Development of the World Health Organization WHOQOL-BREF Quality of Life Assessment

THE WHOQOL GROUP[1–3]

ABSTRACT

Background. The paper reports on the development of the WHOQOL-BREF, an abbreviated version of the WHOQOL-100 quality of life assessment.

Method. The WHOQOL-BREF was derived from data collected using the WHOQOL-100. It produces scores for four domains related to quality of life: physical health, psychological, social relationships and environment. It also includes one facet on overall quality of life and general health.

Results. Domain scores produced by the WHOQOL-BREF correlate highly (0·89 or above) with WHOQOL-100 domain scores (calculated on a four domain structure). WHOQOL-BREF domain scores demonstrated good discriminant validity, content validity, internal consistency and test–retest reliability.

Conclusion. These data suggest that the WHOQOL-BREF provides a valid and reliable alternative to the assessment of domain profiles using the WHOQOL-100. It is envisaged that the WHOQOL-BREF will be most useful in studies that require a brief assessment of quality of life, for example, in large epidemiological studies and clinical trials where quality of life is of interest. In addition, the WHOQOL-BREF may be of use to health professionals in the assessment and evaluation of treatment efficacy.

INTRODUCTION

Quality of life is defined by the World Health Organization Quality of Life (WHOQOL) Group as individuals' perceptions of their position 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. This definition reflects the view that quality of life refers to a subjective evaluation

that is embedded in a cultural, social and environmental context.

The WHOQOL-100 assessment was developed by the WHOQOL Group in 15 international field centres, simultaneously, in an attempt to develop a quality of life assessment that would be applicable cross-culturally. The development of the WHOQOL-100 has been detailed elsewhere (i.e. Orley & Kuyken, 1994; WHOQOL Group 1994*a*, *b*, 1995; Szabo, 1996). In brief, development involved the participation of all 15 field centres in deciding facets of life that were important in the assessment of quality of life, operationalizing facet definitions and contributing items for inclusion within a pilot version assessment. The original pilot version of the WHOQOL included 236 items relating to quality of life. Fifteen field centres piloted this assessment on at least 300 people with a range of health problems. From these data, 100 items were selected for inclusion in a revised version of the assessment: the WHOQOL-100 field trial version.

Table 1. *WHOQOL-BREF domains of quality of life: overall quality of life and general health*

Domain	Facets incorporated within domains
1 Physical health	Pain and discomfort Sleep and rest Energy and fatigue Mobility Activities of daily living Dependence on medicinal substances and medical aids Work capacity
2 Psychological	Positive feelings Thinking, learning, memory and concentration Self-esteem Bodily image and appearance Negative feelings Spirituality/religion/personal beliefs
3 Social relationships	Personal relationships Social support Sexual activity
4 Environment	Freedom, physical safety and security Home environment Financial resources Health and social care: accessibility and quality Opportunities for acquiring new information and skills Participation in and opportunities for recreation/leisure activity Physical environment (pollution/noise/traffic/climate) Transport

The WHOQOL-100 encompasses 24 facets universally regarded by all 15 field centres as important in assessing quality of life, and four general questions that address overall quality of life and health. Four questions regarding each facet are included. Recent analysis of available data has shown that these 24 facets can be most appropriately grouped into four domains: physical, psychological, social relationships and environment (see Table 1). For a more detailed explanation of this analysis, the reader is referred to The WHOQOL Group (1998). All domain scores relating to the WHOQOL-100 in the present paper calculate domains based on a four domain structure, although for the time being WHOQOL-100 data reported elsewhere will continue with the six domain structure.

While the WHOQOL-100 allows a detailed assessment of individual facets relating to quality of life, it may be too lengthy for some uses, for example in large epidemiological studies where quality of life is only one variable of interest. In these instances, assessments will be more willingly incorporated into studies if they are brief, convenient and accurate (Berwick *et al.* 1991). The WHOQOL-BREF Field Trial Version has, therefore, been developed to look at domain level profiles, which assess quality of life.

At a conceptual level, it was agreed by the WHOQOL Group that comprehensiveness ought to be maintained in any abbreviated version of the WHOQOL-100, by selecting at least one question from each of the 24 facets relating to quality of life. Decisions regarding the selection of items for the WHOQOL-BREF were based on the following criteria.

(*i*) Items selected to represent a particular domain should explain a large proportion of variance within that domain.

(*ii*) Items included should explain a substantial proportion of variance within the general facet relating to Overall Quality of Life and General Health perceptions.

(*iii*) The final assessment should demonstrate structural integrity in terms of confirmatory factor analysis.

(*iv*) The final assessment should be able to discriminate between identified groups of subjects (i.e. ill *versus* well subjects).

Data from 20 field centres situated within 18 countries were used to select items for these purposes (see Table 2). One item from each of

Table 2. *Centres included in development of the WHOQOL-BREF*

Data from original pilot of the WHOQOL	N	Data from original centres field testing the WHOQOL-100	N	Data from new centres field testing the WHOQOL-100	N
Bangkok, Thailand	300	Bangkok, Thailand	435	Hong Kong	856
Beer Sheva, Israel	344	Beer Sheva, Israel	464	Leipzig, Germany	527
Madras, India	412	Madras, India	567	Mannheim, Germany	483
Melbourne, Australia	300	Melbourne, Australia	350	La Plata, Argentina	421
New Delhi, India	304	New Delhi, India	82	Port Alegre, Brazil	82
Panama City, Panama	300	Panama City, Panama	117		
Seattle, USA	300	Seattle, USA	192		
Tilburg, The Netherlands	411	Tilburg, The Netherlands	799		
Zagreb, Croatia	300	Zagreb, Croatia	96		
Tokyo, Japan	286	Tokyo, Japan	190		
Harare, Zimbabwe	300	Harare, Zimbabwe	149		
Barcelona, Spain	303	Barcelona, Spain	558		
Bath, England	319	Bath, England	105		
St Petersburg, Russia	300				
Paris, France	323				
Total	4802		4104		2369

the 24 facets contained in the WHOQOL-100 has been included. In addition, two items from the Overall Quality of Life and General Health facet have been included. The WHOQOL-BREF therefore contains a total of 26 questions.

METHOD
The sample
Two data sets were used to select items for inclusion in the WHOQOL-BREF. The first included all data from 15 field centres who participated in the WHOQOL pilot study (see Table 2). This pilot assessment contained 236 questions relating to quality of life. The second dataset included data from the 13 centres who field-tested the WHOQOL-100. A further dataset including data from five new centres, who had not participated at the pilot stage but had field-tested the WHOQOL-100, and had results available, was also used to test the adequacy of items selected.

Procedure
The procedure followed to field-test the WHOQOL and the WHOQOL-100 for new centres was similar. The instrument was piloted on approximately 300 people in each field centre. The sample of respondents to whom the assessment was administered were adults, with 'adult' being culturally defined. The sampling quota applied with regard to: age ($50\% = < 45$ years, $50\% = +45$ years); sex ($50\% = $ male,

$50\% = $ female); and health status (250 persons with disease or impairment and 50 well persons).

This enabled the WHOQOL-BREF to be assessed in several populations. With respect to persons with disease or impairment, this group included patients from a cross-section of primary-care settings, hospitals and community-care settings. Within the new field centres dataset, some centres collected more than the required 300 respondents, with approximately 37% of the total sample consisting of subjects with no health problems.

The procedure followed by centres field testing the WHOQOL-100 who had already participated in the development of the WHOQOL differed. In some of these centres, data were collected from specific populations, dependent on the area of interest of the investigators collecting the data. For example, Barcelona collected a substantial portion of their data from patients diagnosed with schizophrenia, whereas Madras collected data from patients with cataracts, diabetes or cancer. By contrast, a large proportion of data from Seattle, Panama, Tilburg, Zagreb, Bath and Beer-Sheva was collected from healthy subjects. Hence, in this dataset approximately 45% of subjects had no health problems.

Selection of items from the WHOQOL-BREF
As noted in the Introduction, it was agreed that one item from each of the 24 facets ought to be included in an abbreviated version of the

WHOQOL-100. The most general question from each facet (i.e. the item that correlated most highly with the total score, calculated as the mean of all facets) was chosen for inclusion in the WHOQOL-BREF. Individual items selected by this method were then examined by a panel to establish whether the items selected to represent each domain reflected the conceptually derived operationalization of facets of quality of life. That is to say, they constituted a cohesive and interpretable domain, with good construct validity. Of the 24 items selected, six were substituted. Three items from the environmental domain were substituted because they were highly correlated with the psychological domain. A further three items were substituted because it was felt that other items within the facet could better explain the concept.

Calculation of domain scores

Domain scores for the WHOQOL-100 are calculated by taking the mean of all facet scores included in each domain and multiplying by a factor of four. Domain scores for the WHOQOL-BREF were calculated by multiplying the mean of all items included within the domain by four. Potential scores for all domain scores, therefore, range from 4–20. For the following analysis, the WHOQOL-100 is based on a four domain solution, so as to be comparable with the WHOQOL-BREF domain scores.

Statistical analysis

Data analyses were carried out using SPSS (Windows) Version 7. Internal consistency of domains was assessed using Cronbach alpha. Discriminant validity was determined via *t* tests to distinguish differences between ill and well subjects. Test–retest reliability was assessed using Pearson *r* correlations. Contribution of domain scores to assessing quality of life was assessed using multiple regression. Confirmatory factor analysis of the items included in the WHOQOL-BREF was carried out using the EQS package Version 5.0 (Bentler & Wu, 1995).

RESULTS

Descriptive statistics

Domain scores calculated using the WHOQOL-100 and the WHOQOL-BREF were very similar.

Differences between domain scores based on either assessment ranged from 0 to 0·27 (mean difference = 0·11). As expected from these small differences, there were high correlations between domain scores based on the WHOQOL-100 and domain scores calculated using items included in the WHOQOL-BREF. These correlations ranged from 0·89 (for domain 3) to 0·95 (for domain 1).

Internal consistency

Cronbach alpha values for each of the four domain scores ranged from 0·66 (for domain 3) to 0·84 (for domain 1), demonstrating good internal consistency (see Table 3). Cronbach alpha values for domain 3 should be read with caution as they were based on three scores (i.e. the personal relationships, social support and sexual activity facets), rather than the minimum four generally recommended for assessing internal reliability.

Discriminant validity

The WHOQOL-100 has previously been shown to have excellent ability in discriminating between ill and well respondents (see Table 4). The WHOQOL-BREF was shown to be comparable to the WHOQOL-100 in discriminating between these subject groups, with similar values and significant differences between ill and well subjects apparent in all domains (see Table 4).

Table 3. *Internal consistency of the WHOQOL-100 and the WHOQOL-BREF domains*

	Cronbach alpha		
	Original data (N = 4802)	Field data (N = 3882)	New data (N = 2369)
Physical health			
WHOQOL-100	0·88	0·87	0·86
WHOQOL-BREF	0·82	0·84	0·80
Psychological			
WHOQOL-100	0·79	0·79	0·82
WHOQOL-BREF	0·75	0·77	0·76
Social relationships			
WHOQOL-100*	0·72	0·72	0·73
WHOQOL-BREF*	0·66	0·69	0·66
Environment			
WHOQOL-100	0·85	0·85	0·85
WHOQOL-BREF	0·80	0·80	0·80

* Only 3 items, therefore Cronbach alphas may not be reliable.

Table 4. *Discriminant validity of the WHOQOL-100 and the WHOQOL-BREF*

| | Comparisons 'ill' *v.* 'well' *t* tests | | | | | |
| | Original data (N = 4802) | | Field data (N = 3882) | | New data (N = 2369) | |
	t	*P*	*t*	*P*	*t*	*P*
Physical health						
WHOQOL-100	34·1	0·001	39·1	0·001	25·4	0·001
WHOQOL-BREF	31·2	0·001	36·4	0·001	24·2	0·001
Psychological						
WHOQOL-100	13·0	0·001	24·0	0·001	10·3	0·001
WHOQOL-BREF	12·3	0·001	24·0	0·001	10·6	0·001
Social relationships						
WHOQOL-100	9·8	0·001	22·0	0·001	7·8	0·001
WHOQOL-BREF	8·4	0·001	16·2	0·001	6·9	0·001
Environment						
WHOQOL-100	7·0	0·001	21·7	0·001	2·4	0·02
WHOQOL-BREF	6·6	0·001	21·1	0·001	2·8	0·01

Test–retest reliability

Data used to assess test–retest reliability included a majority of well subjects (87% of respondents) from four centres participating in the field trial of the WHOQOL-100. These were Bath (N = 90), Harare (N = 100), Tilburg (N = 116) and Zagreb (N = 85). In all centres, respondents were university students, with the exception of Harare, where subjects were random samples of ill (N = 50) and well (N = 50) respondents. The interval between test and retest ranged from 2–8 weeks. Correlations between items at time points one and two were generally high, ranging from 0·56 for item 8 (How safe do you feel in your daily life?) to 0·84 for item 12 (Have you enough money to meet your needs?). The test–retest reliabilities for domains were 0·66 for physical health, 0·72 for psychological, 0·76 for social relationships and 0·87 for environment.

Confirmatory factor analysis

Confirmatory factor analysis of the WHOQOL-100 assessment, at facet level, suggested that a four domain solution may be most appropriate. This solution is shown in Fig. 1. Items relating to each domain therefore load onto that particular domain. The four domains then all load onto a second order factor, representing global quality of life. This four domain structure was tested using items in the WHOQOL-BREF. In both the dataset relating to the original pilot and the dataset relating to the field trial of the WHOQOL-100, an acceptable fit index (one measure of which is a Comparative Fit Index of 0·90 or higher) was achieved when the data was applied to the four domain structure (see Table 5).

In the dataset including new centres field testing the WHOQOL-100, the initial Comparative Fit Index (CFI) was 0·87, suggesting that alterations to the model were necessary. When three pairs of error variances were allowed to covary (i.e. pain and dependence on medication, pain and negative feelings, home and physical environment) and two items were allowed to cross-load on other domains (i.e. safety on the global domain and medication negatively on the environment domain), the comparative fit index increased to 0·901 (see Table 5).

Multisample analysis was subsequently undertaken to assess whether parameter estimates were similar across all three datasets. All parameter estimates were constrained to be equal across datasets, with the exception of two of the 24 items (item 4 and item 8), as these were known to cross-load on other domains in the case of the new centres dataset. In the multivariate model, the CFI reached 0·900, suggesting that the parameter estimates assessed were equivalent across all datasets.
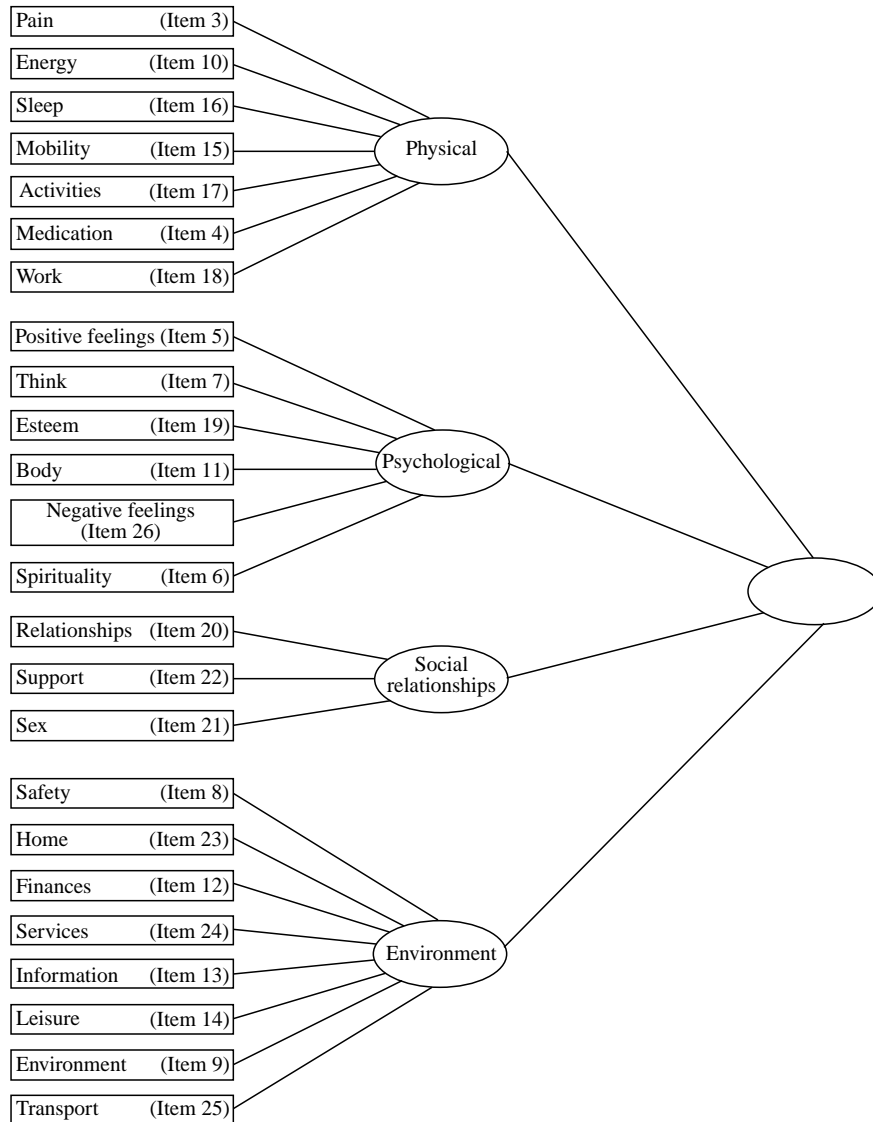
FIG. 1. Four domain confirmatory factor analysis model.

Importance of individual domains in assessing overall quality of life

Multiple regression was used to determine the contribution made by each domain score to explaining the observed variance in the general facet from the WHOQOL-100 assessment (i.e. the overall quality of life and general health facet). As shown in Table 5, all four WHOQOL-BREF domain scores made a significant contribution to explaining variance observed in the general facet relating to overall quality of life and general health, with the physical health domain contributing most highly, and the social relationships domain making least contribution. This suggests that all four domains should be taken into consideration when evaluating overall quality of life.

Table 5. *Comparative fit indices for the four domain model and multiple regression using general health and quality of life facet as the dependent variable and domain scores as independent variables*

Dataset	Comparative fit index	% of overall QOL and general health facet explained	Final equation standardized beta values*			
			Domain 1	Domain 2	Domain 3	Domain 4
Original	0·906	62·9	0·31	0·31	0·16	0·21
Field test	0·903	68·4	0·38	0·23	0·17	0·22
New data	0·901	61·5	0·33	0·29	0·13	0·20

* Significant at $P < 0.001$.

DISCUSSION

The WHOQOL-BREF has been shown to assess adequately domains relevant to quality of life in a large number of cultures worldwide. Domain scores produced by the WHOQOL-BREF have been shown to correlate at around 0·9 with the WHOQOL-100 domain scores, which has itself demonstrated criterion validity. They have also been shown to display good discriminant validity, content validity and test–retest reliability.

Although only one-quarter of the length of the WHOQOL-100, the WHOQOL-BREF incorporates good breadth and comprehensiveness by including items from each of the 24 facets of quality of life included in the longer form. Despite the heterogeneity of facets included within domains, all domains display excellent internal consistency.

As with other measures shortened in such a way, the fact that the WHOQOL-BREF uses a subset of items included within the WHOQOL-100 allows direct comparison between data collected from specific populations using either of the two assessments.

The WHOQOL-BREF remains slightly longer than some other short forms of quality of life assessments (e.g. the SF-12; see Ware *et al.* 1996), but encompasses a larger number of domains that are integral to the assessment of quality of life; notably the social relationships and environment domains that are not always included in other assessments. In research studies where only certain domains of quality of life are of interest, there is the option to include only those domains relevant to the study. However, as we have argued elsewhere, if quality of life is conceptualized as a multi-dimensional construct, each domain is seen as integral to an assessment of quality of life (The WHOQOL Group, 1994*a*, *b*).

It should be noted that analysis of the WHOQOL-BREF was based on taking 26 items from either 100 items, or in the case of data from the initial pilot WHOQOL, from 236 items. While the aim is now to collect and analyse data from the field trial version of the WHOQOL-BREF itself, we would predict similar results to emerge from this procedure. In addition, further field trials aim to address the responsiveness to change and concurrent validity of the WHOQOL-BREF.

The WHOQOL-BREF provides an adequate alternative to the assessment of domain profiles using the WHOQOL-100. It provides a rapid means of scoring domain profiles; it does not however allow assessment of the individual facets within these domains. A balance between detail and length of assessment will, therefore, always be important to consider when selecting between different WHOQOL assessments. It is envisaged that the WHOQOL-BREF will be used primarily in circumstances where a brief assessment of quality of life is appropriate, for example, in routine clinical work, large scale epidemiological studies and in clinical trials.

REFERENCES

Bentler, P. M. & Wu, E. J. C. (1995). *EQS for Windows Users Guide*. Encino, Multivariate Software, Inc.: CA.

Berwick, D. M., Murphy, J. M., Goldman, P. A., Ware, J. E., Barsky, A. J. & Weinstein, M. C. (1991). Performance on a five-item mental health screening test. *Medical Care* **29**, 169–176.

Orley, J. & Kuyken, W. (eds.) (1994) *Quality of Life Assessment: International Perspectives*. Springer Verlag: Heidelberg.

Szabo, S. on behalf of the WHOQOL Group (1996). The World Health Organization Quality of Life (WHOQOL) Assessment

Instrument. In *Quality of Life and Pharmaeconomics in Clinical Trials*, 2nd edn (ed. B. Spilker), pp. 355–362. Lippincott-Raven Publishers: Philadelphia, New York.

Ware, J. E., Kosinski, M. & Keller, S. D. (1996). A 12-item short-form health survey. *Medical Care* **34**, 220–228.

WHOQOL Group (1994 *a*). The development of the World Health Organization quality of life assessment instrument (the WHOQOL). In *Quality of Life Assessment: International Perspectives* (ed. J. Orley and W. Kuyken), pp. 41–57. Springer Verlag: Heidelberg.

WHOQOL Group (1994 *b*). Development of the WHOQOL: rationale and current status. *International Journal of Mental Health* **23**, 24–56.

WHOQOL Group (1995). The World Health Organization Quality of Life assessment (WHOQOL): position paper from the World Health Organization. *Social Science and Medicine* **41**, 1403–1409.

WHOQOL Group (1998). The World Health Organization Quality of Life Assessment (WHOQOL): development and general psychometric properties. *Social Science and Medicine* (in the press).