Expecting a good quality of life in health: assessing people with diverse diseases and conditions using the WHOQOL-BREF

Suzanne M. Skevington BSc PhD* and Farah M. McCrate BSc MSc†

*Professor, Department of Psychology, WHO Centre for the Study of Quality of Life and †Research Officer, Department of Psychology, WHO Centre for the Study of Quality of Life, University of Bath, Bath, UK

Correspondence
Suzanne M. Skevington
WHO Centre for the Study of Quality of Life
University of Bath
Bath BA2 7AY
UK
E-mail: s.m.skevington@bath.ac.uk

Abstract

Background and objectives Fulfilling patient expectations is central to defining a good quality of life (QoL) in health. The WHOQOL-BREF was developed using novel, person-centred methods and is a generic patient-reported outcomes measure (PROM). However, without robust psychometric performance, PROMs cannot be relied upon to assess individuals. This study investigated the WHOQOL-BREF (UK), with this use in mind.

Design Cross sectional with nested repeated measures.

Setting and participants Twenty-seven disease groups or health conditions and healthy people were recruited at 38 UK sites, in a wide range of settings ($n = 4628$).

Interventions 'Treatment as usual'; new and alternative interventions.

Outcome measures WHOQOL-BREF (UK); SF-36.

Results Respondent burden was low, as acceptability and feasibility were high. Internal consistency was excellent (0.92) and test–retest reliability good. Distinctive QoL profiles were found for diverse conditions. Musculoskeletal, psychiatric and cardiovascular patients reported the poorest QoL and also improved most during treatment. Overall, QoL was good, and best for healthy groups, supporting discriminant validity. Compared with the SF-36, WHOQOL physical and psychological domains showed good concurrent validity, although social was weak. Small or moderate effect sizes confirmed responsiveness to change in specified domains for certain conditions and interventions. Age had a small impact on reporting QoL.

Discussion and conclusion The WHOQOL-BREF is found to be a high quality patient-centred generic tool suited to individual assessment in clinics, for research, and audit.

Introduction

Health expectations are integral to the definition of quality of life (QoL) in health and health care. Calnan[1] defined QoL as narrowing the gap between a patient's hopes and expectations. This theme was developed by the World Health Organisation (WHO) in defining QoL as 'An individual's 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'.[2] The WHO definition places fulfilled expectations at the very centre of understandings about how patients decide whether they have a good QoL. Furthermore, cross-cultural data shows that people with the very poorest QoL also see this aspect of their QoL as very important.[3] A large aspirational gap on an issue that is very important can have a disenhancing effect on QoL that is already poor,[3] and this finding is commensurate with definitions above. Such findings highlight a vulnerable group whose expectations have not been fulfilled to the detriment of their QoL and health. Growing consensus among QoL researchers indicates that studying patient expectations is complex, but vital to conceptual and measurement developments.[4] Drawing on the WHOQOL definition, we take the view that health expectations are inherently important to assessing QoL, as they are an embedded component of it.

Recent discussions about designing outcome measures for individual use were prompted by the arrival of new high calibre QoL measures. Without precision instruments, scores from individual patients cannot be relied upon in the same way that groups of patients can be reliably assessed, as individual assessment requires more stringent criteria. Higher precision is needed to produce good patient-reported outcome measures (PROMs). Furthermore, these measures need to be developed by explicitly involving people who will ultimately answer them, rather than exclusive consultation with experts. The WHOQOL is one such instrument to fulfil this brief, and the first cross-cultural generic measure to actively incorporate the language and concepts of users into the measure.[2] This person-centred development process means that the WHOQOL-BREF fulfils perhaps *the* most important prerequisite for a good PROM, namely the involvement of users. It could also fulfil another vital criterion for one-to-one use in clinical and non-clinical settings if it was found to perform to robust psychometric standards for individual use. A main aim was therefore to test the properties of the WHOQOL-BREF, to ascertain whether it was suitable for individual assessment.

Improving the quality of audit is the current focus for using PROMs in UK, although the Department of Health may also use them to enhance clinical governance, support performance pay, increase patient choice and make purchasing decisions.[5] In NHS hospitals, generic and specific PROMs are being administered before and after elective surgery for hip and knee replacement, groin hernia repair and varicose veins. Treatment for six long-term conditions and two cancers will also be similarly evaluated. Consequently, our study aimed to further assess the performance of an internationally sound generic instrument in UK – the WHOQOL-BREF – is timely, relevant and very important within this policy framework.[6]

Several well-standardized generic scales assess subjective QoL across a spectrum of diseases making them amenable to audit and decision making about health-care priorities. In recent years, several clusters of new generic instruments have become available showing high performance on accepted psychometric criteria. Two clusters include widely used measures; the Short Form-36 (SF-36) health status scales[7] and the EuroQol.[8] However, concepts and items in these questionnaires were proposed primarily by health researchers and clinicians, not potential users. Mounting evidence shows that health professionals report a quite different subset of perspectives on a patient's QoL from that persons own subjective views, as they are unable to truly share their subjective experience or derive the same meaning. This therefore raises important questions about the accuracy of results obtained from measures that did not incorporate direct patient input, as they do not entirely

satisfy the prerequisite for patient-centred development that is quintessential to the best PROMs. Who is better able to say what are a persons/patients 'goals, expectations, standards, and concerns' than the person themselves?

A third cluster of measures developed at WHO by the WHOQOL Group designed international person-centred methods to address previous shortcomings. Through a 'bottom-up' process, qualitative information was pooled and integrated from focus groups of patients, health professionals and community members held simultaneously in 15 countries worldwide, including UK.[9] Agreed concepts and language structured a 'universal' framework of QoL in health, generating international translatable items in 15 languages.[10] Cross-cultural survey data from the same centres confirmed a comprehensive and holistic range of 25 internationally approved facets of QoL.[11,12] The WHOQOL has greater semantic and conceptual equivalence between language versions and better applicability and acceptability than before.[13] The WHOQOL-100[14] and its extracted short-form the WHOQOL-BREF[15] provide ideal tools to assess the outcomes of care, multinational clinical trials and cross-cultural research.[13]

The WHOQOL-BREF has been used to assess QoL in people with many diseases and conditions; HIV,[16] liver transplantation,[17] chronic fatigue syndrome (CFS),[18] depression,[19] psychosis,[20] obsessive-compulsive disorder,[21] older adults[22] and serious life events e.g. holocaust,[23] and earthquakes.[24] Although cross-cultural differences between patients with chronic diseases are documented, the SF-36,[25] systematic evidence of QoL in people with diverse diseases, conditions and health for the WHOQOL has not been gathered in a single national group, and providing this information was an important aim. Such information could be used to monitor progress in treatment, to compare different disease stages, and within the same or similar conditions. It could document how QoL in one condition compares to other physical or mental health conditions and health, and facilitate clinical decision making. It could be used to audit procedures and treatments, and incorporate a patient's perspective into

decisions about resource distribution. Through examining the impact of socio-demographic features on QoL as we propose, health inequalities can be accounted for, and the need for adjusted norms assessed. Furthermore, this study aimed to investigate additional psychometric properties of the WHOQOL-BREF (UK) using the largest UK data set so far available. Without information about whether scales are good measures, it is impossible to know whether clinical trials and treatments are truly effective, or whether their results are merely a consequence of measurement error. In demonstrating that the best psychometric properties exist (reliability, validity, sensitivity to change, interpretability; low respondent burden through high feasibility, acceptability and appropriateness) and appropriate cultural translations,[26] essential new information is provided to those who select measures, where formerly, only limited information was available. Lastly, as the working language of the WHOQOL Group is English, knowledge about the performance of the UK instrument is crucial to provide an international benchmark.

Method

Design

A cross-sectional sample was used to survey the QoL of sick and well populations by administering the WHOQOL-BREF; in some samples, simultaneously with the SF-36. Using a repeated-measures design, a nested subgroup of people provided assessments on two occasions, within clinical trials and longitudinal samples. The impact size of an intervention on QoL was assessed before and after 'treatment as usual', or a new/alternative health-care intervention. Data from 'healthy' people obtained twice without an intervening intervention or change, assessed the stability of scores for test–retest reliability.

Settings and participants

A total of 4669 adults were recruited in 38 sites ($n = 663$ to $n = 10$), following formal ethical approval locally. The sample contained 27 dif-

ferent physical or psychological diseases/conditions, and healthy people (see Tables 1 and 2). lived in England [south-west (14), south (5), Midlands (2), London (8)] Scotland (2), Wales (1) and Ireland (1). People were recruited in community, primary care, outpatient, inpatient, rehabilitation settings and social care.

Well people included six samples of university students and student nurses. Community samples contained elderly carers, health visitor clients, cocaine addicts, depressed, neurodis-

ability and dental patients. Older adults received interventions for dementia, social care and skin conditions. Patients from disfigurement clinics and awaiting elective cosmetic surgery were recruited. A prison health promotion scheme was assessed. Some patients with arthritis had skin diseases; others participated in trials of alternative treatments. Patients with diabetes were sampled in primary care; they were newly diagnosed, from self-help groups, or receiving treatment in contrasting settings. Several inpa-

Table 1 Comparing quality of life domains for 27 health conditions using the WHOQOL-BREF

Diseases and conditions	Category (n)	Physical mean (SD)	Psychological mean (SD)	Social mean (SD)	Environment mean (SD)	F	d.f.
Well	H (141)	75.41 (18.72)[1]	70.21 (15.84)	71.37 (19.82)	72.26 (14.78)	5.91*	2.37, 322.65
Students	H (241)	73.66 (14.70)[1,2,3]	63.71 (14.98)	67.91 (19.66)[4,5]	63.94 (12.76)	36.89*	2.35, 564.73
Nurses	H (663)	78.12 (15.47)[1,2,3]	68.84 (15.78)	72.27 (20.80)[4,5]	68.60 (14.06)	93.85*	2.35, 1550.05
Carers	H (61)	61.53 (20.87)	65.78 (14.56)	61.68 (20.34)	68.95 (14.67)[3,5]	4.48*	2.45, 147.25
Dental	H (222)	79.49 (13.97)[1,2,3]	68.28 (14.89)	70.00 (21.20)	68.84 (12.61)	46.09*	2.24, 489.56
Prisoners	P (388)	79.19 (14.54)[1,2,3]	70.89 (13.51)	74.21 (18.73)[4,5]	69.92 (13.77)	60.05*	2.59, 961.84
Chronic fatigue syndrome	L (25)	35.58 (18.31)	53.30 (14.79)[1]	58.00 (21.85)[2]	58.43 (12.37)[3]	10.32*	1.82, 40.02
Cocaine dependency	L (53)	59.03 (18.66)[1,2,3]	48.98 (18.54)	47.48 (23.38)	51.36 (16.92)	6.83*	2.36, 122.53
Sleep disorder	L (45)	57.65 (17.67)	60.70 (13.23)	66.85 (21.80)	72.29 (12.06)[3,6]	8.43*	2.32, 102.17
Disfigurement	PS (472)	67.10 (20.30)[1]	63.31 (18.38)	68.65 (21.54)[4]	70.42 (16.33)[3,6]	23.78*	2.56, 1182.94
Cosmetic surgery	PS (222)	71.34 (20.35)[1,2,3]	57.07 (21.78)	62.71 (26.33)[4]	64.95 (16.51)[6]	37.07*	2.57, 552.87
Depression	MH (50)	45.81 (18.70)[1]	31.78 (18.62)	57.98 (20.87)[2,4]	53.07 (13.89)[3,6]	31.56*	2.24, 103.08
Chronic schizophrenia	MH (10)	67.50 (15.36)	55.23 (18.27)	52.53 (25.82)	64.80 (15.34)	1.65	3.0, 27
Mild dementia	MH (20)	70.00 (14.63)	76.67 (13.81)[1]	75.83 (8.51)[2]	78.91 (9.23)[3]	5.14*	3.0, 57
Neurodegenerative disease	N (45)	54.84 (20.09)	57.31 (17.99)	62.59 (20.99)	67.78 (16.48)[3,6]	6.64*	2.40, 105.50
Arthroplasty surgery	MS (61)	41.34 (15.66)	61.31 (14.58)[1]	69.76 (16.70)[2,4,5]	59.33 (11.99)[3]	64.14*	2.38, 135.39
Arthritis	MS (71)	46.45 (20.91)	58.30 (16.33)[1]	69.48 (22.32)[2,4]	68.20 (15.96)[3,6]	39.63*	2.00, 139.99
Chronic pain	MS (374)	39.04 (20.86)	53.07 (18.68)[1]	59.34 (23.64)[2,4]	58.82 (17.90)[3,6]	165.58*	2.31, 812.28
Skin disorder	D (70)	63.36 (18.82)	68.49 (14.22)	71.70 (19.39)[2]	70.32 (12.71)[3]	6.25*	2.34, 149.97
Heart transplant	CV (40)	38.88 (21.15)	56.19 (23.18)[1]	54.27 (21.28)[2]	58.75 (17.12)[3]	15.46*	3.0, 114
Stroke	CV (19)	58.47 (17.95)	51.68 (17.52)	60.39 (19.46)	65.53 (17.44)[6]	3.80**	3.0, 51
Diabetes	E (524)	67.84 (19.55)	67.66 (16.10)	70.12 (19.66)[4]	71.08 (15.52)[3,6]	10.13*	2.49, 1274.50
Irritable bowel syndrome	GI (358)	57.17 (19.46)	54.90 (17.83)	59.33 (22.51)[4]	64.76 (16.55)[3,5,6]	34.34*	2.45, 865.99
Irritable bowel disorder	G (38)	60.79 (22.00)	61.12 (17.11)	68.13 (24.28)	64.91 (17.46)	2.42	2.25, 80.90
Crohn's disease	GI (117)	54.38 (19.99)	56.99 (19.42)	58.51 (25.07)	64.95 (17.28)[3,5,6]	13.79*	2.51, 288.61
Colitis	GI (153)	48.74 (22.12)	56.67 (19.30)[1]	58.17 (23.49)[2]	63.43 (16.77)[3,5,6]	24.71*	2.49, 368.58
Polycystic ovarian syndrome	UG (87)	66.20 (19.63)[1,2,3]	50.93 (20.33)	49.43 (26.62)	60.97 (15.13)[5,6]	23.62*	2.31, 196.52

H, healthy; P, prisoners; L, lifestyle; PS, plastic surgery; MH, psychiatry; N, neurology; MS, musculoskeletal; D, dermatology; CV, cardiovascular; E, endocrine; GI, gastrointestinal; UG, urinogenital.
*$P < 0.01$; **$P < 0.05$.
Significantly higher mean of:[1]Physical vs. psychological; [2]Physical vs. social; [3]Physical vs. environment; [4]Psychological vs. social; [5]Social vs. environment; [6]Psychological vs. environment.

Table 2 Means (and standard deviations) of WHOQOL-BREF domains for different conditions

	Physical health	Psychological	Social relationships	Environment	F	d.f.
Well (n = 1324–1328)	76.49 (16.19)[1,2,3]	67.82 (15.56)	70.52 (20.67)[4,5]	68.20 (13.81)	149.67*	2.42, 3190.51
Prisoners (n = 381–388)	79.19 (14.54)[1,2,3]	70.89 (13.51)	74.21 (18.73)[4,5]	69.92 (13.77)	60.05*	2.59, 961.84
Lifestyle (n = 121–123)	54.06 (20.20)	54.15 (16.76)	56.71 (23.95)	60.45 (17.12)[3,6]	4.10*	2.30, 275.66
Plastic surgery (n = 686–693)	68.46 (20.39)[1]	61.32 (19.73)	66.75 (23.33)[4]	68.67 (16.57)[5]	40.69*	2.57, 1739.77
Psychiatric (n = 77–80)	54.57 (20.62)[1]	45.93 (25.99)	61.91 (20.80)[4]	61.00 (17.02)[3,6]	18.58*	2.32, 176.41
Neurological (n = 45)	54.84 (20.09)	57.31 (17.99)	62.59 (20.99)	67.78 (16.48)[3,6]	6.64*	2.40, 105.49
Musculoskeletal (n = 493–506)	40.37 (20.44)	54.81 (18.13)[1]	62.03 (23.15)[2,4]	60.19 (17.31)[3,6]	248.73*	2.30, 1107.08
Dermatological (n = 67–70)	63.36 (18.82)	68.49 (14.22)	71.70 (19.39)[2]	70.32 (12.71)[3]	6.25*	2.34, 149.97
Cardiovascular (n = 57–59)	45.19 (22.05)	54.74 (21.47)[1]	56.20 (20.75)[2]	60.93 (17.37)[3,6]	11.59*	2.63, 147.52
Endocrine (n = 519–524)	67.84 (19.55)	67.66 (16.10)	70.12 (19.66)[4]	71.08 (15.52)[3,6]	10.13*	2.49, 1274.50
Gastrointestinal (n = 656–666)	54.99 (20.62)	56.03 (18.45)	59.42 (23.35)[2,4]	64.50 (16.75)[3,5,6]	62.98*	2.51, 1641.47
Urinogenital (n = 86–87)	66.20 (19.63)[1,2,3]	50.93 (20.33)	49.43 (26.62)	60.97 (15.13)[5,6]	23.62*	2.31, 196.52

*$P < 0.01$.
Significantly higher mean of: [1]Physical vs. psychological; [2]Physical vs. social; [3]Physical vs. environment; [4]Psychological vs. social; [5]Social vs. environment; [6]Psychological vs. environment.

tient neurodisability groups received rehabilitation, including speech therapy for stroke. Inpatient care was provided to schizophrenia and heart transplant surgery patients. Irritable bowel disorder (IBD) and patients with CFS mainly received primary care e.g. self-help. Clinic data was collected on sleep disorders and facial hirsuitism in polycystic ovarian syndrome (POS).

Procedure

Participants completed the WHOQOL-BREF (UK). In some samples, they also simultaneously completed the SF-36, enabling concurrent validity to be assessed. Most self-completed the measure(s) without assistance during a consultation. A minority received mailed questionnaires at baseline and/or follow-up. Educational and occupational settings facilitated group administration to healthy samples e.g. nurses.

Instruments

The WHOQOL-BREF is a QoL measure applied to health. It contains 26 items; 24 are scored in one of four QoL domains namely, physical health, psychological, social relationships and environment. Two overall health and QoL items are unscored; a mean provides an overarching, independent general facet. Data from 23 countries (n = 11 801) showed good internal consistency reliability and construct validity for the international WHOQOL-BREF.[15] The WHOQOL-100 containing 100 items shows good psychometric properties in UK[27] and has been validated for chronic pain,[28] depression[29] and psoriasis[30] populations. As all the WHOQOL-BREF items were extracted from the 25 facets of the WHOQOL-100, its validation was expected.

The SF-36 is a generic health status measure evaluating physical and mental health via eight subscales on physical functioning, role physical, bodily pain, general health, vitality, social functioning, role emotional and mental health. Extensive technical work on the US version shows very good psychometric properties.[7]

Analysis

Anonymized samples were checked, merged and assessed for normality (items and domains). Some clinically related diagnostic groups were combined after significance testing (Scheffe)

confirmed similar domain scores. Dummy variables of education and marital status were created by recoding: with and without primary education became 'low education'; married and living-as-married as 'partnered'. Age was banded into six decades (20–79 years), plus youngest (< 19 years) and oldest (> 80).

Repeated-measures ANOVA on QoL domains tested main effects and interactions, using Bonferroni comparisons for health conditions, summarized categories and individual socio-demographic variables. The contribution of socio-demographic characteristics to overall QoL was assessed using stepwise multiple regression with the general facet as the dependent variable. Malahanobis and Cook's distances, case-wise diagnostics and standardized DF beta values were examined, and multivariate outliers excluded. R^2 adjusted and significance of change (F) were recorded.

Additional psychometric properties of the WHOQOL-BREF (UK) were assessed. Test–retest reliability was examined using Pearson correlations (criterion > 0.5). Paired Student's t ($P < 0.05$ two-tailed) examined similarities and differences between domains for well samples at two time periods. Binary answers to the question 'Are you currently ill' distinguished 'known' groups of well and sick people,[31] so assessing discriminant validity (Student's t; two-tailed). Healthy people were expected to report better QoL. Concurrent validity was assessed by correlating dimensions of the WHOQOL-BREF with the SF-36 (criterion $P < 0.01$), substituting Spearman's rank correlations where variables lacked normality. WHOQOL general health, physical and psychological domains, and pain and energy facets, were expected to correlate positively with similar SF-36 concepts; environmental QoL was expected to correlate weakly. Repeated-measures data enabled a preliminary assessment of the responsiveness of WHOQOL-BREF scores to change. Effect size (Cohen's d) compared the impact of different interventions using thresholds for small (0.2), moderate (0.5) and large (0.8) effects;[32] moderate responsiveness is expected from generic domains.

Results

Sample

The sample contained 1446 men and 3006 women with ages ranging from 16 to 105 years (M = 44.5 years; SD = 14.4). Thirty-five percent had received secondary school education, 44% tertiary, 4% primary and 1.5% no education. Fifty percent were married, 8% living-as-married, 21% single, 2% separated, 5% divorced and 6% widowed. Sixty percent were ill, of which 50% were in community care, 19% primary care, 19% outpatients, 4% inpatients and 8% rehabilitation.

Overall, QoL was good (M = 3.5; SD = 0.9); 19% reported very good QoL, 47% good, 22% neither good nor poor QoL, 9% poor and 3% very poor QoL. Eleven percent were very satisfied with their health, 37% satisfied, 24% dissatisfied and 8% very dissatisfied.

Mean QoL domain scores for the total sample exceeded the midpoint of 50 (transformed 0–100) showing QoL in every domain was acceptable to good. Environmental QoL was best (M = 66.8; SD = 16), then social relationships (M = 66.5; SD = 22), physical (M = 65.2; SD = 22) and psychological (M = 62.6; SD = 18).

Acceptability and feasibility

Of 4669 who completed the WHOQOL-BREF, 4628 cases were analysed; 41 had missing data > 20%. There was little item data missing; most being for working capacity (3%) and sex life (8%). Moderate negative skew for five WHOQOL items was within acceptable limits. SF-36 distributions were abnormal for physical functioning, role physical, bodily pain, social functioning and role emotional.

Completion times for the WHOQOL-BREF ranged from 2 to 240 min. Subsample means ranged from 4.5 min (students) to 20 min (rehabilitation). Three percent needed assistance with completion. The findings demonstrate acceptability and feasibility of use.

Quality of life in illness and health

The QoL of different diseases and conditions is compared in Table 2 and summarized in Table 1. Substantial differences were found between groups for every domain (Physical $F_{26,4517} = 84.61$, $P < 0.001$; Psychological $F_{26,4523} = 31.45$, $P < 0.001$; Social $F_{26,4495} = 15.20$, $P < 0.001$; Environmental $F_{26,4541} = 15.73$, $P < 0.001$). *Post hoc* comparisons showed that most differences were for physical QoL, probably because of the high proportion of physical conditions sampled. Domain differences were confirmed for 24 of 27 conditions; not IBD, chronic schizophrenia or stroke ($P < 0.05$). This data shows distinctive profiles for many conditions, particularly depression, cocaine dependency, chronic pain, gastrointestinal disorders and POS.

Predictably, healthier students, nurses, dental patients and prisoners had the best physical QoL. Cocaine addicts, plastic surgery candidates, diabetes and POS patients reported fairly good physical QoL. Physical QoL was poorest and poor in musculoskeletal patients (arthritis/arthroplasty, chronic pain) and cardiovascular patients awaiting heart transplantation. QoL was acceptable for several chronic disease groups: gastrointestinal (Crohn's disease and colitis), chronic neurodegenerative diseases and CFS.

As expected, psychiatric patients, particularly those with depression, reported poor and poorest psychological QoL. Psychological QoL was relatively good for dermatology and endocrine patients, well people and prisoners but only acceptable in CFS, cocaine dependency, chronic pain, stroke and POS. Urinogenital patients and cocaine addicts reported the poorest social QoL; dermatology patients the best. Social QoL was barely acceptable for those with chronic schizophrenia and prior to heart transplantation. Few group differences were found for environmental QoL, being best for endocrine, dermatology and healthy groups, but barely acceptable in depression or addiction. Together these findings add validity to all WHOQOL-BREF domains.

Relations between socio-demographic features and QoL related to health

Significant main and interaction effects were expected for each socio-demographic variable with QoL domains.

Gender
QoL was good for both genders, as domain means ranged from 61 to 68. Although it was expected that women would report poorer QoL than men, this main effect was only marginal ($P = 0.052$). However, a significant interaction between QoL and gender was confirmed ($F_{2,10\,840} = 29.55$, $P < 0.001$), showing that women reported better social QoL than men but poorer QoL on other domains. Psychological QoL was poorest for women and the greatest area of gender inequality.

Education
QoL differences were found between educational levels ($F_{3,3820} = 36.26$, $P < 0.001$). An interaction ($F_{7,9579} = 26.26$, $P < 0.001$) showed that uneducated participants had the poorest QoL ($P < 0.001$), and their physical QoL was particularly unacceptable (M = 47.8). Educational level had least impact on environmental QoL.

Marital status
Differences in marital status ($F_{5,4153} = 31.79$, $P < 0.001$) showed that people with partners reported better psychological, social and environmental QoL than divorcees ($F_{12,10\,496} = 28.63$, $P < 0.001$). Divorced and separated people had barely acceptable QoL. Widowed participants had poorer QoL than married on every domain, but had better social and environmental QoL than singles. Those living-as-married and single (M = 68.3) had good physical QoL. Very good social QoL was reported by those with partners (M = 70.2).

Age
QoL was acceptable to good at all ages, although age-band differences were confirmed ($F_{7,3880} = 5.23$, $P < 0.001$). The interaction

showed that QoL was better for younger people (20s and 30s), than in early old age, particularly 60s ($P = 0.002$; $P = 0.012$, respectively) and 70s ($P = 0.0001$; $P = 0.005$), but not over 80 years ($F_{17,9733} = 32.45$, $P < 0.001$). The largest variations were in physical QoL which tended to decrease across the lifespan (not shown), but social QoL was best for people over 80 years. Environmental QoL was good at every age and relatively stable.

Stepwise multiple regression assessed the relative impact of different socio-demographic variables on general overall QoL and health. As the correlation between gender and QoL was not significant, gender was excluded. Age showed the biggest impact, explaining 4.7% of QoL ($F_{1,3337} = 166.06$, $P < 0.0001$), then educational level (1.3%) ($F_{1,3338} = 44.67$, $P < 0.0001$), followed by marital status (0.2%)($F_{1,3337} = 6.09$, $P < 0.0001$). Together, socio-demographic features account for 6.2% of the variance in overall QoL.

Psychometric properties of the WHOQOL-BREF (UK)

An aim was to further investigate the psychometric properties of the UK WHOQOL-BREF; test–retest reliability, discriminant and concurrent validity, and responsiveness to change; using a substantial heterogeneous national sample. Before validity could be tested, internal consistency reliability was calculated for the overall scale. Although previously confirmed for the international measure,[15] without knowing the internal consistency reliability for the present sample, validation work would be weakened.

Internal consistency reliability
High internal consistency reliability (> 0.90) is necessary if scores are to be reliably used with individuals. For the 24 specific items of the WHOQOL-BREF (UK), internal consistency was confirmed as excellent, as alpha exceeds 0.90, at 0.92. The measure is therefore well suited to individual use.

Test–retest reliability
Repeat data enabled test–retest reliability to be assessed in two relatively stable populations. All correlations between domains from both time points were significant and positive (Table 3). They were strongest for psychological QoL (0.72), then environment (0.70), social relations (0.70) and physical health (0.66). Furthermore, domain scores from both occasions showed no significant differences, so together the findings confirm good test–retest reliability.

Discriminant validity
When well and sick people were compared, domain means were significantly different (Table 4) confirming that sick people had poorer QoL on all important dimensions. Healthy people ($n = 2761$) saw their physical and social QoL as very good, and environmental and psychological QoL as good. Sick people ($n = 1864$) reported that environmental and social QoL was good, and physical and psychological QoL acceptable. The results provide evidence of good discriminant validity for the WHOQOL-BREF (UK).

Concurrent validity
Concurrent validity was investigated by correlating WHOQOL-BREF domains (and selected facet items) with SF-36 dimensions (Table 5). As

Table 3 Test–retest reliability of the WHOQOL-BREF in well samples ($n = 347$)

	Time 1		Time 2			
QoL domain	Mean	SD	Mean	SD	t	P
Physical health	78.47	16.13	78.82	15.36	−0.500	0.617
Psychological	68.63	15.34	68.60	15.17	0.051	0.960
Social relations	72.30	20.42	72.55	20.84	−0.223	0.824
Environment	69.06	13.59	69.37	14.36	−0.530	0.597

Table 4 Discriminant validity: differences between sick and well groups on WHOQOL-BREF domains

Domain	Sick (n = 1864)		Well (n = 2761)			
	Mean	SD	Mean	SD	t	P
Physical	77.38	15.2	57.00	22.7	36.42	0.001
Psychological	68.79	15.0	58.34	19.3	20.65	0.001
Social	71.84	20.0	62.81	23.2	14.05	0.001
Environment	69.62	13.9	64.86	16.9	10.48	0.001

expected, WHOQOL general QoL was strongly, positively associated with SF-36 general health. Physical and psychological dimensions on both instruments were strongly associated, and also each of these with general health. The WHO-QOL physical domain correlated strongly with SF-36 physical dimensions: physical functioning, role physical, bodily pain and vitality (all r or $\rho > 0.7$). Validating these facets, strong correlations were confirmed between the WHOQOL pain item with SF-36 bodily pain, and energy item with SF-36 vitality. The WHOQOL psychological domain was moderately associated with vitality and role emotional (predicted) and social functioning (unpredicted). The WHO-QOL social domain associated weakly with social functioning (predicted); also weakly with role emotional, and moderately with mental health (unpredicted), limiting validation. The results indicate conceptual overlap between social and psychological areas in both measures.

Although environmental QoL was associated moderately with general health, social functioning and mental health, it maps poorly onto other SF-36 dimensions, and validation was not expected. The results indicate that the WHO-QOL-BREF environment domain makes an unusual and original contribution to generic QoL assessment in health.

Responsiveness or sensitivity to change
Preliminary assessment of the responsiveness of domain scores to changes in clinical condition compared effect sizes for 13 diseases and conditions, where repeated measures were collected (Table 6). A limited number of significant effect sizes were found; only 16 out of a possible 56. Of these, six were moderate or large. Seven groups received usual treatment, and change was expected because of their recognized efficacy. Six new or alternative interventions were evaluated, with unknown efficacy. As a very small group of patients with chronic schizophrenia moving residences represented the sole environmental intervention, this was retained to provide indicative domain results.

Significant responsiveness to change across three or more domains was found for arthritis, depression and chronic pain (Table 6). Greatest impact on physical QoL was expected for disorders that were primarily physical in manifestation, and/or where an intervention had a predominantly physical focus. Improvement in

Table 5 Concurrent validity: Pearson correlations (n = 125) between WHO-QOL-BREF domains and SF-36 sub-scales

SF-36	WHOQOL-BREF				
	QoL and general health	Physical	Psychological	Social	Environment
Physical functioning	0.61*	0.74*	0.45*	0.03	0.18**
Role physical	0.64*	0.72*	0.39*	0.24*	0.25*
Bodily pain	0.66*	0.74*	0.36*	0.13	0.21**
General health	0.76*	0.73*	0.60*	0.23**	0.37*
Vitality	0.63*	0.74*	0.67*	0.30*	0.34*
Social functioning	0.60*	0.69*	0.53*	0.33*	0.42*
Role emotional	0.48*	0.48*	0.52*	0.35*	0.33*
Mental health	0.45*	0.48*	0.70*	0.53*	0.41*
Physical score	0.66*	0.79*	0.32*	0.04	0.20**
Mental score	0.59*	0.63*	0.69*	0.50*	0.42*

*P < 0.01; **P < 0.05.

Table 6 Testing the responsiveness of the WHOQOL-BREF showing the effect size of different interventions

Participants	N	Intervention	Time	Physical	Psychological	Social	Environment
Treatment as usual							
Depression	50	Treatment as usual	3 mths	−0.65[1]	−0.78[1]	−0.13	−0.33[2]
Arthritis	18	Treatment as usual	1 year	−0.80[3]	−0.39[2]	0.23[2]	−0.12
IBS	84	Treatment as usual	6 mths	0.01	0.08	0.01	−0.04
IBD	38	Treatment as usual	Variable	−0.15	−0.08	−0.02	0.02
CFS	25	Treatment as usual	Variable	−0.09	−0.25[2]	−0.06	−0.17
Diabetes	164	Treatment as usual	Variable	0.09	0.04	0.18	0.00
Diabetes	50	Treatment as usual	Variable	0.36[2]	0.06	0.19	0.18
New interventions							
Students	42	Negative emotional disclosure	4 weeks	−0.06	−0.18	−0.11	−0.13
IBS	72	Self-help groups	8 mths	−0.20[2]	−0.17	−0.09	−0.04
Chronic schizophrenia	7	Moving residence	Variable	−0.12	−0.06	0.05	0.21[2]
Chronic pain	15	Aromatherapy and massage	4 weeks	−0.20[2]	−0.06	−0.01	−0.14
Chronic pain	12	Massage only	4 weeks	−0.58[1]	−0.19	0.37[2]	−0.21[2]
Arthroplasty	61	Hip or knee replacement surgery	3 mths	−0.51[1]	0.50[1]	0.02	0.10

IBD, irritable bowel disorder; IBS, irritable bowel syndrome; Mths, months.
Effect size criteria: [1]Moderate 0.50; [2]Small 0.20; [3]Large 0.80.

all QoL domains was predicted for depression interventions.[29] Seven of the 13 samples reported improvements to physical QoL; four offered new/alternative interventions and three usual treatments. Moderate or small improvements to physical QoL were confirmed for some musculoskeletal, diabetes, depression and irritable bowel syndrome (IBS) samples.

Four interventions showed changes to psychological QoL. Greatest impact was predicted for psychological disorders e.g. depression, and psychologically focussed interventions e.g. self-management. Substantial improvement was confirmed for depression treatment, but not chronic schizophrenia (small sample), or negative emotional disclosure in students. Patients with CFS reported expected modest improvements to psychological QoL, but none were found in IBS and IBD groups. Arthritis/arthroplasty patients reported unexpected moderate improvements to psychological QoL.

The greatest impact on social QoL was predicted for social interventions but these were not confirmed for IBS self-help groups, or depression. Small, unpredicted improvements were found for arthritis and chronic pain patients. Small improvements to environmental QoL were confirmed for patients with chronic schizophrenia who moved residences, depression treatment

and unexpectedly, chronic pain patients receiving massage.

Discussion

Fulfilling reasonable expectations of individual patients that health care improves QoL can be important to the patient–provider relationship. Being able to accurately assess whether QoL is affected by treatment is therefore essential to this endeavour. Addressing the quality of audit and other services and conditions utilizing patient-centred evidence obtained through PROMs is becoming increasingly common in UK health care. Consequently, our investigation of the performance of this innovative international generic instrument – the WHOQOL-BREF – in multiple diverse diseases, conditions and health, is timely and relevant to these policy initiatives.

The results show that the WHOQOL-BREF has very good to excellent psychometric properties, so endorsing its use as an individual assessment. It is acceptable and feasible to use as there were few refusals and negligible missing data from a large, heterogeneous sample. In view of patient-centred procedures used to develop it, and also the availability of extra national items than can be 'bolted' onto to the international core measures (WHOQOL-100

and WHOQOL-BREF) to 'round out' the cultural concept,[33] it is highly adapted to the user's culture. Together with our findings, these features indicate low respondent burden to users. These qualities are essential if it is to be routinely used in busy clinics, with very sick patients, and for monitoring and evaluation in large national population surveys.

The reliability of the WHOQOL-BREF is excellent. At 0.92, its internal consistency exceeds the criterion necessary for individual use. Test–retest reliability was also very good. Concurrent assessment with the SF-36 confirmed good validity for physical and psychological domains but less so for the social domain, probably because of the absence of a clearly defined SF-36 social subscale. There was negligible evidence that environmental QoL was assessed by the SF-36 and further validation evidence should be sought for the latter two domains. Although a health status measure, the SF-36 is still seen as a 'gold standard' measure for QoL. However, the WHOQOL-BREF provides a very good, more holistic cross-cultural QoL assessment. Environmental QoL represents a substantial addition to the QoL concept in health, and its inclusion in the WHOQOL was consensually endorsed internationally during person-centred development, then confirmed with cross-cultural survey data. Assessment of environmental QoL has many potential new applications in public health and population health.

Sick people had lower scores than well, indicating poorer QoL, and good discriminant validity for the measure. Furthermore, differences between multiple diagnostic groups and well samples provided supplementary detail to the broader picture. Patients with serious long-term chronic physical and mental health had poor QoL and the poorest, although responsiveness to change tests showed that from this low baseline, some diagnostic groups were also most likely to show the biggest QoL improvements attributed to treatment.

Evidence of responsiveness was partial, and strong effect sizes scarce. Where they occurred, they were best for physical and psychological QoL. *A priori* predictions about changes in certain diagnostic groups and for specified QoL domains were partly confirmed. Greatest responsiveness to change in the physical domain was for physical conditions, like chronic pain and arthritis. Change in the psychological domain confirmed that treatment can improve QoL in mental health, primarily depression,[29] and for physical disorders with predominant psychological components e.g. CFS. Moving into a community residence from psychiatric wards improved environmental QoL for patients with chronic schizophrenia, but because of the sample size, these results necessarily remain indicative. Longitudinal data from large homogeneous samples is needed to further test responsiveness in the WHOQOL-BREF. Of particular value would be evaluations of social and environmental interventions.

In general, usual treatment had only limited impact on QoL. Even when small samples were discounted, changes during standard treatments were generally small and sparse. Such small effect sizes could be interpreted as a limitation to the responsiveness of a generic instrument like the WHOQOL-BREF, or as ineffectiveness, either in the standard treatment itself, or in the application of the treatment in clinical practice.[34] In the case of new and alternative treatments tested, there were few substantial effect sizes, also opening up the possibility that these treatments might not be effective. However, this evidence of responsiveness will increase confidence in using the UK instrument until new data are available.[35] Work on the responsiveness of WHOQOL-BREF scores to change is incomplete. New data is needed to investigate responsiveness using other methods. There is limited consensus about which test is best, although investigation of the minimal significant clinical difference could be important, and comparing an independent measure of subjective clinical change to score changes in the WHOQOL-BREF.

Well people had good or very good QoL on all domains, and this substantial sample provides the first benchmark to clinicians and researchers who use the WHOQOL-BREF, to

direct the help they might provide to patients. This data provides a baseline for healthy QoL against which those afflicted by disease may now be compared. Although the sample contained a rich heterogeneous cross section of diseases, conditions and health, from people living in diverse locations in the British Isles, it was neither randomized nor fully representative, and this limits our ability to present norms. However, in practice, health professionals do not deal with representative samples of the population but individual patients, so the study information still has pragmatic value in benchmarking.

Quality of life for many UK patients was acceptable or better, indicating that expectations about a good QoL resulting from health care may be fulfilled for many diagnostic groups on some domains at least. The value of the WHO-QOL-BREF profile is that it enables specific but important QoL areas that are unsatisfactory to patients, to be rapidly identified. Such information could highlight priorities about which patient groups and what domains most need clinical attention. Moreover, it assists in selecting an appropriate intervention. We found poor QoL in musculoskeletal patients in chronic pain or with arthritis, heart transplant candidates, CFS and cocaine dependency, clearly indicating resource priorities. As depressed psychiatric patients had the poorest QoL of all, this reinforces the need to invest in improving well-being in mental health.[36] Our findings support current national initiatives to prioritize these areas using PROMs results.[6]

The study aimed to investigate the QoL of people from different socio-economic backgrounds in UK and consider health inequalities. Age band had a small influence on reporting QoL, although not continuous across the age range, suggesting that age-related norms may not be necessary. Other characteristics like a lack of education had a deleterious impact on QoL especially physical, indicating the patent need to differentiate how health care is delivered to different educational levels. Single and partnered people had better QoL than those who had lost a partner but gender differences were marginal. A further limitation to this study is

sample bias. More women were recruited than men, and nearly half the sample completed tertiary education. Participants were native English speakers, as British ethnic language versions of the WHOQOL-BREF were not available, so the QoL of British ethnic minority groups remains unknown.

Through showing the way that the WHO-QOL-BREF has been developed as a PROM, we conclude that it is one of the best of its kind, and deserves serious consideration when measures are selected for use in NHS practice. These results provide an empirical basis for those wishing to use the WHOQOL-BREF to examine the QoL of individual patients and well people in the community.

Acknowledgements

Univ. of Strathclyde, Glasgow; K. Webb, Univ. of Exeter; J. Noeken, M. Musa, A. Beattie, P. Andreou, I. Mills, I. Hutton & R. Day, Univ. of Bath.

References

1 Calnan KC. Quality of life in cancer patients – a hypothesis. *Journal of Medical Ethics*, 1984; **10:** 124.

2 The WHOQOL Group. Development of the WHO-QOL: rationale and current status. *International Journal of Mental Health*, 1994; **23:** 24–56.

3 Skevington SM, O'Connell KA. Can we identify the poorest quality of life? Assessing the importance of quality of life using the WHOQOL-100. *Quality of Life Research*, 2004; **13:** 23–34.

4 Staniszewska S. Patient expectations and health-related quality of life. *Health Expectations*, 1998; **2:** 93–104.

5 Browne J. Can we (should we!) use patient-reported outcome measures to compare the quality of health care providers. Presentation to the 16th Annual meeting of ISOQOL, New Orleans, USA, October, 2009.

6 Department of Health. *Guidance on the Routine Collection of Patient Reported Outcome Measures (PROMS)*. NHS 60; London: 2009.

7 Ware JE, Kosinski M, Keller SD. *SF-36 Physical and Mental Health Summary Scales: A User*'s Manual. Boston, MA: The Health Institute, New England Medical Centre, 1994.

8 EuroQol Group. EuroQol: a new facility for the measurement of health-related quality of life. *Health Policy*, 1990; **16:** 199–208.

9 Skevington SM, Mac Arthur P, Somerset M. Developing items for the WHOQOL: a study of contemporary beliefs about quality of life related to health in Britain. *British Journal of Health Psychology*, 1997; **2:** 55–72.

10 Skevington SM, Sartorius N, Amir M, The WHO-QOL Group. Developing methods for assessing quality of life in different cultural settings: the history of the WHOQOL instruments. *Social Psychiatry and Psychiatric Epidemiology*, 2004; **39:** 1–8.

11 The WHOQOL Group. The World Health Organization quality of life assessment (WHOQOL): position paper from the World Health Organisation. *Social Science and Medicine*, 1995; **41:** 1403–1409.

12 The WHOQOL Group. The World Health Organization quality of life assessment (WHOQOL): development and general psychometric properties. *Social Science and Medicine*, 1998; **46:** 1569–1585.

13 Bowden A, Fox-Rushby JA. Systematic and critical review of the process of translation and adaptation of generic health-related quality of life measures in Africa, Asia, Eastern Europe, the Middle East, and South America. *Social Science and Medicine*, 2003; **57:** 1289–1306.

14 The WHOQOL Group. Development of the World Health Organization WHOQOL-BREF quality of life assessment. *Psychological Medicine*, 1998; **28:** 551–558.

15 Skevington SM, Lotfy M, O'Connell KA. 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. *Quality of Life Research*, 2004; **13:** 299–310.

16 Chandra PS, Deepthivarma S, Jairam KR, Thomas T. Relationships of psychological morbidity and quality of life to illness-related disclosure among HIV-infected persons. *Journal of Psychosomatic Research*, 2003; **54:** 199–203.

17 O'Carroll RE, Smith K, Couston M, Cossar JA, Hayes PC. A comparison of the WHOQOL-100 and the WHOQOL-BREF in detecting change in quality of life following liver transplant. *Quality of Life Research*, 2000; **9:** 121–124.

18 Van Heck GL, de Vries J. Quality of life of patients with chronic fatigue syndrome. *Journal of Chronic Fatigue Syndrome*, 2003; **10:** 17–35.

19 Noerholm V, Groenvold M, Bjorner JB, Rasmussen N-A, Bech P. Quality of life in the Danish population – normative data and validity of the WHOQOL-BREF using Rasch and item response theory models. *Quality of Life Research*, 2004; **13:** 531–540.

20 Herrman H, Hawthorne G, Thomas R. Quality of life assessment in people living with psychosis. *Social Psychiatry and Psychiatric Epidemiology*, 2002; **37:** 510–518.

21 Stengler-Wenzke K, Kroll M, Riedel-Heller S, Matschinger H, Angermeyer MC. Quality of life in obsessive-compulsive disorder: the different impact of obsessions and compulsion. *Psychopathology*, 2007; **40:** 282–289.

22 Chachamovich E, Trentini C, Fleck MPA. Assessment of the psychometric performance of the WHO-QOL-BREF instrument in a sample of Brazilian older adults. *International Psychogeriatrics*, 2007; **19:** 635–646.

23 Amir M, Lev-Wiesel R. Time does not heal all wounds: Quality of life and psychological distress of people who survived the holocaust. *Journal of Traumatic Stress*, 2003; **16:** 295–299.

24 Lin M-R, Huang W, Huang C, Hwang H-F, Tsai L-W, Chiu Y-N. The impact of the Chi-Chi earthquake on quality of life among elderly survivors in Taiwan – a before and after study. *Quality of Life Research*, 2002; **11:** 379–388.

25 Alonso J, Ferrer M, Gandek B *et al.*; The IQOLA Project Group. Health-related quality of life associated with chronic conditions in eight countries: results from

the International Quality of Life Assessment (IQ-OLA) project. *Quality of Life Research*, 2004; **13:** 283–298.

26 Fitzpatrick R. *Patient-Reported Outcomes: Opportunities and Challenges. MRC Workshop on Patient Reported Outcome Measures*. London: Royal College of Physicians, 2009.

27 Skevington SM. Measuring Quality of life in Britain: introducing the WHOQOL-100. *Journal of Psychosomatic Research*, 1999; **47:** 449–459.

28 Skevington SM, Carse MS, Williams AC. Validation of the WHOQOL-100: pain management improves quality of life for chronic pain patients. *The Clinical Journal of Pain*, 2001; **17:** 264–275.

29 Skevington SM, Wright A. Changes in the quality of life of patients receiving antidepressant medication in primary care: validation of the WHOQOL-100. *British Journal of Psychiatry*, 2001; **178:** 261–267.

30 Skevington SM, Bradshaw J, Hepplewhite A, Dawkes K, Lovell C. How does psoriasis affect quality of life? Validating the WHOQOL-100. *British Journal of Dermatology*, 2006; **154:** 680–691.

31 Streiner DL, Norman GR. *Health Measurement Scales: A Practical Guide to their Development and Use*, 2nd edn. Oxford: Oxford University Press, 1996.

32 Howell DC. *Statistical Methods for Psychology*, 5th edn. London: Duxbury, Thomson Learning, 2002.

33 Skevington SM, Bradshaw J, Saxena S. Selecting national items for the WHOQOL: conceptual and psychometric considerations. *Social Science and Medicine*, 1999; **48:** 473–487.

34 Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *Journal of Chronic Diseases*, 1987; **40:** 171–178.

35 Hevey D, McGee HM. The effect size statistics: useful in health outcomes research? *Journal of Health Psychology*, 1998; **3:** 163–170.

36 Layard R. *Happiness: Lessons from a New Science*. London: Penguin/Allyn and Lane, 2005.