover, no missing values from this dataset were detected.

Item Description

Table 1 shows that PF showed the highest mean score (87.60 ± 22.72) and GH the lowest (63.64 ± 21.79). Furthermore, a ceiling effect of more than 30% was reported in PF, RP, BP, SF, and RE, while only 10 participants (0.83%) reported full health across all eight scales. The floor effect was less than 2% for PF, RP, BP, GH, and VT; however, an overall floor effect was not found. Notably, no items had missing and out of range values.

Item-scale Correlation

Table 2 displays construct validity between item-scale correlation and component summaries. All items correlated with their hypothesized scales, that is, with Spearman's rho correlation coefficients from 0.78 to 1.00. Correlations between the eight scales and their component summaries showed that the PCS correlated more with PF, RP, BP, and GH ($r = 0.65-0.79$, $P < 0.01$), while the MCS correlated more with SF, RE, VT, and MH ($r = 0.65-0.71$, $P < 0.01$). Notably, the highest correlation coefficients were found in these pairs: PF₂/PCS ($r = 0.79$, $P < 0.01$) and MH₂/MCS ($r = 0.71$, $P < 0.01$). Regarding discriminant validity, all items correlated more with their hypothesized scales and component summaries than their counterparts.

Table 3 presents the correlation coefficients between the SF-12v2 items and 3L/5L dimensions: MO had a moderate correlation with the PF₁/PF₂ ($r = -0.33/-0.42$ for the 3L, $r = -0.36/-0.45$ for the 5L, all $P < 0.01$); BP had moderate correlation with BP₁ ($r = -0.33$ for the 3L, $r = -0.35$ for the 5L, all $P < 0.01$); and AD had moderate correlation with MH₂ ($r = -0.32$ for the 3L, $r = -0.35$ for the 5L, all $P < 0.01$).

Known-group Validity

Table 4 presents Thai SF-12v2 scale scores by sociodemographic and health-related factors. Results supported most hypothesized associations. As expected, women reported significantly lower SF-12 scores on all scales than did men. The elderly (≥ 60 years) had significantly lower scores than other age groups on most scales except VT and MH. Highly educated participants (\geq bachelor's degree) reported the significantly highest scale scores in PF, RP, and GH. Participants with high incomes ($\geq 30,000$ baht/month) tended to report significantly higher scores than their counterparts except for PF and VT. Participants suffering from diseases reported significantly lower scores than healthy participants, except for VT. However,

Table 1: Item description of the Thai SF-12v2 items

SF-12v2 item (scale)	Response value frequencies, n (%)				
	1	2	3	4	5
Limitations in moderate physical activities (PF ₁)	27 (2.25)	221 (18.42)	952 (79.33)	-	-
Limitations in climbing several flight of stairs (PF ₂)	41 (3.42)	238 (19.83)	921 (76.75)	-	-
Accomplished less due to physical problems (RP ₁)	5 (0.42)	36 (3.00)	184 (15.33)	261 (21.75)	714 (59.50)
Limited in kind of work or activities due to physical problems (RP ₂)	3 (0.25)	31 (2.58)	181 (15.08)	229 (19.08)	756 (63.00)
Pain interference with work inside or outside home (BP)*	4 (0.33)	21 (1.75)	160 (13.33)	399 (33.25)	616 (51.33)
Health rating in general (GH)*	7 (0.58)	190 (15.83)	587 (48.92)	347 (28.92)	69 (5.75)
Having a lot of energy (VT)*	3 (0.25)	66 (5.50)	401 (33.42)	518 (43.17)	212 (17.67)
Physical or emotional interference with social activities (SF)	0 (0.00)	18 (1.50)	195 (16.25)	334 (27.83)	653 (54.42)
Accomplished less due to emotional problems (RE ₁)	2 (0.17)	29 (2.42)	197 (16.42)	270 (22.50)	702 (58.50)
No carefulness in work or activities due to emotional problems (RE ₂)	2 (0.17)	37 (3.08)	202 (16.83)	258 (21.50)	701 (58.42)
Feeling calm and peaceful (MH ₁)*	4 (33.33)	47 (3.92)	320 (26.67)	626 (52.17)	203 (16.92)
Feeling downhearted and blue (MH ₂)	3 (0.25)	20 (1.67)	185 (15.42)	376 (31.33)	616 (51.33)
SF-12v2 scales	1	2	3	4	5
	Mean	SD	Range	% Ceiling	%Floor
PF	87.60	22.72	0-100	72.6	1.5
RP	84.86	20.84	0-100	57.1	0.1
BP	83.38	19.92	0-100	51.3	0.3
GH	63.64	21.79	0-100	5.8	0.6
VT	68.13	20.59	0-100	17.7	0.3
SF	83.79	20.06	25-100	54.4	0.0
RE	83.96	20.78	12.5-100	54.8	0.0
MH	76.66	15.74	12.5-100	12.3	0.0

*Item recoded; thus, higher scores indicate a better health condition

Response options for PF: 1=Yes, limited a lot, 2=Yes, limited a little, 3=No, not limited at all

Response options for GH: 1=Excellence, 2=Very good, 3=Good, 4=Fair, 5=Poor

Response options for RP, RE, VT, SF, MH: 1=All of the time, 2=Most of the time, 3=Some of the time, 4=A little of the time, 5=None of the time

Response options for BP: 1=Not at all, 2=A little bit, 3=Moderately, 4=Quite a bit, 5=Extremely

PF: Physical functioning, RP: Role limitation due to physical problems, BP: Bodily pain, GH: General health perceptions, SF: Social functioning, VT: Vitality, RE: Role limitations due to emotional problems, MH: Mental health

smokers reported significantly higher PF scale scores than non-smokers. Similarly, drinkers tended to report significantly higher scale scores in PF, GH, and SF than non-drinkers.

EFA

Table 5 displays the results of factor analysis. A two-factor structure (physical and mental health) was identified with Eigenvalues over 1.00 accounting for 64% of the total variance from the SF-12v2 questionnaire. The results demonstrated that the PF, RP, BP, GH, and RE subscales were loaded higher on the physical health component and that the VT and MH subscales were loaded higher on the mental health component. SF was loaded on the physical and mental health components.

Reliability

Internal consistency reliabilities were 0.82 for the PF, 0.92 for the RP, and 0.92 for the RE, while the MH scale was 0.43. For all SF-12v2 items, reliability was 0.89. For test-retest reliability, 236 of 400 participants (59%) reported no health

status change between initial and second assessments. As Table 6 shows, ICCs ranged from 0.49 to 0.64. Notably, RE had the highest reproducibility, with ICCs of 0.64 (95% CI: 0.53–0.72), while VT had the lowest, with ICCs of 0.49 (95% CI: 0.35–0.61). Moreover, ICCs of PCS and MCS were 0.52 (95% CI: 0.38–0.63) and 0.66 (95% CI: 0.56–0.74), respectively.

Responsiveness

Table 6 also shows all SF-12v2 scales' responsiveness. At 2 weeks after the initial assessment, 164 samples (41%) indicated health status changes (improved or worsened). Reporting health improvement were 136 samples, with positive mean change scores of all SF-12v2 scales except for RP and BP ($P > 0.05$). In this group, all SF-12v2 scales' SES values ranged from 0.05 to 0.39, while their SRM values ranged from 0.04 to 0.34. Conversely, only 28 participants reported worsened health status with negative mean change scores from baseline. For all SF-12v2 scales, SES and SRM ranges were 0.21–0.92 and 0.17–0.68, respectively.

Table 2: Item-scale correlation for the SF-12v2 scales and two component summary scales

	PF	RP	BP	GH	SF	RE	VT	MH	PCS	MCS
PF										
PF ₁	0.88	0.54	0.43	0.36	0.30	0.47	0.19	0.19	0.76	0.17
PF ₂	0.92	0.55	0.42	0.38	0.28	0.45	0.20	0.17	0.79	0.15
RP										
RP ₁	0.59	0.97	0.58	0.42	0.44	0.70	0.26	0.30	0.73	0.41
RP ₂	0.58	0.94	0.57	0.40	0.44	0.70	0.27	0.30	0.72	0.41
BP										
BP ₁	0.46	0.59	1.00	0.39	0.54	0.58	0.35	0.34	0.73	0.42
GH										
GH ₁	0.40	0.43	0.39	1.00	0.26	0.42	0.24	0.27	0.65	0.29
SF										
SF ₁	0.30	0.45	0.54	0.26	1.00	0.53	0.33	0.42	0.36	0.67
RE										
RE ₁	0.46	0.70	0.56	0.40	0.49	0.96	0.27	0.40	0.47	0.67
RE ₂	0.49	0.71	0.56	0.40	0.53	0.96	0.26	0.40	0.49	0.66
VT										
VT ₁	0.21	0.27	0.35	0.24	0.33	0.27	1.00	0.58	0.25	0.66
MH										
MH ₁	0.10	0.18	0.18	0.19	0.22	0.23	0.54	0.78	0.02	0.65
MH ₂	0.23	0.33	0.37	0.25	0.46	0.43	0.40	0.79	0.17	0.71

All values were correlation coefficients and significant at $P < 0.01$. PF: Physical functioning, RP: Role limitation due to physical problems, BP: Bodily pain, GH: General health perceptions, SF: Social functioning, VT: Vitality, RE: Role limitations due to emotional problems, MH: Mental health, PCS: Physical component score, MCS: Mental component score

SF-12v2 Scale Scores by Age and Gender

Table 7 presents SF-12v2 scale scores by age and gender. Similar to known-group results, older samples reported lower scale scores than youngsters in most scales and the two component summaries except for MH and MCS. Moreover, female participants tended to report lower SF-12v2 scale scores than males in PF, SF, and RE scales. Notably, scales related to physical health (PF, RP, and BP) had higher scores than those related to mental health (SF and VT).

DISCUSSION

This is the first study to investigate the SF-12v2's psychometric properties and to demonstrate its scale scores by age and gender in the general Thai population. Overall results supported the SF-12v2's application in monitoring and measuring population health status. Moreover, because this study revealed no missing and out of range values, the collected data had satisfactory quality.

Due to both ceiling and floor effects of less than 15%,^[52] most SF-12v2 scales exceeded the acceptable percentage, except for GH and MH. However, for all SF-12v2 scales, floor effects were within the acceptable percentage. The high ceiling effect was due to exclusion of participants with acute or life threatening disease and cognitive impairment or disability, so the majority of participants (70%) were healthy.

Similar to many previous studies,^[37,38,51,57,58] results showed that all item-scale correlation coefficients are higher than those

of items and other scales determining good convergent and discriminant validity. Moreover, item-component correlations were consistent with hypothesized correlations defined by previous studies.^[27,29,59] With regard to the correlations between the SF-12v2 items and EQ-5D dimensions, most of the hypothesized correlations were confirmed except for MH₁. MH₁ was the item dealing with the feeling of calm and peacefulness, which has a different meaning from the AD dimension of the EQ-5D questionnaire. Therefore, weak or no association was detected from this correlation analysis. Nevertheless, this study demonstrated the moderate correlations between the MH scores and AD dimension of the 5L ($r = -0.30$ for the 5 L), indicating that the MH scale had a good convergent validity with the AD dimension of the EQ-5D questionnaire. However, the strength of correlations was not as strong as expected. Different recall periods between these two instruments may account for these discrepancies because the EQ-5D asked the respondents to rate their current health status, whereas the SF-12v2 asked respondents to rate their health status in the past 4 weeks.

Known-group results showed that SF-12v2 scales could distinguish HRQoL scores among sample subgroups based on sociodemographic status. Hypothesized associations were confirmed because female, older, and poorly educated participants and those with lower incomes and chronic diseases reported lower SF-12v2 scale scores than their counterparts. These results are similar to those reported in previous studies.^[37,38,51,60] Moreover, this study showed that increasing age had a positive impact on MH and MCS, a result also

Table 3: Correlations between SF-12v2 items and EQ-5D-3L (3L) and EQ-5D-5L (5L) dimensions

Dimensions	MO		SC		UA		PD		AD	
	3L	5L	3L	5L	3L	5L	3L	5L	3L	5L
PF										
PF ₁	-0.33	-0.36	-0.16	-0.17	-0.29	-0.28	-0.33	-0.31	-0.28	-0.25
PF ₂	-0.42	-0.45	-0.16	-0.19	-0.29	-0.29	-0.36	-0.36	-0.24	-0.27
RP										
RP ₁	-0.27	-0.30	-0.12	-0.13	-0.26	-0.28	-0.31	-0.31	-0.26	-0.27
RP ₂	-0.28	-0.30	-0.12	-0.14	-0.29	-0.30	-0.32	-0.32	-0.25	-0.26
BP										
BP ₁ *	-0.22	-0.26	-0.11	-0.13	-0.21	-0.22	-0.33	-0.35	-0.24	-0.27
GH										
GH ₁ *	-0.24	-0.29	-0.09	-0.08	-0.20	-0.19	-0.29	-0.30	-0.25	-0.24
SF										
SF ₁	-0.14	-0.15	-0.09	-0.10	-0.15	-0.15	-0.20	-0.17	-0.26	-0.25
RE										
RE ₁	-0.20	-0.21	-0.09	-0.10	-0.23	-0.22	-0.24	-0.21	-0.30	-0.33
RE ₂	-0.20	-0.21	-0.08	-0.10	-0.26	-0.24	-0.25	-0.25	-0.31	-0.33
VT										
VT ₁ *	-0.14	-0.16	-0.01	-0.01	-0.07	-0.07	-0.19	-0.13	-0.14	-0.16
MH										
MH ₁ *	-0.11	-0.13	-0.03	-0.02	-0.03	-0.01	-0.11	-0.09	-0.08	-0.09
MH ₂	-0.16	-0.17	-0.02	-0.01	-0.02	-0.08	-0.17	-0.15	-0.32	-0.35

*Item recoded; thus, higher scores indicate a better health condition. All values were correlation coefficients and were significant at $P < 0.01$. PF: Physical functioning, RP: Role limitation due to physical problems, BP: Bodily pain, GH: General health perceptions, SF: Social functioning, VT: Vitality, RE: Role limitations due to emotional problems, MH: Mental health, MO: Mobility, SC: Self-care, UA: Usual activities, PD: Pain/discomfort, AD: Anxiety/depression

consistent with previous studies.^[29,61-64] A possible explanation is that although older participants had some chronic diseases, their conditions were quite stable, and none suffered from mental disorders. However, this association deserves future investigation with patient groups. Furthermore, this study revealed that smokers and drinkers had higher PF scores than their counterparts. These results were similar to a previous Thai SF-36v2 study^[29] that reported about the failure to draw an association between tobacco and alcohol use and SF-36v2 item scores. I reasoned that lower number of smokers and drinkers than their counterparts and most of the healthy samples (70%) were recruited into this study. Therefore, this finding should be investigated to determine whether there is a possibility of an unexpected association in future research. Moreover, using multi-group analysis to compare the perceptions of the Thai SF-12v2 items between healthier and less healthy respondents by using structural equation modeling is recommended for further research.

Similar to what the SF-12v2 studies in both the general population and the clinical area have yielded,^[27,39,51,65,66] in this study, the principal component analysis with varimax rotation yielded a two-factor structure of the Thai SF-12v2 questionnaire. Most of the hypothesized SF-12v2 subscales and factor loadings were confirmed, except for SF and RE, which were loaded higher on the physical health component. Similar to the findings of Seon-Ha Kim's study of the general Korean population,^[37] in this study, the findings of the EFA

demonstrated that RE and SF items were loaded into the same factor as RP although this study yielded a three-factor structure. This may be because the respondents' perceptions for those two items may differ from those of the remaining populations^[38,39] wherein those two items were loaded on mental health problems. Unlike the findings from item-scale correlation, this study revealed that RE and SF were highly correlated with MCS. I reasoned that the MCS score is generally derived from the summation of the RE and SF scale scores, thereby resulting in high correlations between those scales observed. Nevertheless, these two scales had high correlations with other scales loaded on the physical component in the factor analysis. A possible explanation is that most of the samples were healthy (70%) and those with physical health problems were primarily recruited, as these recruited samples might perceive these two scales mainly asking them about their physical activities similar to PH and RP scales, although both RE and SF asked about the emotional health problems affecting their role-functioning and social activities. Therefore, both RE and SF scales should be performed for cognitive debriefing whether general Thai samples understand what those two scales mean and reinvestigated for psychometric properties in general Thai samples with varied health conditions.

Internal consistency reliability was satisfactory for PF, RP, and RE scales because their Cronbach's alpha exceeded 0.7; however, it was 0.43 for the MH scale, similar to previous studies.^[37,59] MH₁ (feeling calm and peaceful) was negatively

Table 4: Known-group validity of the SF-12v2 scale scores based on sociodemographic and health-related factors

Participants characteristics	Sample size (n)	SF-12v2 scale scores, mean (SD)							
		PF	RP	BP	GH	VT	SF	RE	MH
Gender									
Male	560	90.80 (19.90)	86.85 (19.83)	84.73 (19.41)	65.59 (21.38)	69.29 (20.30)	85.36 (19.32)	85.92 (19.68)	77.32 (15.13)
Female	640	84.80 (24.60)	83.13 (21.55)	82.19 (20.29)	61.93 (22.02)	67.11 (20.80)	82.42 (20.59)	82.25 (21.57)	76.07 (16.24)
<i>P</i> -value ^a		<0.01	<0.01	0.03	0.04	0.07	0.01	<0.01	0.17
Age									
20–29	258	94.38 (15.36)	89.10 (18.95)	87.50 (17.95)	69.50 (20.69)	69.38 (20.86)	84.88 (19.46)	86.39 (19.99)	75.68 (14.94)
30–39	263	91.35 (19.26)	89.02 (17.13)	86.88 (17.31)	67.57 (18.92)	67.68 (19.93)	85.08 (18.87)	86.64 (18.64)	76.43 (15.42)
40–49	268	91.04 (18.33)	87.27 (19.25)	86.10 (17.18)	65.04 (20.93)	70.15 (20.18)	85.35 (19.77)	85.96 (19.31)	78.45 (15.78)
50–59	252	80.85 (27.20)	78.82 (24.00)	78.17 (23.08)	57.04 (22.53)	65.67 (21.46)	81.45 (21.07)	80.06 (22.85)	75.55 (17.04)
≥60 years	159	75.31 (28.82)	76.65 (22.15)	74.53 (21.51)	55.72 (23.45)	67.30 (20.27)	80.97 (21.32)	78.38 (22.53)	77.36 (15.16)
<i>P</i> -value ^b		<0.01	<0.01	<0.01	<0.01	0.11	0.04	<0.01	0.19
Education level									
None or primary school	414	83.03 (25.75)	82.19 (22.53)	79.59 (21.66)	59.50 (22.36)	66.49 (20.40)	82.43 (20.53)	82.19 (21.69)	76.48 (16.43)
Secondary school	512	89.36 (21.08)	85.33 (20.07)	83.89 (18.88)	65.68 (21.34)	67.87 (21.11)	83.50 (20.18)	84.30 (20.57)	76.34 (15.86)
College degree	102	89.46 (18.43)	87.99 (18.36)	89.71 (14.65)	63.38 (19.72)	71.57 (18.96)	86.76 (19.17)	86.40 (18.87)	76.47 (14.64)
≥Bachelor's degree	172	92.30 (20.11)	88.08 (19.58)	87.21 (19.59)	67.65 (21.48)	70.78 (20.06)	86.19 (18.78)	85.76 (20.07)	78.13 (14.25)
<i>P</i> -value ^b		<0.01	<0.01	<0.01	<0.01	0.04	0.08	0.12	0.62
Household income (baht/month)									
Low: 0–30,000 baht/month	1111	87.38 (22.90)	84.48 (21.06)	83.01 (20.13)	63.15 (21.90)	67.96 (20.83)	83.44 (20.25)	83.39 (21.07)	76.27 (15.81)
High: ≥30,000 baht/month	89	90.45 (20.12)	89.61 (17.30)	87.92 (16.48)	69.66 (19.54)	70.22 (17.22)	88.20 (16.89)	91.01 (15.19)	81.46 (13.99)
<i>P</i> -value ^a		0.17	<0.01	<0.01	<0.01	0.24	0.01	<0.01	<0.01
Smoking									
Non-smokers	865	86.79 (89.70)	84.88 (20.97)	83.58 (19.82)	63.37 (21.95)	68.44 (20.89)	83.27 (20.12)	84.28 (20.78)	76.71 (15.73)
Smokers	335	89.70 (20.25)	84.81 (20.55)	82.84 (20.20)	64.33 (21.40)	67.31 (19.79)	85.15 (19.86)	83.13 (20.79)	76.53 (15.78)
<i>P</i> -value ^a		0.03	0.96	0.56	0.50	0.40	0.15	0.39	0.86
Alcohol									
Non-drinkers	748	86.13 (24.25)	83.99 (20.96)	82.82 (20.14)	62.55 (22.37)	67.68 (20.63)	82.65 (20.47)	83.24 (21.23)	77.22 (15.48)
Drinkers	452	90.04 (19.71)	86.31 (20.58)	84.29 (19.54)	65.43 (20.70)	68.86 (20.52)	85.67 (19.22)	85.15 (19.99)	75.72 (16.12)
<i>P</i> -value ^a		<0.01	0.06	0.22	0.03	0.34	0.01	0.12	0.11
Current disease									
No	844	91.23 (18.99)	87.60 (19.20)	86.20 (17.87)	67.06 (20.22)	68.66 (20.44)	85.01 (19.42)	85.89 (19.78)	77.58 (15.26)
Yes	356	79.00 (27.95)	78.37 (23.06)	76.69 (22.77)	55.52 (23.21)	66.85 (20.90)	80.90 (21.24)	79.39 (22.35)	74.47 (16.63)
<i>P</i> -value ^a		<0.01	<0.01	<0.01	<0.01	0.17	<0.01	<0.01	<0.01

^aThe student's *t*-test. ^bAnalysis of variance. PF: Physical functioning, RP: Role limitation due to physical problems, BP: Bodily pain, GH: General health perceptions, SF: Social functioning, VT: Vitality, RE: Role limitations due to emotional problems, MH: Mental health

worded, so this might have contributed to inconsistent response to two items (MH₁/MH₂) if participants completed them without understanding their meaning. However, results showed fair-to-good test-retest reliability for all of SF-12v2 and the two component summaries (ICCs: 0.49–0.66). Compared with studies by Saban *et al.*^[67] and Cheak-Zamora *et al.*,^[68] this study had PCS and MCS's ICCs slightly higher than Saban *et al.* (PCS = 0.44, MCS = 0.47), while they were lower than those reported in Cheak-Zamora *et al.* (PCS = 0.78, MCS = 0.60). This might be due to longer intervals; Cheak-Zamora *et al.* and Saban *et al.* employed 1-year and 3-month periods, respectively, while this study used a 2-week period for retest reliability. Previous evidence has suggested that a period of over 2 weeks is considered quite long, resulting in samples' changed health status, in turn affecting results of test-retest reliability.^[69,70]

Regarding responsiveness, the improved group's mean changes were slightly less than the worsened group's across all

Table 5: Thai SF-12v2 structure derived from principal component analysis with varimax rotation

Scales	Factor loadings	
	Physical	Mental
PF	0.807	0.026
RP	0.840	0.203
BP	0.731	0.327
GH	0.630	0.151
VT	0.125	0.845
SF	0.508	0.490
RE	0.776	0.314
MH	0.207	0.853

PF: Physical functioning, RP: Role limitation due to physical problems, BP: Bodily pain, GH: General health perceptions, SF: Social functioning, VT: Vitality, RE: Role limitations due to emotional problems, MH: Mental health

SF-12v2 scales and the two component summaries. Moreover, the mean change for the improved group had negative values for some SF-12v2 scales—RP and BP; this unexpected finding was not statistically significant, however. Notably, the improved group's SES and SRM values were less than the worsened group's, except for the SF, VT, and MCS, ranging from trivial to small; those of the worsened group were diverse, ranging from small to large. This finding indicated that the Thai SF-12v2 was more sensitive to health deterioration than to health improvement, consistent with a previous study of hypertensive patients.^[71] Nevertheless, future studies should reinvestigate with various patient groups.

Although this study's sample size was not large enough to generate norm-based scores for the general Thai population,^[72] SF-12v2 scale scores by age and gender groups could perhaps be used to project the SF-12v2's normative scores. Compared to the previous study,^[73] it followed a similar trend of decreasing SF-12v2 scale scores with advancing age, except for RE, MH, and MCS; it also reported that women had lower scores than men. Nevertheless, comparison of Thai SF-12v2 scale scores with varied cultural populations has not yet been achieved because the general Thai population's norm-based scores are not yet available. Therefore, to compare scale scores with other specific populations, a Thai population norm for the SF-12v2 should be established with a larger sample.

This study has some limitations to be addressed. First, although it employed a multistage random sampling to select representative areas, all samples were selected for convenience—a non-probability sampling method that might cause sampling bias. To minimize sampling bias, however, all samples were selected according to age and gender to represent the national Thai census. Second, this study did not establish population norm-based scores or assess some psychometric properties, including content validity and criterion validity, so these should be reinvestigated both in patient groups and in the general Thai sample with varied health condition. Third,

Table 6: Test-retest reliability and responsiveness of the SF-12v2 scales and two component summary scores

SF-12v2 scales	Reliability Intraclass correlation coefficients (95% CI)	Responsiveness							
		Improved				Worsened			
		Baseline scores (SD)	Mean change (SD)	SES	SRM	Baseline scores (SD)	Mean change (SD)	SES	SRM
PF	0.50 (0.35–0.61)	87.13 (23.96)	5.37** (23.73)	0.22	0.23	82.14 (25.33)	−9.82 (31.43)	0.39	0.31
RP	0.50 (0.36–0.61)	83.46 (21.56)	−1.02 (26.28)	0.05	0.04	80.36 (20.25)	−10.71* (19.75)	0.53	0.54
BP	0.53 (0.39–0.63)	84.56 (19.50)	−0.93 (25.99)	0.05	0.04	82.14 (17.82)	−15.18 (25.77)	0.85	0.59
GH	0.53 (0.39–0.63)	62.76 (22.25)	8.67** (25.17)	0.39	0.34	54.11 (21.35)	−19.64 (29.41)	0.92	0.68
SF	0.57 (0.45–0.67)	79.41 (22.54)	7.22 (27.30)	0.32	0.26	82.14 (24.40)	−5.36* (24.87)	0.22	0.22
RE	0.64 (0.53–0.72)	80.79 (23.12)	2.69 (26.16)	0.12	0.10	77.68 (17.47)	−5.80* (17.83)	0.33	0.33
VT	0.49 (0.35–0.61)	68.57 (18.76)	4.44* (24.60)	0.24	0.18	61.61 (20.95)	−4.46 (25.51)	0.21	0.17
MH	0.52 (0.38–0.63)	75.18 (15.74)	2.78** (17.05)	0.18	0.16	72.32 (17.79)	−6.70* (17.83)	0.38	0.38
PCS	0.52 (0.38–0.63)	52.22 (6.82)	0.89 (7.37)	0.13	0.12	50.50 (6.66)	−5.17** (7.94)	0.78	0.65
MCS	0.66 (0.56–0.74)	50.90 (7.51)	1.84* (8.79)	0.25	0.21	49.95 (7.82)	−1.65 (8.31)	0.21	0.20

* $P < 0.05$, ** $P < 0.01$. PF: Physical functioning, RP: Role limitation due to physical problems, BP: Bodily pain, GH: General health perceptions, SF: Social functioning, VT: Vitality, RE: Role limitations due to emotional problems, MH: Mental health, PCS: Physical component score, MCS: Mental component score, SES: Standardized effect sizes, SRM: Standardized response mean

Table 7: Scores of the SF-12v2 scales and two component summary scores stratified by age and gender (mean±SD)

Age	Gender	PF	RP	BP	GH	VT	SF	RE	MH	PCS	MCS
20–29	Female (n=127)	93.30 (16.50)	89.37 (18.10)	88.39 (16.61)	69.57 (20.67)	68.31 (20.28)	82.87 (20.82)	84.55 (21.16)	74.02 (15.95)	54.67 (4.62)	50.76 (7.67)
	Male (n=131)	95.42 (14.14)	88.84 (19.81)	86.64 (19.19)	69.43 (20.78)	70.42 (21.44)	86.83 (17.90)	88.17 (18.69)	77.29 (13.76)	54.07 (5.02)	52.64 (6.35)
30–39	Female (n=138)	88.95 (21.58)	87.32 (18.00)	86.41 (18.65)	67.68 (19.25)	67.03 (21.14)	84.96 (19.23)	85.69 (19.26)	75.72 (16.60)	53.14 (5.56)	51.90 (6.91)
	Male (n=125)	94.00 (16.00)	90.90 (15.98)	87.40 (15.76)	67.44 (18.63)	68.40 (18.55)	85.20 (18.54)	87.70 (17.96)	77.20 (14.03)	54.04 (4.68)	52.20 (6.67)
40–49	Female (n=144)	89.06 (20.21)	86.46 (19.36)	86.11 (15.86)	63.37 (19.59)	70.49 (19.31)	85.24 (19.25)	85.33 (19.46)	79.17 (14.86)	52.31 (5.64)	53.02 (6.56)
	Male (n=124)	93.35 (15.65)	88.21 (19.15)	86.09 (18.67)	66.98 (22.30)	69.76 (21.21)	85.48 (20.44)	86.69 (19.18)	77.62 (16.81)	53.49 (5.68)	52.47 (7.87)
50–59	Female (n=141)	76.60 (29.00)	75.53 (25.13)	74.65 (24.09)	53.83 (22.85)	62.77 (22.28)	79.61 (21.26)	77.84 (23.88)	74.20 (18.07)	48.12 (8.27)	51.13 (8.45)
	Male (n=111)	86.26 (23.77)	83.00 (21.90)	82.66 (21.00)	61.13 (21.53)	69.37 (19.85)	83.78 (20.68)	82.88 (21.25)	77.25 (15.56)	51.29 (6.54)	52.39 (6.50)
≥60	Female (n=90)	72.50 (29.04)	74.44 (22.56)	72.50 (20.54)	52.72 (23.26)	66.94 (20.12)	77.78 (22.31)	75.69 (22.67)	77.50 (14.54)	46.98 (7.23)	52.11 (7.16)
	Male (n=69)	78.99 (28.32)	79.53 (21.44)	77.17 (22.58)	59.64 (23.28)	67.75 (20.61)	85.14 (19.33)	81.88 (22.02)	77.17 (16.04)	49.20 (7.38)	53.05 (6.73)
Total	Female (n=640)	84.80 (24.60)	83.13 (21.55)	82.19 (20.29)	61.93 (22.02)	67.11 (20.80)	82.42 (20.59)	82.25 (21.57)	76.07 (16.24)	51.29 (6.95)	51.79 (7.41)
	Male (n=560)	90.80 (19.90)	86.85 (19.83)	84.73 (19.41)	65.59 (21.38)	69.29 (20.30)	85.36 (19.32)	85.92 (19.68)	77.32 (15.13)	52.78 (5.98)	52.51 (6.84)

PF: Physical functioning, RP: Role limitation due to physical, BP: Bodily pain, GH: General health, VT: Vitality, SF: Social functioning, RE: Role limitation due to emotion, MH: Mental health, PCS: Physical component score, MCS: Mental component score

SF-12v2 requires permission for use from OPTUM, and the users had to pay the license fees. The response options also vary among the eight SF-12v2 scales. This variation might cause confusion and a major obstacle for the SF-12v2 users. Fourth, this study did not perform confirmatory factor analysis to verify the factor structure of the Thai SF-12v2 questionnaire and cognitive debriefing, especially for RE and SF scales. Thus, these scales should be reinvestigated in samples with varied health conditions in the future study.

CONCLUSION

These preliminary results suggest that the SF-12v2 is valid and reliable, and it can be an alternative to the Thai SF-36v2 for HRQoL measurement. Future studies are recommended to establish the population norm, investigate some other psychometric properties, including content, concurrent validities, and construct validity from confirmatory factor analysis and cognitive debriefing especially for RE and SF scales, in patient groups and in general Thai sample with varied health conditions, and examine whether the perception of the Thai SF-12v2 items differs between healthy and less healthy respondents.

ACKNOWLEDGMENTS

This work was financially supported by a Research Grant of Burapha University through the National Research Council of Thailand (Grant No. Rx2/2562). However, the results and opinions in this report have not been endorsed by this funding agency or elsewhere. I would like to express my gratitude to all general Thai samples from five provinces across Thailand.

I would also like to thank all trained interviewers for their assistance with the interview process.

REFERENCES

1. Weldring T, Smith SM. Patient-reported outcomes (PROs) and patient-reported outcome measures (PROMs). *Health Serv Insights* 2013;6:61-8.
2. Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. *Cochrane Handbook for Systematic Reviews of Interventions*. New York, United States: John Wiley & Sons; 2019.
3. Badia X, Herdman M. The importance of health-related quality-of-life data in determining the value of drug therapy. *Clin Ther* 2001;23:168-75.
4. Hwang JS, Wang JD. Integrating health profile with survival for quality of life assessment. *Qual Life Res* 2004;13:1-10.
5. Osoba D. What has been learned from measuring health-related quality of life in clinical oncology. *Eur J Cancer* 1999;35:1565-70.
6. Lin XJ, Lin IM, Fan SY. Methodological issues in measuring health-related quality of life. *Tzu Chi Med J* 2013;25:8-12.
7. Song T, Ding YW, Sun Y, He YN, Qi DJ, Wu Y, et al. A population-based study on health-related quality of life among urban community residents in Shenyang, Northeast of China. *BMC Public Health* 2015;15:921.
8. Burstrom K, Johannesson M, Diderichsen F. Swedish population health-related quality of life results using the EQ-5D. *Qual Life Res* 2001;10:621-35.
9. Bullinger M, Quitmann J. Quality of life as patient-reported outcomes: Principles of assessment. *Dialogues Clin Neurosci* 2014;16:137-45.
10. Saffari M, Lin CY, O'Garro K, Koenig HG, Sanaeinasab H, Pakpour AH. Psychometric properties of persian diabetes-mellitus specific quality of life (DMQoL) questionnaire in a population-based sample of Iranians. *Int J Diabetes Dev Ctries*

- 2019;39:218-27.
11. Pakpour AH, Chen CY, Lin CY, Strong C, Tsai MC, Lin YC. The relationship between children's overweight and quality of life: A comparison of Sizing Me Up, PedsQL and Kid-KINDL. *Int J Clin Health Psychol* 2019;19:49-56.
 12. Lin CY, Hwang JS, Wang WC, Lai WW, Su WC, Wu TY, *et al.* Psychometric evaluation of the WHOQOL-BREF, Taiwan version, across five kinds of Taiwanese cancer survivors: Rasch analysis and confirmatory factor analysis. *J Formos Med Assoc* 2019;118:215-22.
 13. Andresen EM, Meyers AR. Health-related quality of life outcomes measures. *Arch Phys Med Rehabil* 2000;81 Suppl 12:S30-45.
 14. Wang R, Wu C, Zhao Y, Yan X, Ma X, Wu M, *et al.* Health related quality of life measured by SF-36: A population-based study in Shanghai, China. *BMC Public Health* 2008;8:292.
 15. Loge JH, Kaasa S. Short form 36 (SF-36) health survey: Normative data from the general Norwegian population. *Scand J Soc Med* 1998;26:250-8.
 16. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med* 1993;118:622-9.
 17. Liem YS, Bosch JL, Arends LR, Heijnenbroek-Kal MH, Hunink MG. Quality of life assessed with the medical outcomes study short form 36-item health survey of patients on renal replacement therapy: A systematic review and meta-analysis. *Value Health* 2007;10:390-7.
 18. Ware JE Jr., Gandek B. Methods for testing data quality, scaling assumptions, and reliability: The IQOLA project approach. *International quality of life assessment. J Clin Epidemiol* 1998;51:945-52.
 19. Anderson RT, Aaronson NK, Wilkin D. Critical review of the international assessments of health-related quality of life. *Qual Life Res* 1993;2:369-95.
 20. Coons SJ, Rao S, Keininger DL, Hays RD. A comparative review of generic quality-of-life instruments. *Pharmacoeconomics* 2000;17:13-35.
 21. Feroz AH, Islam MN, Ten Klooster PM, Hasan M, Rasker JJ, Haq SA. The bengali short form-36 was acceptable, reliable, and valid in patients with rheumatoid arthritis. *J Clin Epidemiol* 2012;65:1227-35.
 22. Li L, Wang HM, Shen Y. Chinese SF-36 health survey: Translation, cultural adaptation, validation, and normalisation. *J Epidemiol Community Health* 2003;57:259-63.
 23. Tseng HM, Lu JR, Gandek B. Cultural issues in using the SF-36 health survey in Asia: Results from Taiwan. *Health Qual Life Outcomes* 2003;1:72.
 24. Fukuhara S, Bito S, Green J, Hsiao A, Kurokawa K. Translation, adaptation, and validation of the SF-36 health Survey for use in Japan. *J Clin Epidemiol* 1998;51:1037-44.
 25. Sabbah I, Drouby N, Sabbah S, Retel-Rude N, Mercier M. Quality of life in rural and urban populations in Lebanon using SF-36 health survey. *Health Qual Life Outcomes* 2003;1:30.
 26. Lam CL, Gandek B, Ren XS, Chan MS. Tests of scaling assumptions and construct validity of the Chinese (HK) version of the SF-36 Health Survey. *J Clin Epidemiol* 1998;51:1139-47.
 27. Montazeri A, Goshtasebi A, Vahdaninia M, Gandek B. The short form health survey (SF-36): Translation and validation study of the Iranian version. *Qual Life Res* 2005;14:875-82.
 28. Leurmarnkul W, Meetam P. Properties testing of the retranslated SF36 (Thai version). *Thai J Pharm Sci* 2005;29:69-88.
 29. Kangwanrattanakul K, Auamnoy T. Psychometric testing of the health-related quality of life measurement, SF-36v2, in the general population of Thailand. *Expert Rev Pharmacoecon Outcomes Res* 2019;19:313-20.
 30. Su CT, Ng HS, Yang AL, Lin CY. Psychometric evaluation of the short form 36 health survey (SF-36) and the World Health Organization quality of life scale brief version (WHOQOL-BREF) for patients with schizophrenia. *Psychol Assess* 2014;26:980-9.
 31. Gandek B, Ware JE, Aaronson NK, Apolone G, Bjorner JB, Brazier JE, *et al.* Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: Results from the IQOLA Project. *International Quality of Life Assessment. J Clin Epidemiol* 1998;51:1171-8.
 32. Turner-Bowker D, Hogue SJ. Short form 12 health survey (SF-12). In: Michalos AC, editor. *Encyclopedia of quality of life and well-being research*. Dordrecht, Netherlands: Springer; 2014. p. 5954-7.
 33. Ware J Jr., Kosinski M, Keller SD. A 12-item short-form health survey: Construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;34:220-33.
 34. Lam ET, Lam CL, Fong DY, Huang WW. Is the SF-12 version 2 health survey a valid and equivalent substitute for the SF-36 version 2 health survey for the Chinese? *J Eval Clin Pract* 2013;19:200-8.
 35. Ware J, Kosinski M, Turner-Bowker D, Gandek B. How to Score SF-12 Items. SF-12 v2: How to Score Version 2 of the SF-12 Health Survey; 2002. p. 29-38.
 36. Damásio BF, Andrade TF, Koller SH. Psychometric properties of the Brazilian 12-item short-form health survey version 2 (SF-12v2). *Paidéia (Ribeirão Preto)* 2015;25:29-37.
 37. Kim SH, Jo MW, Ahn J, Ock M, Shin S, Park J. Assessment of psychometric properties of the Korean SF-12 v2 in the general population. *BMC Public Health* 2014;14:1086.
 38. Montazeri A, Vahdaninia M, Mousavi SJ, Asadi-Lari M, Omidvari S, Tavousi M. The 12-item medical outcomes study short form health survey version 2.0 (SF-12v2): A population-based validation study from Tehran, Iran. *Health Qual Life Outcomes* 2011;9:12.
 39. Su SW, Wang D. The reliability and validity of short form-12 health survey version 2 for Chinese older adults. *Iran J Public Health* 2019;48:1014-24.
 40. Linde L, Sørensen J, Østergaard M, Hørslev-Petersen K, Rasmussen C, Jensen DV, *et al.* What factors influence the health status of patients with rheumatoid arthritis measured by the SF-12v2 health survey and the health assessment questionnaire? *J Rheumatol* 2009;36:2183-9.
 41. Brown TM, Lee WC, Joshi AV, Pashos CL. Health-related quality of life and productivity impact in haemophilia patients with inhibitors. *Haemophilia* 2009;15:911-7.
 42. Lee CE, Browell LM, Jones DL. Measuring health in patients with cervical and lumbosacral spinal disorders: Is the 12-item short-form health survey a valid alternative for the 36-item short-form health survey? *Arch Phys Med Rehabil* 2008;89:829-33.
 43. Yang M, Wallenstein G, Hagan M, Guo A, Chang J, Kornstein S. Burden of premenstrual dysphoric disorder on health-related quality of life. *J Womens Health* 2008;17:113-21.
 44. Sutton D, Raines DA. Health-related quality of life following a surgical weight loss intervention. *Appl Nurs Res* 2010;23:52-6.
 45. McBride O, Adamson G, Bunting BP, McCann S. Assessing the general health of diagnostic orphans using the short form health survey (SF-12v2): A latent variable modelling approach. *Alcohol Alcohol* 2009;44:67-76.
 46. Sloan RA, Sawada SS, Martin CK, Church T, Blair SN. Associations between cardiorespiratory fitness and health-related quality of life. *Health Qual Life Outcomes* 2009;7:47.
 47. Shah RM, Banahan BF 3rd, Holmes ER, Patel AS, Barnard M, Khanna R, *et al.* An evaluation of the psychometric properties of the sf-12v2 health survey among adults with hemophilia. *Health Qual Life Outcomes* 2018;16:229.
 48. Krittayaphong R, Bhuripanyo K, Raungratanaamporn O, Chotinaiwatarakul C, Chaowalit N, Punlee K, *et al.* Reliability of Thai version of SF-36 questionnaire for the evaluation of quality of life in cardiac patients. *J Med Assoc Outcomes* 2000;83 Suppl 2:S130-6.
 49. Jirarattanaphochai K, Jung S, Sumananont C, Saengnipanthkul S. Reliability of the medical outcomes study short-form survey

- version 2.0 (Thai version) for the evaluation of low back pain patients. *J Med Assoc Outcomes* 2005;88:1355-61.
50. Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. *Med Care* 2004;42:851-9.
 51. Kontodimopoulos N, Pappa E, Niakas D, Tountas Y. Validity of SF-12 summary scores in a Greek general population. *Health Qual Life Outcomes* 2007;5:55.
 52. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, *et al.* Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34-42.
 53. Cohen P. *Regression and Correlation. Statistic in Medicine.* Boston: Little Brown and Company; 1974.
 54. Rosner B. *Fundamentals of Biostatistics.* Ontario: Nelson Education; 2015.
 55. Cohen J. *Statistical Power Analysis for the Behavioral Sciences.* London, United Kingdom: Routledge; 2013.
 56. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. *Med Care* 1990;632-42.
 57. Larson CO, Schlundt D, Patel K, Beard K, Hargreaves M. Validity of the SF-12 for use in a low-income African American community-based research initiative (REACH 2010). *Prev Chronic Dis* 2008;5:A44.
 58. Jakobsson U, Westergren A, Lindskov S, Hagell P. Construct validity of the SF-12 in three different samples. *J Eval Clin Pract* 2012;18:560-6.
 59. Fong DY, Lam CL, Mak KK, Lo WS, Lai YK, Ho SY, *et al.* The short form-12 health survey was a valid instrument in Chinese adolescents. *J Clin Epidemiol* 2010;63:1020-9.
 60. Brazier JE, Harper R, Jones NM, O'Cathain A, Thomas KJ, Usherwood T, *et al.* Validating the SF-36 health survey questionnaire: New outcome measure for primary care. *BMJ* 1992;305:160-4.
 61. Su SW, Wang D. Health-related quality of life and related factors among elderly persons under different aged care models in Guangzhou, China: A cross-sectional study. *Qual Life Res* 2019;28:1293-303.
 62. Plantinga L, Lim SS, Bowling CB, Drenkard C. Association of age with health-related quality of life in a cohort of patients with systemic lupus erythematosus: The Georgians organized against lupus study. *Lupus Sci Med* 2016;3:e000161.
 63. Mishra GD, Hockey R, Dobson AJ. A comparison of SF-36 summary measures of physical and mental health for women across the life course. *Qual Life Res* 2014;23:1515-21.
 64. Garratt AM, Stavem K. Measurement properties and normative data for the Norwegian SF-36: Results from a general population survey. *Health Qual Life Outcomes* 2017;15:51.
 65. Cernin PA, Cresci K, Jankowski TB, Lichtenberg PA. Reliability and validity testing of the short-form health survey in a sample of community-dwelling African American older adults. *J Nurs Meas* 2010;18:49-59.
 66. Zhao LC, Yang Z, Hu X, He Y, Liu DP, Li NX. Equivalence of SF-12v2 and SF-36v2 in assessing health related quality of life in a general population in Chengdu. *J Sichuan Univ Med Sci Ed* 2018;49:87-92.
 67. Saban KL, Penckofer SM, Androwich I, Bryant FB. Health-related quality of life of patients following selected types of lumbar spinal surgery: A pilot study. *Health Qual Life Outcomes* 2007;5:71.
 68. Cheak-Zamora NC, Wyrwich KW, McBride TD. Reliability and validity of the SF-12v2 in the medical expenditure panel survey. *Qual Life Res* 2009;18:727-35.
 69. Streiner DL, Norman GR, Cairney J. *Health Measurement Scales: A Practical Guide to their Development and Use.* USA: Oxford University Press; 2015.
 70. Marx RG, Menezes A, Horovitz L, Jones EC, Warren RF. A comparison of two time intervals for test-retest reliability of health status instruments. *J Clin Epidemiol* 2003;56:730-5.
 71. Wan EY, Yu EY, Chin WY, Choi EP, Wu T, Lam CL. Evaluation of the responsiveness of short form-12 health survey version 2 (SF-12v2) in Chinese patients with hypertension in primary care. *Qual Life Res* 2019;28:2851-7.
 72. Gandek B, Ware JE Jr. Methods for validating and norming translations of health status questionnaires: The IQOLA project approach. *International quality of life assessment. J Clin Epidemiol* 1998;51:953-9.
 73. Lam C, Wong C, Lam ET, Lo YY, Huang WW. Population norm of Chinese (HK) SF-12 health survey-version 2 of Chinese adults in Hong Kong. *Hong Kong Pract* 2010;32:77-86.