ORIGINAL ARTICLE

Taiwanese Version of the EQ-5D: Validation in a Representative Sample of the Taiwanese Population

Ting-Jung Chang,[1,2] *Yen-Huei Tarn,*[3]* *Ching-Lin Hsieh,*[4] *Wen-Shyong Liou,*[5]
James W. Shaw,[6] *Xue Grace Chiou*[6]

Background/Purpose: We know of no validated Taiwanese-language instrument to measure a utility of the patient's health. Our aim was to evaluate the reliability and validity of a Taiwanese version of the EuroQol instrument (EQ-5D) in a Taiwanese population.

Methods: Questionnaires containing the Taiwanese versions of the EQ-5D and the Short-Form 12 Health Survey (SF-12) were sent to 12,923 people in Taiwan in December 2002. Concurrent validity of the EQ-5D was analyzed by assuming that subjects with problems in any EQ-5D dimensions had decreased SF-12 scores. Discriminant validity of the EQ-5D was analyzed by assuming that subjects with the following characteristics had lowered EQ-5D indexes and scores on the EQ-5D visual analog scale (VAS): more chronic diseases than others, serious illness, more hospitalizations in the past year than others, poor general health, and more outpatient visits than others. Test–retest reliability was analyzed in a subgroup of respondents who were evaluated twice within a month by using the intraclass correlation coefficient and the κ method.

Results: The general survey response rate was 12.7% (1644 of 12,923). SF-12 scores were lower in subjects reporting problems on EQ-5D dimensions than in others without such problems ($p < 0.01$). Subjects with more health problems than others had lower EQ-5D indexes and VAS scores ($p < 0.01$). The physical dimension of the EQ-5D was more strongly correlated with the SF-12 Physical Component Summary than with the Mental Component Summary; this finding satisfied the *a priori* hypothesis. For test–retest reliability of items on the EQ-5D, κ values ranged from 0.49 to 1 ($p < 0.001$).

Conclusion: The Taiwanese EQ-5D instrument appears to be a moderately valid and reliable tool for measuring the health status of the general population in Taiwan. [*J Formos Med Assoc* 2007;106(12):1023–1031]

Key Words: health surveys, quality of life, questionnaires, reproducibility of results, Taiwan

A number of instruments have been developed to measure health-related quality of life (HRQOL).[1] One such instrument is the EuroQol, or EQ-5D.[2,3] The EQ-5D is a preference-based instrument based on multi-attribute health-status classifiers. Population-specific health-state preferences have been derived for this tool. This instrument is being used with increasing frequency in both clinical and health services research, and it has been validated in numerous international studies.[4–6]

[1]Medical Equipment Review Committee, Tri-Service General Hospital, [2]Graduate Institute of Life Sciences, National Defense Medical Center, [3]HTA Task Force, Center For Drug Evaluation, [4]Department of Occupational Therapy, National Taiwan University, and [5]Bureau of Medical Affairs, Ministry of National Defense, Taipei, Taiwan, and [6]Department of Pharmacy Administration, College of Pharmacy, University of Illinois at Chicago, USA.

Received: March 16, 2007
Revised: April 30, 2007
Accepted: August 7, 2007

***Correspondence to:** Dr Yen-Huei Tarn, HTA Task Force, Center For Drug Evaluation, 1/F, No. 15-1, Section 1, Hangjou South Road, Taipei 100, Taiwan.
E-mail: yhtarn@cde.org.tw

The EQ-5D has been translated into many languages, including 60 official versions.[7,8] The usefulness and the construct validity of the various language versions of the EQ-5D have been tested in different community respondents and patient groups, including patients examined in general practice, those with rheumatoid arthritis, and those receiving cardiac rehabilitation.[9–12] Its generic character, proven efficacy, and brevity make the EQ-5D particularly attractive for use in the assessment of patients and in the economic evaluation of medications and treatments.

The aforementioned preferences for the EQ-5D have been most extensively linked to the preferences of the general population of the United Kingdom and the United States.[1,13,14] However, how these preferences apply to the Taiwanese population remains unclear. Moreover, to our knowledge, the psychometric properties of the Taiwanese version of the EQ-5D have not been assessed in Taiwan.

The purposes of this study were to evaluate the construct and concurrent validity as well as the test–retest reliability of the Taiwanese version of the EQ-5D in a large random sample of the general population in Taiwan.

Methods

Subjects and study design

The island of Taiwan comprises 22 cities and counties. The population in these geographical areas aged 20–64 years in 2000 (roughly 14 million citizens) was the sampling frame. According to data from the Measurement and Valuation of Health study by the EuroQol Group, an estimated sample of at least 1618 is needed to achieve a significant difference of 0.05 among individual health statuses, with an α value of 0.05 and a β value of 0.2.[15] Because of the difficulty in answering the questionnaire, we assumed a low response rate of around 15%. Therefore, we conducted oversampling of 12,923 individuals. They were randomly selected by applying a probability proportional to the sizes of the population of each city and county.

To conduct this postal survey, we mailed a questionnaire containing the EQ-5D and the Short-Form 12 Health Survey (SF-12) version 1 to these subjects in December 2002.[2] The EQ-5D questionnaire contained an EQ-5D self-classifier, a visual analog scale (VAS), sociodemographic questions (SDQ), and a 16-item health-status valuation questionnaire. Because of the difficulty of answering the valuation questionnaire, we excluded individuals who were older than 64 years or younger than 20 years.

Two weeks after we mailed the first questionnaire, we sent a reminder to individuals who had not responded. The original questionnaire and an explanation letter were sent to individuals whose answers were incomplete. A gift of NT$50 was sent to each respondent.

In the first questionnaire, the SDQ, an item asked if the subject would be willing to receive another similar questionnaire in the future. Respondents who responded yes were the candidates for retesting. We randomly selected 302 retest samples from 1145 respondents who consented to receive a second questionnaire. Retest measurement was conducted within 1 month.

Instruments

With the EQ-5D, a single index score of –0.59 to 1.00 can be generated for each respondent. The EQ-5D value set was first developed by using a time trade-off technique for a sample of these health states from a representative sample of the general population in the United Kingdom.[3] EQ-5D scores are calculated by subtracting the relevant tariffs from 1. For example, for health state 11223, the preference value is $1 - 0.036 - 0.123 - 0.236 - 0.269 - 0.081 = 0.225$, where a negative value means that a health state is worse than death.

The EQ-5D self-classifier assesses the five dimensions of mobility, self-dare, usual activity, pain/discomfort, and anxiety/depression on three levels (none, moderate, extreme). The VAS consisted of a 20-cm, vertical scale. Respondents were asked to classify and rate their health status on the day of the survey. The SDQ consisted of sociodemographic data and supplementary clinical information.

The last part was the 16-health status valuation questionnaire, which was designed to create the tariffs to calculate the preference value of every health state for another study; these data were not used for this validation study.

In developing the Taiwanese EQ-5D instrument, an official Taiwanese version was obtained from the EuroQol Group; however, it is not a validated instrument in Taiwan.[8] We independently created another translation of the English version. The final Taiwanese document was translated back into English and sent to the EuroQol Group for review. The wording of one item was not appropriate, and the Taiwanese characters were corrected. Two other words were changed to standardized translation terms. The Taiwanese EQ-5D was administered to 10 native Chinese-speaking Taiwan ese from diverse sociodemographic backgrounds. In this pilot study, subjects commented on the demographic part of the instrument, and changes were made when necessary. This final Taiwanese version was used in the validation study.

The SF-12 consists of 12 items measuring physical and mental health, and it yields two summary scores: the Mental Component Summary (MCS) and Physical Component Summary (PCS). High scores reflect improved health status. The Chinese Short Form-36 (SF-36) has been shown to be reliable and valid in general or medical populations.[16,17] The Chinese version of the SF-12 was derived from the validated version of the Chinese SF-36, and its exact wording was used in this study.

Data analysis

Concurrent validity of the EQ-5D was analyzed by making two assumptions. The first assumption was that subjects with some/moderate or extreme problems on any EQ-5D dimension have SF-12 PCS and MCS scores that are lower than those of other subjects. The second assumption was that today's (the testing day's) EQ-5D index or EQ VAS score is highly correlated with the SF-12 PCS and SF-12 MCS scores in the recent month.[18,19] Pearson's correlation coefficients were used to describe correlations among the EQ index, the VAS score, and the SF-12 subscale scores. For the EQ-5D index, we used EQ-5D weights derived from the general population of Japan; for the SF-12 PCS and SF-12 MCS scores, we used SF-12 weights derived from the general population of the United States.[3,20,21]

Discriminant validity of the EQ-5D was analyzed by assuming that subjects with the following characteristics had lowered EQ-5D indexes and VAS scores: more chronic diseases than other subjects, serious illness, more hospitalizations in the past year than other subjects, poor general health, and more outpatient visits in the past 2 weeks than other subjects.[3,10,14] For the EQ-5D index, we used EQ-5D weights derived from the general population of the United Kingdom and Japan.

Convergent validity was analyzed by assuming a high correlation between EQ-5D physical dimension scores (i.e. mobility, self-care, usual activities, or pain/discomfort) and SF-12 PCS scores and between EQ-5D mental dimension scores (i.e. anxiety/depression) and SF-12 MCS scores.[12,22]

Divergent validity was analyzed by assuming a low correlation between EQ-5D physical dimension scores and SF-12 MCS scores and between EQ-5D physical dimension scores and EQ-5D mental dimension scores. For the EQ-5D dimension, we used EQ-5D weights derived from the general population of Japan.

Pearson's correlation coefficients were used to describe correlations among the EQ-5D dimension and the SF-12 subscale scores. Spearman's correlation coefficients were used to describe correlations among the EQ-5D dimensions. Coefficients ≥ 0.5 indicated a strong correlation; 0.35–0.50 a moderate correlation; and < 0.35 a weak correlation.[23]

Test–retest reliability of EQ-5D indexes and VAS scores was determined by using intraclass correlation coefficients (ICCs), and response consistency on the five dimensions of the EQ-5D was determined by using the agreement method and Cohen's κ.[24] An ICC ≥ 0.7 is considered acceptable for test–retest reliability.[25] Values < 0.4 indicated poor agreement; 0.41–0.6 indicated moderate agreement; 0.61–0.8 substantial agreement; and > 0.8 almost perfect agreement.[26]

Reliability of the EQ-5D was assessed by correlating the first and retest sets of scores for subjects who indicated no change on the today's health status transition question. The question was: "compared with my general level of health over the past 12 months, my health state today is: better, much the same, worse."

Data were analyzed using STATA version 7.0 (StataCorp LP, College Station, TX, USA) and SPSS version 13 (SPSS Inc., Chicago, IL, USA). Analysis of variance and independent *t* test were used to test for differences in mean scores among groups with different attribute variables. Scheffé's test was used for *post hoc* comparisons. A χ^2 test was performed to evaluate sample characteristics. A *p* value of less than 0.05 indicated significant difference.

Results

We sent questionnaires to 12,923 individuals from the general population, 1768 of whom returned them. Questionnaires from 1644 were valid for analysis. Therefore, the survey response rate was 12.7%. The exclusion criterion was a missing response to any item on the SF-12 or EQ-5D self-classifier.

The Figure shows the geographic distributions of people aged 20–64 years in the general population of Taiwan,[27] of the sample given the postal survey, and of respondents. For these groups, the distribution trends were almost the same.

Table 1 summarizes the comparative information regarding the Taiwanese population and the response sample. The mean age (± standard deviation) of the respondents was 41.9 ± 12.1 years (range, 20–64 years). Many respondents were highly educated and unemployed. About 45% had at least a college education. The mean age of the respondents was almost the same as that of the 12,923 individuals who were sampled (mean age, 41.7 years) and similar to that of the population 20–64 years (mean age, 39.4 years).

Tables 2 and 3 provide the evidence of concurrent validity. Subjects with no or some/moderate problems on any EQ-5D dimension had increased or decreased scores on the SF-12 PCS and MCS, respectively ($p < 0.01$, *t* test; Table 2). Because fewer than 30 respondents with usable data reported extreme problems, the extreme category was combined with the some/moderate

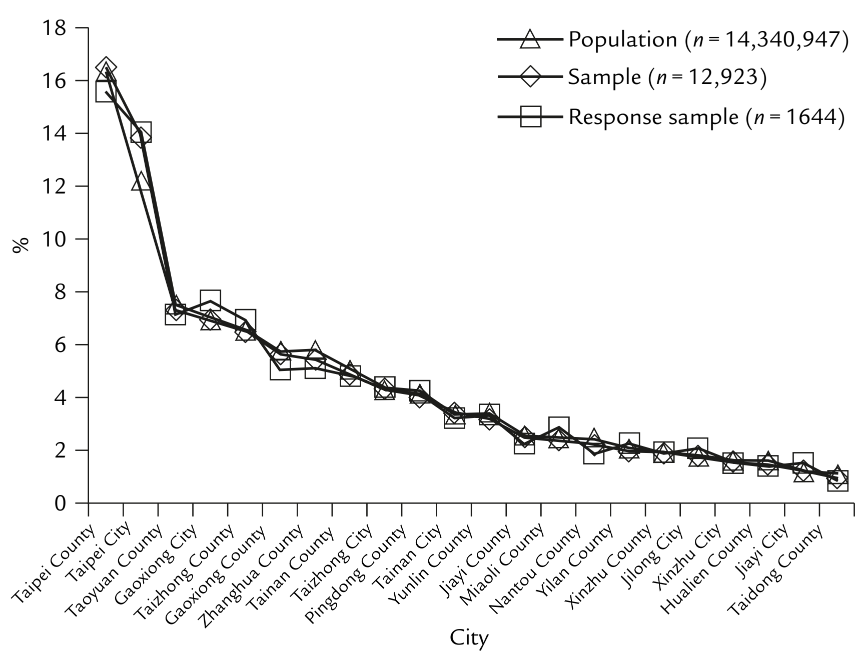

Figure. Geographical distributions of the general Taiwanese population, postal survey samples and respondents aged 20–64 years.

Table 1. Characteristics of 1644 respondents aged 20–64 years

	Response sample			Population	
	n	Mean (SD)	%	%	Mean
Age	1618	41.9 (12.1)			39.4†
Salary in past 1 yr (US$)	1400	11,991 (10,839)			
Inpatient visits in past 1 yr	1632	0.1 (0.6)			
Outpatient visits in past half mo	1631	0.9 (2.4)			
Out of pocket expenses for disease in past 1 yr (US$)	1489	335 (683)			
Female	828		50.4	49.6	
Marital status					
Unmarried/separated/divorced	488		29.7	32.1	
Married	1089		66.2	65.4	
Widowed	49		3.0	2.4	
Education level*					
≤ Senior high school	843		51.3	72.1	
College or university	649		39.5	25.8	
≥ Master's degree	88		5.4	2.2	
Employment status*					
Employed	1068		65.0	70.6	
Unemployed/homemaker/retired	545		33.2	29.4	

**$p < 0.01$, χ^2 test; †estimated mean age = (subgroup median age × number of people in subgroup) ÷ (number of people aged 20–64 years). SD = standard deviation.*

Table 2. Test of concurrent validity ($n = 1644$)*

EQ-5D dimension	*n*	SF-12 PCS†	SF-12 MCS†
Mobility			
No problem	1581	50.4	45.6
With problem	63	34.4	39.2
Self-care			
No problem	1619	50.0	45.4
With problem	25	32.0	37.6
Usual activity			
No problem	1531	50.7	46.0
With problem	113	36.6	36.8
Pain/discomfort			
No problem	1111	52.5	47.1
With problem	533	44.0	41.6
Anxiety/depression			
No problem	1175	50.8	48.6
With problem	469	47.1	37.2

**All $p < 0.01$, independent t test; †using US Norm-Based Standardization of Scale Scores Weight. PCS = Physical Component Summary; MCS = Mental Component Summary.*

category. EQ-5D indexes and VAS scores on the study day were moderately to strongly correlated with SF-12 PCS and SF-12 MCS scores in the recent month (Table 3).

For each EQ-5D dimension, subjects reported (moderate or extreme) problems (Table 2). Dimensions in which subjects reported the most problems were pain/discomfort (moderate, $n=519$ [31.6%]; extreme, $n=14$ [0.9%]), followed by anxiety/depression problems (moderate, $n=450$ [27.4%]; extreme, $n=19$ [1.2%]).

Subjects who reported serious illnesses, poor general health on the testing day, more outpatient visits, inpatient visits, and chronic diseases than other subjects had low EQ-5D indexes and EQ VAS scores ($p<0.01$, t test; Table 4).

Table 5 shows the data for convergent and divergent validity. The EQ-5D had two dimensions: mental health (anxiety/depression) and physical health (mobility, self-care, usual activities, pain/discomfort). Convergent validity was demonstrated by the response on EQ-5D physical health dimension, which was correlated more with the SF-12 PCS ($r=0.26$–0.51) than with the SF-12 MCS ($r=0.09$–0.26). Additional evidence came from the response to the anxiety dimension of the EQ-5D, which was more strongly correlated with the SF-12 MCS ($r=0.53$) than with the PCS ($r=0.22$). Divergent validity was demonstrated by the response to the EQ-5D physical health dimension, which was

Table 3. Correlation of SF-12 subscale, EQ-5D index, EQ VAS score for concurrent validity ($n=1644$)

	SF-12 PCS*	SF-12 MCS*
SF-12 MCS*	0.10	–
EQ-5D index†	0.53‡	0.42‡
EQ VAS score	0.45‡	0.49‡

**Using US Norm-Based Standardization of Scale Scores Weight; †using Japanese TTO value set; ‡hypotheses which are supported. PCS = Physical Component Summary; MCS = Mental Component Summary.*

Table 4. Test of discriminant validity

Health condition	EQ-5D index* (n)	EQ-5D index† (n)	EQ VAS score (n)	Assumption
Experienced serious illness				
No	0.81 (1468)	0.86 (1468)	79.3 (1423)	
Yes	0.72 (173)	0.70 (173)	64.3 (168)	No > Yes‡
General health today				
Excellent	0.84 (62)	0.89 (62)	91.8 (61)	
Very good	0.83 (597)	0.89 (597)	86.4 (585)‡	Excellent > Very good
Good	0.81 (452)	0.85 (452)	77.3 (436)	Very good > Good‡
Fair	0.76 (462)	0.79 (462)	68.3 (444)	Good > Fair‡
Poor	0.64 (57)	0.54 (57)	50.6 (55)	Fair > Poor‡
Outpatient visits in past half mo				
0	0.82 (1090)	0.87 (1090)	80.7 (1060)	
1–2	0.77 (395)	0.81 (395)	73.4 (380)	0 > 1–2‡
≥ 3	0.73 (145)	0.74 (145)	67.0 (140)	1–2 > ≥ 3‡
Inpatient visits in past yr				
0	0.80 (1526)	0.85 (1526)	78.4 (1483)	
≥ 1	0.73 (106)	0.73 (106)	68.4 (99)	0 > ≥ 1‡
Number of chronic diseases				
0	0.82 (687)	0.88 (687)	82.4 (669)	
1–2	0.79 (847)	0.83 (847)	76.0 (820)	0 > 1–2‡
≥ 3	0.72 (110)	0.71 (110)	61.9 (105)	1–2 > ≥ 3‡

**Using Japanese TTO value set; †using UK TTO value set; ‡$p<0.01$, ANOVA or independent t test.*

Table 5. Correlation of the SF-12 subscales and five dimensions of the EQ-5D for convergent and divergent validity ($n = 1644$)*†

EQ-5D	SF-12 PCS‡	SF-12 MCS‡	Anxiety/depression§
Mobility	**–0.32**	***–0.11***	***0.14***
Self-care	**–0.26**	***–0.09***	***0.11***
Usual activity	**–0.41**	***–0.24***	***0.26***
Pain/discomfort	**–0.51**	***–0.26***	***0.40***
Anxiety/depression	***–0.22***	**–0.53**	–

**Hypotheses are shown for convergent validity (bold) and divergent validity (bold italics), and hypotheses which are supported are underlined; †coefficients ≥ 0.5 indicate strong correlation, 0.35–0.50 moderate correlation, and < 0.35 weak correlation; ‡Pearson's correlation; §Spearman's correlation. PCS = Physical Component Summary; MCS = Mental Component Summary.*

weakly correlated with the SF-12 MCS ($r = 0.09$–0.26). Also supporting divergent validity was the response to the EQ-5D anxiety dimension, which was also weakly correlated with the SF-12 PCS ($r = 0.22$). Divergent validity was also demonstrated by the weak correlation ($\rho = 0.11$–0.26) between results from the EQ-5D physical and mental dimensions, with the exception of pain/discomfort ($\rho = 0.40$).

In the retest samples, 184 of 302 subjects returned the retest questionnaire, and results of 112 were usable for analysis. The test–retest survey response rate was 37.1%. The exclusion criteria were any missing item on the SF-12 or EQ-5D self-classifier ($n = 4$), a difference in sex between the test and retest surveys ($n = 6$), a 1-year difference in age between the test and retest surveys ($n = 14$), and a change in the subject's health status on the testing day between the test and retest surveys ($n = 48$).

The ICCs of reliability were 0.51 (95% confidence interval [CI], 0.36–0.63; $p < 0.000$) for the EQ-5D index and 0.70 (95% CI, 0.58–0.78; $p < 0.000$) for the VAS score. Agreements for EQ-5D items regarding mobility, self-care, usual activities, pain/discomfort, and anxiety/depression were 98.2%, 100%, 98.2%, 84.8%, and 83%, respectively; Cohen's κ values for these items were 0.49 (95% CI, 0.31–0.68; $p < 0.0001$), 1.0, 0.74 (95% CI, 0.56–0.92; $p < 0.0001$), 0.58 (95% CI, 0.40–0.76; $p < 0.0001$), and 0.50 (95% CI, 0.32–0.67; $p < 0.0001$), respectively ($n = 112$; interval, 3–5 weeks). The study demonstrated moderate evidence to support test–retest reliability of the Taiwanese version of the EQ-5D classifier.

Discussion

This study represents the first application of the Taiwanese version of the EQ-5D, a generic instrument, in a large and representative population in Taiwan. The respondents appeared to be well represented in terms of geography, age, sex and marital status. However, their education level and unemployment rate were higher than those in the general population of people aged 20–64 years. The EQ-5D questionnaire contained a self-classifier, a VAS, and SDQ, as well as a 16-health status valuation questionnaire that was difficult to answer. Individuals with high levels of education might have been able to answer the 16-health status valuation questionnaire more easily than others, and unemployed people might have had more time to answer the questions than others.

Two possible reasons for the low response rate were difficulty in answering the valuation questionnaire and the fact that the questionnaires were mailed. Offering NT$50 as an incentive, we re-sent questionnaires to non-respondents and to respondents whose answers were incomplete; however, the response rate did not increase much.

The EQ-5D and SF-12 have similar constructs; correlations among the EQ-5D, SF-12 PCS and SF-12 MCS were moderate. Petrou and Hockley reported that the correlation coefficients for the EQ-5D with the SF-12 PCS and the MCS were 0.65 and 0.54 respectively, whereas coefficients for the EQ-VAS with the SF-12 PCS and MCS were 0.65 and 0.44, respectively ($n = 321$).[28] We found that the correlations between these instruments were moderate to strong (Table 3).

The Taiwanese version of the EQ-5D appeared to have acceptable concurrent discriminant validity. The EQ-5D was able to distinguish between groups of self-reported health statuses. These patterns suggested that the EQ-5D could detect differences in HRQOL among people with different degrees of health (Table 2). In addition, subjects with serious illness, poor general health, and more outpatient visits, inpatient visits, and chronic diseases than others had relatively low EQ-5D indexes and VAS scores (Table 4). Petrou and Hockley reported mean EQ-5D and SF-6D utility scores for each of the self-reported health subgroups. Both multi-attribute utility measures significantly differed among subjects who described their health status as very good, good, fair, bad, or very bad ($p < 0.001$). In addition, both measures generated utility scores that decreased monotonically with deteriorating self-reported health status ($p < 0.001$, test for linear trend).[28] Our findings confirmed that HRQOL was correlated with disease activity. These EQ-5D results corresponded well with this theory.

Convergent and divergent validities of EQ-5D dimensions were supported by the expected relationship with scores on the SF-12 subscales. The strongest correlations between the measures were observed between similar constructs, that is, between the SF-12 MCS with anxiety/depression on the EQ-5D and between the SF-12 PCS with pain/discomfort on the EQ-5D. In contrast, weak correlations were observed between different constructs, that is, between the SF-12 MCS and mobility, self-care, usual activities and pain/discomfort on the EQ-5D, and between SF-12 PCS and anxiety/depression on the EQ-5D.

Haywood et al determined the test–retest reliability of the EQ-5D by studying HRQOL in patients with ankylosing spondylitis. At 2 weeks, the ICC was 0.83 for the EQ-5D and 0.80 for the VAS ($n = 176$).[29] Konig et al reported that the agreement between responses and EQ-5D items ranged from 80.4% (anxiety/discomfort) to 100% (self-care). For the VAS score and the EQ-5D index, ICCs were 0.77 and 0.89, respectively ($n = 52$, interval of 19.5 days).[30] Their study demonstrated consistent results on five dimensions of the EQ-5D self-classifier, on the VAS, and on the EQ-5D index. Because this study distribution was strongly skewed toward improved health states, 96%, 98%, 93%, 67% and 71% of the subjects scored at the top for the dimensions of mobility, self-care, usual activity, pain/discomfort and anxiety/depression, respectively. This skewing toward improved health states reflected a homogeneous population, making for low reliability.[31]

Validity and reliability were not as good as they were in the reference study. One reason might have been because the study was a postal survey and not an interview survey. The study design did not permit us to explain items and answer respondents' questions face to face. Another reason was that the retest interval was 3–5 weeks longer than that of the other study, which was 2 weeks.

This study had a few limitations. First, the response rate was low because the study was a postal survey and because the questionnaire contained 16 items for health status valuation that were difficult to answer. This low rate raises issues about the representativeness and generalizability of the sample to the whole population of Taiwan. Second, we used EQ-5D weights derived from the general population of the United Kingdom and Japan and SF-12 weights derived from the general population of the United States because such weights are lacking for Taiwan.[3,20,21] Third, we used the Chinese version of the SF-12 to examine the validity of the EQ-5D. Although the Chinese version of the SF-36 has been well validated,[16,17] the SF-12, to our knowledge, has not been validated.

The cross-culturally adapted Taiwanese version of the EQ-5D was well accepted and demonstrated acceptable psychometric properties, including moderate-to-substantial validity and reliability in a highly educated general Taiwanese population aged 20–64 years. Our data provide a basis for further studies of the Taiwanese EQ-5D instrument. For example, we will establish the EQ-5D health-status weights in Taiwan. A potential advantage of the EQ-5D is its brevity. When one requires a limited health instrument with a minimal burden on respondents for use in routine practice or in health economic analyses, we recommend the EQ-5D self-classifier and the EQ VAS.

References

1. Hawthorne G, Richardson J, Day NA. A comparison of the Assessment of Quality of Life (AQoL) with four other generic utility instruments. *Ann Med* 2001;33:358–70.
2. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med* 2001;33:337–43.
3. Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997;35:1095–108.
4. Badia X, Roset M, Herdman M, et al. A comparison of United Kingdom and Spanish general population time trade-off values for EQ-5D health states. *Med Decis Making* 2001;21:7–16.
5. Wang H, Kindig DA, Mullahy J. Variation in Chinese population health related quality of life: results from a EuroQol study in Beijing, China. *Qual Life Res* 2005;14:119–32.
6. Prieto L, Sacristan JA, Hormaechea JA, et al. Psychometric validation of a generic health-related quality of life measure (EQ-5D) in a sample of schizophrenic patients. *Curr Med Res Opin* 2004;20:827–35.
7. EuroQol Group. *EQ-5D: An Instrument to Describe and Value Health.* Available at: http://www.euroqol.org/
8. Ellis JJ, Eagle KA, Kline-Rogers EM, et al. Perceived work performance of patients who experienced an acute coronary syndrome event. *Cardiology* 2005;104:120–6.
9. Glick HA, Polsky D, Willke RJ, et al. A comparison of preference assessment instruments used in a clinical trial: responses to the visual analog scale from the EuroQol EQ-5D and the Health Utilities Index. *Med Decis Making* 1999; 19:265–75.
10. Luo N, Chew LH, Fong KY, et al. Validity and reliability of the EQ-5D self-report questionnaire in Chinese-speaking patients with rheumatic diseases in Singapore. *Ann Acad Med Singapore* 2003;32:685–90.
11. Schweikert B, Hahmann H, Leidl R. Validation of the EuroQol questionnaire in cardiac rehabilitation. *Heart* 2006;92:62–7.
12. Lubetkin EI, Jia H, Gold MR. Construct validity of the EQ-5D in low-income Chinese American primary care patients. *Qual Life Res* 2004;13:1459–68.
13. Luo N, Johnson JA, Shaw JW, et al. Self-reported health status of the general adult U.S. population as assessed by the EQ-5D and Health Utilities Index. *Med Care* 2005;43: 1078–86.
14. Kind P, Dolan P, Gudex C, et al. Variations in population health status: results from a United Kingdom national questionnaire survey. *BMJ* 1998;316:736–41.
15. Williams A. *The Measurement and Valuation of Health: A Chronicle.* Centre for Health Economics, University of York, England, 1995:17.
16. Li TC, Lee YD, Lin CC, et al. Quality of life of primary caregivers of elderly with cerebrovascular disease or diabetes hospitalized for acute care: assessment of well-being and functioning using the SF-36 health questionnaire. *Qual Life Res* 2004;13:1081–8.
17. Wang SJ, Fuh JL, Lu SR, et al. Quality of life differs among headache diagnoses: analysis of SF-36 survey in 901 headache patients. *Pain* 2001;89:285–92.
18. Cleemput I, Kesteloot K, Moons P, et al. The construct and concurrent validity of the EQ-5D in a renal transplant population. *Value Health* 2004;7:499–509.
19. Johnson JA, Coons SJ. Comparison of the EQ-5D and SF-12 in an adult US sample. *Qual Life Res* 1998;7:155–66.
20. Tsuchiya A, Ikeda S, Ikegami N, et al. Estimating an EQ-5D population value set: the case of Japan. *Health Econ* 2002;11:341–53.
21. Ware JE, Kosinski M, Keller SD. *SF-12: How To Score the SF-12 Physical & Mental Health Summary Scales.* Lincoln, RI: Quality Metric Incorporated and Boston, MA: The Health Assessment Lab, 1998.
22. Salyers MP, Bosworth HB, Swanson JW, et al. Reliability and validity of the SF-12 health survey among people with severe mental illness. *Med Care* 2000;38:1141–50.
23. Juniper E, Gordon H, Roman J. How to develop and validate a new health-related quality of life instrument. In: Spilker B, ed. *Quality of Life and Pharmacoeconomics in Clinical Trial*. Philadelphia: Lippincott-Raven, 1996:49–56.
24. Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures. Statistics and strategies for evaluation. *Control Clin Trials* 1991;12:142S–58S.
25. Hays R, Anderson R, Revicki D. Assessing reliability and validity of measurement in clinical trials. In: Staquet MJ, Hayes RD, Fayers PM, eds. *Quality of Life Assessment in Clinical Trials: Methods and Practice*. Oxford: Oxford University Press, 1998:169–82.
26. Landis J, Koch G. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
27. http://www.moi.gov.tw/stat/english/index.asp Taipei, 2000.
28. Petrou S, Hockley C. An investigation into the empirical validity of the EQ-5D and SF-6D based on hypothetical preferences in a general population. *Health Econ* 2005;14: 1169–89.
29. Haywood KL, Garratt AM, Dziedzic K, et al. Generic measures of health-related quality of life in ankylosing spondylitis: reliability, validity and responsiveness. *Rheumatology (Oxford)* 2002;41:1380–7.
30. Konig HH, Ulshofer A, Gregor M, et al. Validation of the EuroQol questionnaire in patients with inflammatory bowel disease. *Eur J Gastroenterol Hepatol* 2002;14:1205–15.
31. Bartko JJ. Measurement and reliability: statistical thinking considerations. *Schizophr Bull* 1991;17:483–9.