

# Reliability and Validity of the English-, Chinese- and Malay-Language Versions of the World Health Organization Quality of Life (WHOQOL-BREF) Questionnaire in Singapore

Yin Bun Cheung,<sup>1,2</sup>*PhD*, Khung Keong Yeo,<sup>3,4</sup>*MBBS, ABIM*, Kok Joon Chong,<sup>5</sup>*BSc (Hons)*, Eric YH Khoo,<sup>5,6</sup>*MD, FRCP*, Hwee Lin Wee,<sup>7,8</sup>*PhD*

## Abstract

**Introduction:** The World Health Organization Quality of Life (WHOQOL-BREF) questionnaire is a 26-item questionnaire that evaluates 4 domains of quality of life (QoL), namely Physical, Psychological, Social Relationships and Environment. This study aimed to evaluate the validity and reliability of the WHOQOL-BREF among Singapore residents aged 21 and above. **Materials and Methods:** We recruited participants from the general population by using multistage cluster sampling and participants from 2 hospitals by using convenience sampling. Participants completed either English, Chinese or Malay versions of the WHOQOL-BREF and the EuroQoL 5 Dimension 5 Levels (EQ-5D-5L) questionnaires. Confirmatory factor analysis, known-group validity, internal consistency (Cronbach's alpha) and test-retest reliability using the intraclass correlation coefficient (ICC) were performed. **Results:** Data from 1316 participants were analysed (Chinese: 46.9%, Malay: 41.0% and Indian: 11.7%; 57.5% men, mean standard deviation [SD, range] age: 51.9 [15.68, 24 to 90] years); 154 participants took part in the retest in various languages (English: 60, Chinese: 49 and Malay: 45). Tucker-Lewis Index (TLI) was 0.919, 0.913 and 0.909 for the English, Chinese and Malay versions, respectively. Standardised root mean square residual (SRMR) was 0.067, 0.074 and 0.094, respectively. Cronbach's alpha exceeded 0.7 and ICC exceeded 0.4 for all domains in all language versions. **Conclusion:** The WHOQOL-BREF is valid and reliable for assessing QoL in Singapore. Model fit is reasonable with room for improvement.

Ann Acad Med Singapore 2017;46:461-9

**Key words:** Health-related quality of life, Linking values, Mapping functions, Patient-reported outcomes, Preference-based measures

## Introduction

Health-related quality of life (HRQoL) is a multidimensional concept that describes the aspects pertaining to health. It is generally agreed that HRQoL assesses the well-being and functioning of an individual, comprising physical and emotional aspects while opinions are divided as to whether it should also include the individual's social aspects.<sup>1</sup> HRQoL is increasingly recognised as an important measure of

patient-reported outcomes and treatment efficacy, and requires reliable assessment tools that demonstrate reliability and validity for a specific patient population.<sup>2</sup>

The World Health Organization Quality of Life (WHOQOL-BREF) questionnaire—an abbreviated version of the WHOQOL-100 test—was designed for studies that require a brief assessment of quality of life (QoL), such as epidemiological studies, clinical trials and routine clinical work.<sup>3,4</sup> It was also designed as a cross-cultural

<sup>1</sup>Centre for Quantitative Medicine, Duke-NUS Medical School, Singapore

<sup>2</sup>Tampere Centre for Child Health Research, University of Tampere and Tampere University Hospital, Finland

<sup>3</sup>National Heart Centre Singapore, Singapore

<sup>4</sup>Duke-NUS Medical School, Singapore

<sup>5</sup>Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore

<sup>6</sup>Division of Endocrinology, University Medicine Cluster, National University Hospital, Singapore

<sup>7</sup>Department of Pharmacy, Faculty of Science, National University of Singapore, Singapore

<sup>8</sup>Saw Swee Hock School of Public Health, National University of Singapore, Singapore

Address for Correspondence: Dr Wee Hwee Lin, Department of Pharmacy, Faculty of Science, National University of Singapore, Block S4A, Level 3, 18 Science Drive 4, Singapore 117543.

Email: phawhl@nus.edu.sg

questionnaire, and has since been used in a large number of countries across various continents including Africa,<sup>5,6</sup> South America,<sup>7</sup> Europe,<sup>6,8-10</sup> Oceania<sup>11</sup> and Asia.<sup>12-16</sup> Even though the WHOQOL-BREF was designed as a cross-cultural questionnaire, Singapore was not one of the 20 centres in 18 countries involved in its development.<sup>3</sup> As Singapore has a multiethnic population that uses several languages, the psychometric properties of the WHOQOL-BREF test need to be evaluated to suit the Singaporean population. The Chinese version of the WHOQOL-BREF, for instance, was tested in Mainland China, Taiwan and Hong Kong.<sup>12,14,15</sup> Fortunately, the factor structure of the WHOQOL-BREF was replicated in all 3 Chinese-speaking populations. However, there were previous studies which reported that the factor structure of the WHOQOL-BREF may not be replicable in some populations. For example, in a study that evaluated the factor structure of the Spanish version of the WHOQOL-BREF across 9 Spanish-speaking countries, it was reported that the model fit statistics from the confirmatory factor analysis were poor for most of the countries.<sup>17</sup> The WHOQOL-BREF has previously been used in Singapore,<sup>18</sup> on a separate sample of patients with schizophrenia and on a sample of patients with diabetes mellitus.<sup>19</sup> It has not been used in the general population. Furthermore, the study among patients with schizophrenia did not evaluate the psychometric properties of the WHOQOL-BREF. Hence, this study aimed to determine the validity and reliability of the WHOQOL-BREF for assessing HRQoL in 3 languages – English, Chinese, and Malay – in a multiethnic Singaporean population.

## Materials and Methods

### *Participants and Study Design*

Singapore is a multiethnic society, comprising 3 main ethnic groups – Chinese, Malays and Indians. This study encompasses participants from the general population as well as 2 clinic samples. Ethics approval for this study was obtained from the National Healthcare Group Domain Specific Review Board (Ref. 2013/00747) and the SingHealth Centralised Institutional Review Board (Ref. 2015/2041). Participants provided written informed consent. Participants from the general population were recruited based on a multistage cluster sampling method using postcodes as the primary sampling unit (PSU), followed by the selection of households and then the selection of respondents. Three call attempts (1<sup>st</sup> attempt and 2 callbacks) were made at different days and times of the week. Contacted participants of each household residing in public housing (i.e. high rise flats built by the government) were selected based on a prespecified quota for language of interviews, age and gender within each ethnic group. More than 80% of Singaporeans reside in public housing. A list of potential

households was selected using a sampling frame maintained by the survey company engaged to conduct this survey. The face-to-face interviews were conducted in the participants' homes between October 2014 and January 2015. A subset of participants was revisited within 2 weeks for assessment of test-retest reliability. The interviewers were trained in survey data collection and were effectively bilingual in English and Mother Tongue (i.e. Chinese or Malay). We paired interviewers and participants by ethnicity to improve response rate. For example, Malay interviewers would be given a list of Malay households to follow-up with.

The clinic samples were chosen to enrich the dataset, such that a wider spread of health status was represented. The clinic participants were drawn from two separate studies in outpatient clinics at the National Heart Centre Singapore (NHCS) and the Division of Endocrinology at the National University Hospital (NUH). Recruitment was conducted by research assistants via convenience sampling in the clinics while patients were waiting to see the doctor. Patients with recent acute myocardial infarction (STEMI), haemodynamic instability or gestational diabetes were excluded. Interviews were carried out between March 2015 and February 2016.

In both the general population and clinic samples, an eligible participant had to be a Singapore resident aged 21 years and above who spoke English, Chinese (Mandarin) or Malay. Participants who spoke only Tamil were excluded as the questionnaires were not available in Tamil. All participants read and signed the written informed consent form prior to commencement of the interviews.

### *Questionnaires*

The WHOQOL-BREF and EuroQoL 5 Dimension 5 Levels (EQ-5D-5L) questionnaires were used, in addition to a socioeconomic and clinical questionnaire to capture information such as age, gender, ethnicity and self-reported medical conditions. For the Chinese version of the questionnaires, simplified Chinese was used.

### WHOQOL-BREF

The WHOQOL-BREF is a subjective evaluation of individuals' perceptions of their positions in life, in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns.<sup>4</sup> It is not equivalent to a health status measure such as the EQ-5D-5L. The WHOQOL-BREF is a 26-item questionnaire that includes 1 item from each of the 24 facets contained in the WHOQOL-100 and 2 additional items on overall quality of life and general health. The 24 items are organised into 4 domains, namely WHOQOL-BREF Physical (WHO-PHY), Psychological (WHO-PSY), Social Relationships (WHO-SOC) and Environment (WHO-ENV). Three negatively phrased items were

reversed scored. Domain scores were computed by taking the mean of the scores of the items that constitute the domain and multiplied by 4 so that the scores were directly comparable with those derived from WHOQOL-100. Mean substitution was performed for missing data; missing item scores were replaced with the mean of the non-missing item scores in the same domain if at least 50% of the items in the domain were non-missing. The English and Chinese versions of WHOQOL-BREF were obtained from the WHO website while the Malay (Malaysian) version was obtained from the authors of a Malay study.<sup>20</sup>

### EQ-5D-5L

The EQ-5D-5L is a health status measure comprising 5 dimensions: mobility (EQ-MOB), self-care (EQ-SC), usual activities (EQ-UA), pain/discomfort (EQ-PN) and anxiety/depression (EQ-AD) with 5 response options for each dimension (no, slight, moderate, severe, extreme problems/unable).<sup>21</sup> Participants indicate their health status for the day by selecting 1 response option per dimension, giving rise to a 5-digit health profile, which is in turn linked with a utility score. We used the Japanese value set from the EuroQoL Group's crosswalk project as a Singapore dataset is not available.<sup>13,22</sup> The Japanese value set has a possible range of -0.111 to 1. Participants with missing data were excluded from analysis.

### Statistical Analyses

We conducted confirmatory factor analysis (CFA) with ordered categorical factor indicators using a robust weighted least squares estimator (WLSMV) with a diagonal weight matrix in MPlus 8.0 (Muthén & Muthén, California, USA). The raw WHOQOL-BREF item scores were used for this purpose and items with missing responses were deleted listwise. The model structure was defined in 2 ways: first, according to the original 4-factor model and then we added a second-order factor contributed by the 4 factors. Based on the rule of thumb by Hu and Bentler,<sup>23</sup> we defined good model fit as a Tucker-Lewis Index (TLI) = 0.95 and standardised root mean square residual (SRMR) = 0.08.

Known-group validity was investigated with a *priori* hypotheses that differences in mean WHOQOL-BREF scale scores between individuals with and without a specific chronic medical condition will exceed half a standard deviation (SD) (rule of thumb for defining minimally important difference).<sup>24</sup> We limited the analyses to 5 medical conditions that had a prevalence of at least 5% across the 3 language groups: diabetes, hypertension, dyslipidaemia, heart disease and mental illness. In addition, we hypothesised that the difference in mean WHO-PHY scores between participants with and without problem on EQ-MOB as well as the difference

in mean WHO-PSY scores between participants with and without problem on EQ-AD will exceed half a SD.

Two aspects of reliability were assessed: internal consistency and test-retest reliability. Internal consistency was assessed using Cronbach's alpha; test-retest reliability was assessed using intraclass correlation coefficient (ICC). Cronbach's alpha exceeding 0.7 indicates acceptable reliability for group level; 0.9 for individual level.<sup>25</sup> ICC was used in a subset of study participants 2 weeks after the first interview. We defined ICC <0.40 as poor; 0.40-0.59 as fair; 0.60-0.74 as good; and 0.75 and above as excellent.<sup>26</sup>

### Results

The study included a total of 1316 participants, with 892 participants from the general population, 221 participants from NHCS and 203 participants from NUH. In terms of ethnicity, there were 617 Chinese, 539 Malays, 154 Indians and 6 Others. The English-, Chinese- and Malay-language versions were completed by 454, 443 and 419 participants, respectively.

A larger proportion of those who completed the English-language version were younger and had tertiary education compared to those who completed the Chinese- and Malay-language versions. The prevalence of high blood pressure and high cholesterol was significantly higher among those who completed the Chinese- and Malay-language versions compared to those who completed the English version, presumably because these participants were older (Table 1).

A higher proportion of participants who completed the Malay-language version lived in small (1- or 2-room) public housing. WHO-PSY scores were 3.69 points lower ( $P < 0.001$ ), WHO-SOC were 3.53 points lower ( $P < 0.001$ ) and WHO-ENV scores were 2.12 points lower ( $P = 0.022$ ) for someone who completed the Chinese-language version compared to someone who completed the English-language version after adjusting for age, gender, educational level and housing type (Table 2).

We followed-up with 154 participants for retest. The number of participants analysed for the English-, Chinese- and Malay-language versions were 60, 49 and 45, respectively. Mean SD age of the participants in the retest was 39.7 years (13.51, 24.0 to 84.0), 52.7 years (13.62, 24.0 to 77.0) and 49.7 years (14.97, 24.0 to 83.0) for the English-, Chinese- and Malay-language versions, respectively. Hence, they were generally younger than those who participated in the first round of interviews.

### Confirmatory Factor Analysis

In the 4-factor CFA, TLI = 0.919 and SRMR = 0.067 for the English-language version (n = 399). For the Chinese-language version (n = 367), TLI = 0.913 and SRMR =

Table 1. Characteristics of Study Participants (n = 1316)

	English Language Survey (n = 454)	Chinese Language Survey (n = 443)	Malay Language Survey (n = 419)	Pearson's Chi- Squared Statistics/ F-Statistics, Where Applicable	P Value
<b>Gender, n (%)</b>					
Male	277 (61.0)	250 (56.4)	229 (54.7)	3.89	0.143
<b>Age group (years)</b>					
Mean (SD, range)	49.1 (16.16, 24.0 – 90.0)	55.0 (15.33, 24.0 – 84.0)	51.8 (14.96, 24.0 – 85.0)	16.14	<0.001
21 – 29	82 (18.1)	39 (8.8)	48 (11.5)	43.44	<0.001
30 – 39	90 (19.8)	59 (13.3)	61 (14.6)		
40 – 49	80 (17.6)	76 (17.2)	82 (19.6)		
50 – 59	80 (17.6)	93 (21.0)	106 (25.3)		
60 – 69	75 (16.5)	108 (24.4)	81 (19.3)		
>70	47 (10.4)	68 (15.4)	41 (9.8)		
<b>Ethnicity, n (%)</b>					
Chinese	174 (38.3)	443 (100)	0 (0)		
Malay	124 (27.3)	0 (0)	415 (99.0)		
Indian	152 (33.5)	0 (0)	2 (0.5)		
Others	4 (0.9)	0 (0)	2 (0.5)		
<b>Education, n (%)</b>				148.54	<0.001
Primary or less	39 (8.6)	141 (31.8)	113 (27.0)		
Secondary or less	232 (51.1)	180 (40.6)	259 (61.8)		
Tertiary	183 (40.3)	122 (27.5)	47 (11.2)		
<b>Housing, n (%)</b>				39.93	<0.001
Smaller public	279 (61.5)	290 (65.5)	329 (78.5)		
Larger public	140 (30.8)	122 (27.5)	83 (19.8)		
Private	33 (7.3)	30 (6.8)	5 (1.2)		
Others	2 (0.4)	1 (0.2)	2 (0.5)		
<b>Chronic medical conditions, n (%)</b>					
Diabetes mellitus	123 (27.1)	120 (27.1)	119 (28.4)	2.87	0.580
High blood pressure	129 (28.4)	171 (38.6)	145 (34.6)	12.42	0.014
High cholesterol	84 (18.5)	118 (26.6)	94 (22.4)	8.51	0.014
Heart disease	128 (28.2)	121 (27.3)	113 (27.0)	2.08	0.721
Mental illnesses	66 (14.5)	80 (18.1)	53 (12.7)	8.40	0.078
<b>Mean (SD, range) WHOQOL-BREF domain scores</b>					
PHY	74.9 (15.5, 12.5 – 100)	71.7 (15.6, 7.1 – 100)	72.9 (18.0, 10.7 – 100)	4.20	0.015
PSY	72.5 (14.1, 4.2 – 100)	67.1 (15.9, 12.5 – 100)	71.0 (13.8, 16.7 – 100)	16.04	<0.001
SOC	74.0 (16.1, 0 – 100)	69.0 (13.4, 16.7 – 100)	74.1 (14.1, 33.3 – 100)	17.70	<0.001
ENV	72.0 (14.1, 18.8 – 100)	68.2 (14.6, 9.4 – 100)	69.0 (12.5, 34.4 – 100)	9.37	<0.001
<b>Mean (SD, range) EQ-5D utility scores</b>					
	0.92 (0.10, 0.25 – 1)	0.91 (0.10, 0.29 – 1)	0.91 (0.15, 0.02 – 1)	1.02	0.362

ENV: Environment; PHY: Physical; PSY: Psychological; SD: Standard deviation; SOC: Social Relationships; WHOQOL-BREF: World Health Organization Quality of Life

Table 1. Characteristics of Study Participants (n = 1316) (Cont'd)

	English Language Survey (n = 454)	Chinese Language Survey (n = 443)	Malay Language Survey (n = 419)	Pearson's Chi- Squared Statistics/ F-Statistics, Where Applicable	P Value
With problems on EQ-5D, n (%)					
Mobility	50 (11.0)	54 (12.2)	76 (18.1)	10.62	0.005
Self-care	11 (2.4)	4 (0.9)	20 (4.8)	12.61	0.002
Usual activities	32 (7.1)	34 (7.7)	47 (11.2)	5.53	0.063
Pain or discomfort	146 (32.2)	173 (39.1)	139 (33.2)	5.41	0.067
Anxiety or depression	108 (23.8)	111 (25.1)	73 (17.4)	8.30	0.016

ENV: Environment; PHY: Physical; PSY: Psychological; SD: Standard deviation; SOC: Social Relationships; WHOQOL-BREF: World Health Organization Quality of Life

0.074. For the Malay-language version (n = 363), TLI = 0.909 and SRMR = 0.094. In the second-order CFA, the findings were similar to those of the 4-factor CFA. That is, TLI = 0.918, 0.913 and 0.907 for the English-, Chinese- and Malay-language versions, respectively while SRMR = 0.068, 0.074 and 0.096 for the English-, Chinese- and Malay-language versions, respectively. Factor loading exceeded 0.4 for all items in the 3 language versions. For example, in the English-language version, factor loading of items ranged from 0.465 to 0.891 for WHO-PHY; from 0.543 to 0.830 for WHO-PSY; from 0.819 to 0.873 for WHO-SOC and from 0.682 to 0.798 for WHO-ENV.

#### Known-Group Validity

The prevalence of chronic medical conditions ranged from 14.5% (mental illness) to 28.4% (high blood pressure) among those who completed the English-language version; from 18.1% (mental illness) to 38.6% (high blood pressure) among those who completed the Chinese-language version;

and 12.7% (mental illness) to 34.6% (high blood pressure) among those who completed the Malay-language version (Table 1). Apart from WHOQOL-BREF PSY (English) that did not differentiate between participants with and without heart disease ( $P > 0.05$ ), differences in mean scores between participants with and without a given chronic medical condition were statistically significant (Table 3).

Known-group validity of the WHOQOL-BREF domains with EQ-5D-5L was well supported with the differences in mean scores between known-groups exceeding the minimally important difference. Mean (SD) WHO-PHY score was 77.0 (13.81) for participants without problem (n = 404) versus 57.5 (17.71) for participants with any problem (n = 50) on EQ-MOB among those completing the English-language versions ( $P < 0.001$ ). The difference in mean scores between the two groups were 21 and 29 points with pooled standard deviations of 15.63 and 17.97 for those completing the Chinese- and Malay-language versions, respectively ( $P < 0.001$ ). Mean (SD) WHOQOL-BREF PSY score was 75.9 (11.69) for participants without problem (n = 346) versus 61.6 (15.48) for participants with any problem (n = 108) on EQ-AD among those completing the English-language versions ( $P < 0.001$ ). The difference in mean scores between the 2 groups were 17.5 and 14.4 points with pooled SD of 15.92 and 13.80 for those completing the Chinese- and Malay-language versions, respectively ( $P < 0.001$ ).

#### Internal Consistency

For internal consistency, WHOQOL-BREF met the group-level reliability criteria ( $> 0.7$ ), with Cronbach's alpha values ranging from 0.82 to 0.86 across the 4 domains for the English-language version, from 0.73 to 0.85 for the Chinese-language version, and from 0.82 to 0.89 for the Malay-language version (Table 4).

Table 2. WHOQOL-BREF Domain Scores after Adjusting for Age, Gender, Educational Level and Housing Type

WHOQOL-BREF Domains	Regression Coefficients (95% Confidence Interval)			
	Chinese Language	P Value	Malay Language	P Value
PHY	-0.25	0.812	0.23	0.833
PSY	-3.69	<0.001	0.14	0.894
SOC	-3.53	<0.001	1.44	0.157
ENV	-2.12	0.022	-0.92	0.337

ENV: Environment; PHY: Physical Health; PSY: Psychological; SOC: Social Relationships; WHOQOL-BREF: World Health Organization Quality of Life

Table 3. Known-Group Validity of WHOQOL-BREF Domains Based on Co-Existing Chronic Medical Conditions by Language Versions

English Language	Has Diabetes (n = 123)	Has High Blood Pressure (n = 129)	Has High Blood Cholesterol (n = 84)	Has Heart Disease (n = 128)	Has Mental Illness (n = 66)
WHOQOL-BREF Physical Health					
Mean (SD) domain scores	62.7 (16.87)	63.2 (16.20)	64.6 (16.14)	68.5 (15.06)	60.4 (15.60)
Mean (SE) difference	-16.6 (1.45) <sup>§</sup>	-16.2 (1.43) <sup>§</sup>	-12.6 (1.79) <sup>§</sup>	-8.8 (1.57) <sup>§</sup>	-16.8 (1.91) <sup>§</sup>
WHOQOL-BREF Psychological					
Mean (SD) domain scores	66.4 (18.14)	65.7 (16.62)	66.6 (19.02)	71.4 (15.19)	63.6 (18.86)
Mean (SE) difference	-8.4 (1.44) <sup>‡</sup>	-9.6 (1.40) <sup>§</sup>	-7.2 (1.67) <sup>*</sup>	-1.6 (1.47)	-10.4 (1.81) <sup>§</sup>
WHOQOL-BREF Social Relationships					
Mean (SD) domain scores	66.2 (18.59)	65.6 (17.45)	68.4 (19.24)	70.0 (16.24)	67.4 (19.51)
Mean (SE) difference	-10.6 (1.63) <sup>§</sup>	-11.6 (1.59) <sup>§</sup>	-6.8 (1.92) <sup>*</sup>	-5.5 (1.67) <sup>*</sup>	-7.6 (2.12) <sup>*</sup>
WHOQOL-BREF Environment					
Mean (SD) domain scores	66.5 (16.30)	65.2 (14.54)	67.6 (17.86)	69.7 (15.34)	65.3 (16.33)
Mean (SE) difference	-7.5 (1.45) <sup>*</sup>	-9.4 (1.40) <sup>§</sup>	-5.3 (1.70) <sup>†</sup>	-3.2 (1.47) <sup>‡</sup>	-7.7 (1.84) <sup>*</sup>
Chinese Language	Has Diabetes (n = 120)	Has High Blood Pressure (n = 171)	Has High Blood Cholesterol (n = 118)	Has Heart Disease (n = 121)	Has Mental Illness (n = 80)
WHOQOL-BREF Physical Health					
Mean (SD) domain scores	63.6 (14.20)	64.4 (16.35)	62.5 (15.21)	63.1 (14.49)	61.5 (16.79)
Mean (SE) difference	-11.42 (1.58) <sup>§</sup>	-12.1 (1.41) <sup>§</sup>	-12.6 (1.57) <sup>§</sup>	-12.0 (1.57) <sup>§</sup>	-12.5 (1.84) <sup>§</sup>
WHOQOL-BREF Psychological					
Mean (SD) domain scores	60.2 (17.41)	62.2 (16.94)	60.0 (16.07)	61.5 (16.54)	62.5 (15.53)
Mean (SE) difference	-9.7 (1.63) <sup>§</sup>	-8.2 (1.50) <sup>§</sup>	-9.7 (1.65) <sup>§</sup>	7.9 (1.66) <sup>*</sup>	-5.7 (1.95) <sup>§</sup>
WHOQOL-BREF Social Relationships					
Mean (SD) domain scores	64.1 (14.896)	66.2 (14.40)	64.4 (13.51)	64.5 (14.03)	66.0 (13.67)
Mean (SE) difference	-6.8 (1.39) <sup>§</sup>	-4.6 (1.29) <sup>*</sup>	-6.3 (1.41) <sup>*</sup>	-6.3 (1.38) <sup>*</sup>	-3.6 (1.64) <sup>‡</sup>
WHOQOL-BREF Environment					
Mean (SD) domain scores	63.0 (15.13)	64.2 (15.70)	61.5 (15.18)	62.2 (15.16)	63.9 (14.27)
Mean (SE) difference	-7.4 (1.52) <sup>*</sup>	-6.7 (1.39) <sup>§</sup>	-9.2 (1.51) <sup>§</sup>	-8.4 (1.51) <sup>§</sup>	-5.3 (1.79) <sup>†</sup>
Malay Language	Has Diabetes (n = 119)	Has High Blood Pressure (n = 145)	Has High Blood Cholesterol (n = 94)	Has Heart Disease (n = 113)	Has Mental Illness (n = 53)
WHOQOL-BREF Physical Health					
Mean (SD) domain scores	108 (23.8)	111 (25.1)	73 (17.4)	8.30	0.016
Mean (SD) domain scores	58.5 (18.51)	60.0 (18.22)	56.4 (18.25)	56.9 (17.90)	54.2 (16.91)
Mean (SE) difference	-20.1 (1.68) <sup>§</sup>	-19.7 (1.58) <sup>§</sup>	-21.2 (1.83) <sup>§</sup>	-21.9 (1.66) <sup>§</sup>	-21.4 (2.43) <sup>§</sup>
WHOQOL-BREF Psychological					
Mean (SD) domain scores	64.4 (16.58)	64.7 (14.65)	63.7 (15.34)	62.3 (15.21)	60.1 (15.14)
Mean (SE) difference	-9.3 (1.43) <sup>§</sup>	-9.7 (1.34) <sup>§</sup>	-9.5 (1.55) <sup>§</sup>	-11.9 (1.40) <sup>§</sup>	-12.5 (1.94) <sup>§</sup>

SD: Standard deviation; SE: Standard error; WHOQOL-BREF: World Health Organization Quality of Life

\**P* < 0.001.†*P* < 0.01.‡*P* < 0.05.

§Difference is minimally important (i.e. exceeds half a standard deviation).

Table 3. Known-Group Validity of WHOQOL-BREF Domains Based on Co-Existing Chronic Medical Conditions by Language Versions (Cont'd)

Malay Language	Has Diabetes (n = 119)	Has High Blood Pressure (n = 145)	Has High Blood Cholesterol (n = 94)	Has Heart Disease (n = 113)	Has Mental Illness (n = 53)
WHOQOL-BREF Social Relationships					
Mean (SD) domain scores	68.4 (15.75)	68.4 (13.83)	67.1 (14.87)	66.7 (14.60)	68.8 (15.92)
Mean (SE) difference	-8.0 (1.48)*	-8.7 (1.39)*§	-9.0 (1.60)*§	-10.2 (1.48)*§	-6.1 (2.06)†
WHOQOL-BREF Environment					
Mean (SD) domain scores	63.9 (13.80)	63.2 (12.66)	62.4 (12.06)	61.9 (12.05)	60.8 (12.34)
Mean (SE) difference	-7.0 (1.31)*§	-8.8 (1.21)*§	-8.5 (1.41)*§	-9.6 (1.30)*§	-9.3 (1.78)*§

SD: Standard deviation; SE: Standard error; WHOQOL-BREF: World Health Organization Quality of Life

\* $P < 0.001$ .

† $P < 0.01$ .

‡ $P < 0.05$ .

§Difference is minimally important (i.e. exceeds half a standard deviation).

### Test-Retest Reliability

The number of participants who completed the retest in English, Chinese and Malay was 60, 49 and 45, respectively. The ICC ranged from 0.58 (fair; PSY) to 0.83 (excellent; ENV) in the English-language version (Table 4). In the Chinese-language version, ICC ranged from 0.64 (good; SR) to 0.88 (excellent; PHY). For the Malay-language version, ICC ranged from 0.58 (fair; SR) to 0.69 (good; ENV).

### Discussion

This study is the first in Singapore to evaluate the WHOQOL-BREF in the English-, Chinese- and Malay-language versions. It is also one of the few studies evaluating the test-retest reliability of the WHOQOL-BREF. Two aspects of reliability were assessed: internal consistency and test-retest reliability. The WHOQOL-BREF performed well in both aspects. Where validity is concerned, the WHOQOL-BREF demonstrates known-group validity, being able to discriminate between known groups defined by chronic medical conditions as well as by responses to EQ-MOB and EQ-AD. However, results of the confirmatory factor analysis suggest that model fit does not meet the threshold. Hu and Bentler<sup>23</sup> suggested that a two-index presentation strategy is more effective than a single-index presentation strategy in rejecting reasonable proportions of various types of true-population and mis-specified models. On the basis of this two-index presentation strategy, the factor structure of the WHOQOL-BREF is not supported even though the threshold for good model fit is met with SRMR but not TLI. Seeing as the outcomes from SRMR and TLI did not agree with each other, it is probably due to the nature of model mis-specification. It has been suggested that SRMR is better at detecting model mis-specification related to factor covariance whereas TLI is better at detecting model mis-specification related to factor loadings.<sup>23</sup>

The findings of this study should be interpreted in light of its limitations. First, given that the true QoL may have changed during the 2-week interval, we may have underestimated the true test-retest reliability in this study, the extent to which is not known. Second, the WHOQOL-BREF is designed to be self-administered. We have, however, administered the questionnaire through face-to-face interviews. While we know that different modes of

Table 4. Internal Consistency and Test-Retest Reliability of the English-, Chinese- and Malay-Language Versions of the WHOQOL-BREF

	Cronbach's $\alpha$	Intraclass Correlation Coefficient* (95% CI)
English-Language	n = 454	n = 60
Physical Health	0.83	0.62 (0.37 to 0.77)
Psychological	0.85	0.58 (-0.13 to 0.82)
Social Relationships	0.82	0.68 (0.46 to 0.81)
Environment	0.86	0.83 (0.72 to 0.90)
Chinese-Language	n = 443	n = 49
Physical Health	0.82	0.88 (0.79 to 0.93)
Psychological	0.85	0.67 (0.14 to 0.85)
Social Relationships	0.73	0.64 (0.37 to 0.80)
Environment	0.83	0.81 (0.66 to 0.89)
Malay-Language	n = 419	n = 45
Physical Health	0.89	0.68 (0.46 to 0.81)
Psychological	0.82	0.58 (0.06 to 0.79)
Social Relationships	0.83	0.66 (0.38 to 0.81)
Environment	0.82	0.69 (0.44 to 0.83)

CI: Confidence interval; WHOQOL-BREF: World Health Organization Quality of Life

\* $< 0.40$ : poor;  $0.40$  to  $0.59$ : fair;  $0.60$  to  $0.74$ : good; and  $0.75$  and above: excellent.

All values are expressed to 2 significant figures.

administration may elicit different degrees of true responses (e.g. respondents tend to report socially desirable behaviours in face-to-face interviews),<sup>27</sup> the bias is consistent across all the questionnaires and does not affect the accuracy of the psychometric analyses which are largely comparative in nature. In addition, we have yet to examine measurement invariance of the WHOQOL-BREF. For example, the Taiwanese-Chinese version of the WHOQOL-BREF was reported to be equivalent between men and women and across various educational levels.<sup>28</sup> This should be undertaken as part of future work. Alternative analysis such as Rasch analysis may also be explored in future work.<sup>29</sup>

## Conclusion

Based on our findings, the WHOQOL-BREF could be a useful tool for assessing HRQoL in Singapore, showing good known-group validity, high internal consistency and test-retest reliability.

## Acknowledgement

*The authors would like to acknowledge the contribution of the late Professor Michael Power from the Department of Psychology, National University of Singapore, who provided critical feedback on an early draft of this manuscript. They would also like to acknowledge the support of the interviewers and clinic staff at the National Heart Centre Singapore and the Division of Endocrinology, National University Hospital during data collection. They would also like to acknowledge the support of research staff who helped to put together the ethics approval application forms.*

*The authors declare that they have no competing interests. This study was funded by the Ministry of Health Health Services Research Competitive Research Grant (HSRG-0045-2013). The study sponsor does not have a role in the study design, the collection, analysis, and interpretation of data, the writing of the article and the decision to submit it for publication. All researchers have access to all the data. Anonymised patient level data used in this analysis are available from the corresponding author upon the submission of a specific research question. To avoid duplication, please check against the list of research questions that the authors or other interested parties have planned. This is posted at [http://pharmacy.nus.edu.sg/profile\\_wwehweelin/#1481339201570-f200c061-8db6](http://pharmacy.nus.edu.sg/profile_wwehweelin/#1481339201570-f200c061-8db6).*

*In the informed consent form, patients were informed that data collected are the property of the National University Hospital and National Heart Centre Singapore. In the event of any publication regarding this study, their identity will remain confidential.*

## REFERENCES

1. Torrance GW. Utility approach to measuring health-related quality of life. *J Chronic Dis* 1987;40:593-603.
2. McHorney CA, Ware JE, Jr., Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care* 1994;32:40-66.
3. Development of the World Health Organization WHOQOL-BREF quality of life assessment. The WHOQOL Group. *Psychol Med* 1998;28:551-8.
4. Skevington SM, Lotfy M, O'Connell KA, WHOQOL Group. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. *Qual Life Res* 2004;13:299-310.
5. Ohaeri JU, Awadalla AW, El-Abassi AH, Jacob A. Confirmatory factor analytical study of the WHOQOL-Bref: experience with Sudanese general population and psychiatric samples. *BMC Med Res Methodol* 2007;7:37.
6. Colbourn T, Masache G, Skordis-Worrall J. Development, reliability and validity of the Chichewa WHOQOL-BREF in adults in Lilongwe, Malawi. *BMC Res Notes* 2012;3:346.
7. Cruz LN, Polanczyk CA, Camey SA, Hoffmann JF, Fleck MP. Quality of life in Brazil: normative values for the WHOQOL-bref in a southern general population sample. *Qual Life Res* 2011;20:1123-9.
8. Trompenaars FJ, Masthoff ED, Van Heck GL, Hodiament PP, De Vries J. Content validity, construct validity, and reliability of the WHOQOL-Bref in a population of Dutch adult psychiatric outpatients. *Qual Life Res* 2005;14:151-60.
9. Hanestad BR, Rustøen T, Knudsen O, Jr., Lerdal A, Wahl AK. Psychometric properties of the WHOQOL-BREF questionnaire for the Norwegian general population. *J Nurs Meas* 2004;12:147-59.
10. Noerholm V, Groenvold M, Watt T, Bjorner JB, Rasmussen NA, Bech P. Quality of life in the Danish general population--normative data and validity of WHOQOL-BREF using Rasch and item response theory models. *Qual Life Res* 2004;13:531-40.
11. Krägeloh CU, Kersten P, Rex Billington D, Hsu PH, Shepherd D, Landon J, et al. Validation of the WHOQOL-BREF quality of life questionnaire for general use in New Zealand: confirmatory factor analysis and Rasch analysis. *Qual Life Res* 2013;22:1451-7.
12. Xia P, Li N, Hau KT, Liu C, Lu Y. Quality of life of Chinese urban community residents: a psychometric study of the mainland Chinese version of the WHOQOL-BREF. *BMC Med Res Methodol* 2012;12:37.
13. Sreedevi A, Cherkil S, Kuttikkattu DS, Kamalamma L, Oldenburg B. Validation of WHOQOL-BREF in Malayalam and Determinants of Quality of Life Among People With Type 2 Diabetes in Kerala, India. *Asia Pac J Public Health* 2016;28:62S-9S.
14. Leung KF, Wong WW, Tay MS, Chu MM, Ng SS. Development and validation of the interview version of the Hong Kong Chinese WHOQOL-BREF. *Qual Life Res* 2005;14:1413-9.
15. Yao G, Chung CW, Yu CF, Wang JD. Development and verification of validity and reliability of the WHOQOL-BREF Taiwan version. *J Formos Med Assoc* 2002;101:342-51.
16. Min SK, Kim KI, Lee CI, Jung YC, Suh SY, Kim DK. Development of the Korean versions of WHO Quality of Life scale and WHOQOL-BREF. *Qual Life Res* 2002;11:593-600.
17. Benitez-Borrego S, Guardia-Olmos J, Urzua-Morales A. Factorial structural analysis of the Spanish version of WHOQOL-BREF: an exploratory structural equation model study. *Qual Life Res* 2014;23:2205-12.
18. Sim K, Chan YH, Chua TH, Mahendran R, Chong SA, McGorry P. Physical comorbidity, insight, quality of life and global functioning in first episode schizophrenia: a 24-month, longitudinal outcome study. *Schizophr Res* 2006;88:82-9.

19. Tan LSM, Dianne CTB, Khoo YHE, Griva K, New M, Lee YS, et al. WHOQOL-BREF among Singaporean Patients with Type II Diabetes Mellitus: What Does It Measure? *British Journal of Medicine & Medical Research* 2014;4:3293-316.
  20. Hasanah CI, Naing L, Rahman AR. World Health Organization Quality of Life Assessment: brief version in Bahasa Malaysia. *Med J Malaysia* 2003;58:79-88.
  21. Janssen MF, Pickard AS, Golicki D, Gudex C, Niewada M, Scalone L, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Qual Life Res* 2013;22:1717-27.
  22. Shiroiwa T, Fukuda T, Ikeda S, Igarashi A, Noto S, Saito S, et al. Japanese population norms for preference-based measures: EQ-5D-3L, EQ-5D-5L, and SF-6D. *Qual Life Res* 2016;25:707-19.
  23. Hu LT, Bentler PM. Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives. *Structural Equation Modeling* 1999;6:1-55.
  24. Norman GR, Sloan JA, Wyrwich KW. The truly remarkable universality of half a standard deviation: confirmation through another look. *Expert Rev Pharmacoecon Outcomes Res* 2004;4:581-5.
  25. Kline P. *A handbook of test construction: introduction to psychometric design*. London; New York: Methuen; 1986. p. 1-23.
  26. Cicchetti DV, Sparrow SA. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *Am J Ment Defic* 1981;86:127-37.
  27. Cheung YB, Goh C, Thumboo J, Khoo KS, Wee J. Quality of life scores differed according to mode of administration in a review of three major oncology questionnaires. *J Clin Epidemiol* 2006;59:185-91.
  28. Lin CY, Li YP, Lin SI, Chen CH. Measurement equivalence across gender and education in the WHOQOL-BREF for community-dwelling elderly Taiwanese. *Int Psychogeriatr* 2016;28:1375-82.
  29. Lin CY, Yang SC, Lai WW, Su WC, Wang JD. Rasch models suggested the satisfactory psychometric properties of the World Health Organization Quality of Life-Brief among lung cancer patients. *J Health Psychol* 2017;22:397-408.
-