Résumé

Essai pratique du WHOQOL-100 au Canada

Anita E. Molzahn et Ginette Pagé

L'objectif de ce projet était de mettre à l'essai le WHOQOL-100, un instrument servant à évaluer la qualité de vie (QDV) conçu pour faciliter les comparaisons interculturelles. L'instrument a été testé à partir d'un échantillon de commodité de 144 personnes. L'âge moyen des participants était de 58 ans, et 41 % d'entre eux étaient des femmes. La fidélité de test-retest était de 0,86 pour le domaine physique, 0,78 pour le domaine psychologique, 0,91 pour l'indépendance, 0,87 pour les relations sociales, 0,77 pour l'environnement et 0,60 pour la spiritualité. La cohérence interne était de 0,77 pour le domaine physique, 0,79 pour le domaine psychologique, 0,89 pour l'indépendance, 0,71 pour les relations sociales, 0,80 pour l'environnement et 0,89 pour la spiritualité/les croyances personnelles. Les corrélations entre les divers domaines et les mesures correspondantes effectuées par d'autres instruments, notamment le SF-36, la Self-Anchoring Striving Scale et la Visual Analogue Scale for Pain, ont étayé la validité convergente de l'instrument. Celui-ci a permis de différencier les populations saines des populations malsaines, ce qui vient étayer sa validité conceptuelle.

Mots clés : Qualité de vie, mesure

CJNR 2006, Vol. 38 Nº 3, 106–123

Field Testing the WHOQOL-100 in Canada

Anita E. Molzahn and Ginette Pagé

The purpose of this project was to test the WHOQOL-100, an instrument for assessing quality of life (QOL), developed to facilitate cross-cultural comparison. The instrument was tested with a convenience sample of 144 people. The mean age of participants was 58 years and 41% were female. Test-retest reliability was .86 for the physical domain, .78 for the psychological domain, .91 for independence, .87 for social relationships, .77 for environment, and .60 for spirituality. Consistency reliability was .77 for physical, .79 for psychological, .89 for independence, .71 for social relationships, .80 for environment, and .89 for spirituality/personal beliefs. Correlations of the various domains with other instruments, including the SF-36, the Self-Anchoring Striving Scale, and the Visual Analogue Scale for Pain, supported the convergent validity of the instrument. The instrument was able to discriminate between healthy and ill populations, providing support for construct validity.

Keywords: Quality of life, measurement

Introduction

The concept of quality of life (QOL) has received considerable attention around the world in both the lay and the professional literature. However, measurement of the concept is complex. Although many instruments have been developed, they have numerous deficiencies, including excessive focus on disease symptoms, lack of responsiveness to change, and lack of reliability and validity (Carver, Chapman, Salazar, Stadnyk, & Rockwood, 1999; Harrison, Juniper, & Mitchell-DiCenso, 1996). These limitations have made it difficult if not impossible to compare research findings across settings and countries. Further, comparison across countries can be problematic because translation of instruments may not take cultural differences into account (Bowden & Fox-Rushby, 2003). We need a reliable and valid instrument for making accurate cross-cultural comparisons of QOL. The purpose of this project was to test an instrument for assessing QOL, the WHOQOL-100, which was developed as part of a multi-country consortium and the World Health Organization. The Canadian research team joined the international research group at a later time, after pilot testing of the WHOQOL-100 had been completed. As a result, our findings have not been previously reported and have not been compiled with the data from the other international centres.

Literature Review

Conceptualization and Measurement of QOL

QOL has been defined and measured in many different ways. The question of what constitutes QOL has been a focus of exploration and investigation since the days of the early philosophers. Aristotle and Aquinas wrote extensively about "the good life" (Adler, 1970, 1971; Molzahn & Kikuchi, 1998). The introduction of the concept into health research came at a time when the predominance of traditional medical outcomes such as mortality and morbidity was criticized because these outcomes did not represent the full range of potential outcomes of interventions. Recent years have seen the publication of numerous empirical studies (Fry, 2000; Spilker, 1996) addressing QOL in various health and illness contexts.

While some researchers have described QOL in terms of different indicators, such as social indicators, societal resources, infant mortality, economic situation, and community QOL, there is widespread agreement among researchers that the concept of QOL relates to the individual and must be based on individuals' subjective assessment of their own lives (Harrison et al., 1996). Further, it is generally agreed that QOL is a multidimensional construct (Fayers, 2000; Fry, 2000; Spilker, 1996).

Most researchers include physical, psychological, and social domains in their conceptualizations and instruments. Williams (1999) argues that the ideal assessment of QOL would include indicators of the person's physical health, social well-being, psychological functioning, environment, and spiritual well-being. Some researchers have chosen to simply use a set of independent scales that reflect one or more QOL domains. For example, in their study of QOL in assisted living settings, Mitchell and Kemp (2000) elected to use life satisfaction, depression, and satisfaction with the facility as indicators of QOL. Other researchers have used measures of depression, self-esteem, life satisfaction, functional status, and health status.

There is no single ideal measure of QOL. One of the well-known early measures, the Karnofsky scale, actually measures functional status (Karnofsky & Borchenal, 1949). The Karnofsky scale also involves ratings by health professionals, which are known to differ from ratings by the individuals themselves (McPherson & Addington-Hall, 2003; Molzahn, Northcott, & Dossetor, 1997). One of the best known and most popular measures used in QOL studies, the SF-36 (Ware & Sherbourne, 1992), is described by the authors as a measure of health and functional status. Many researchers mislabel these measures; it is important that the concept under study be clarified. QOL is broader than health and func-

tional status; it is inherently subjective and pertains to all aspects of life that are important to the person. Health status typically refers to well-being in relation to the physical, psychological, and social aspects of one's life, while functional status pertains to what one is and is not able to do for oneself.

The creation, adaptation, and testing of measures have increased exponentially over the past 20 years in Europe and North America. Spilker (1996) describes 215 measures. Bowling (1995, 1997) and McDowell and Newell (1996) provide guides for the selection and use of some of these. The journal *Quality of Life Research* is filled with papers addressing the psychometric properties of various instruments. A number of Web sites list various QOL instruments and their properties (see American Thoracic Society, 2004; MAPI Research Institute, 2004). The MAPI Research Institute Web site includes reviews of 454 measures (not all of which actually measure QOL).

There remains a need to address cross-cultural measurement of QOL. It has been difficult to compare results across studies and countries when such a wide variety of instruments are used. Further, it has been difficult for researchers to use the same measure, because many measures are not appropriate for a variety of cultural groups. In recent years several international consortia of researchers have come together to address measurement of QOL in a cross-cultural context (e.g., Bonomi et al., 1996; Ware, Gandek, & IQOLA Project Group, 1994). The WHOQOL Group is one of these. Bowden and Fox-Rushby (2003) evaluated the processes used to translate and adapt nine generic health-related QOL instruments for use in Africa, Asia, Eastern Europe, the Middle East, and South America. They argue that many of the measures focus on the scales rather than on the concept and note that the WHOQOL approach is more likely to establish equivalence of the instrument across countries.

Development of the WHOQOL-100

The rationale for the development of the WHOQOL-100 — its conceptual background and proposed uses and the steps taken to develop the pilot version — has been described in detail in a number of publications (Power, Harper, Bullinger, & WHOQOL Group, 1999; World Health Organization Quality of Life [WHOQOL] Group, 1996, 1998a, 1998b). An agreed definition of QOL provided the starting point: "Quality of life is the individuals' perception 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. It is a broad ranging concept affected in a complex way by the persons' physical health, psychological state, level of independence, social relationships and their relationship to salient features of their environment" (Szabo, 1996, p.

356). This definition drew on the World Health Organization's comprehensive approach to health and health care and was further supported by the focus group work that was used to inductively identify the components of QOL (Skevington, Sartorius, Amir, & WHOQOL Group, 2004).

The WHOQOL-100 was initially developed in 15 different countries (WHOQOL Group, 1995, 1996, 1998a, 1998b) and more than 30 centres are now involved in the project. Development of the pilot WHOQOL-100 included input at a conceptual level by culturally diverse centres. Thus, the base instrument was not provided by any one centre and then merely translated into other languages. Rather, a general instrument was produced through an iterative process that included the development of an agreed definition of QOL, agreed definitions of facets or particular characteristics of QOL and a large item pool reflecting those definitions, and, finally, an agreed set of items for the pilot version. Translation and back-translation methodology was used at all stages to ensure the applicability of different items and facets across different cultural groups. Response scales appropriate for each culture, rather than being merely translations from English, were also constructed (Sartorius & Kuyken, 1994). The net effect of this approach was the production of a subjectively defined measure of QOL.

The resulting instrument, the WHOQOL-100, consists of 100 items and six domains: physical health; psychological well-being; level of independence; social relationships; environment; and spirituality, religion, and personal beliefs. It has 24 facets or subdomains and four questions per facet. The factor structure was confirmed using confirmatory factor analysis, and the instrument was found to discriminate between healthy and ill populations (WHOQOL Group, 1998a). Some countries have developed supplementary national items for their language version of the instrument. No significant differences have been found in the performance of the scales when national items are added (Skevington, 1999). The instrument is available free of charge, in more than 40 languages, with the permission of the investigators from the respective collaborating centre (see World Health Organization, 2004).

Method

Sample

Convenience sampling was used for this study. The inclusion criteria were that participants be at least 19 years of age, able to read English, and willing to complete the questionnaire. A total of 144 people took part. All participants resided on Vancouver Island in the Canadian province of British Columbia. Most participants were recruited through a hospital outpatient clinic. They included people on dialysis (n = 32), people

awaiting joint replacement surgery (n = 52), and people with various other health problems. A "healthy" population was recruited through a university and through snowball sampling. Posters were placed on university bulletin boards inviting individuals to complete a questionnaire. Of these people, 30 (20.6%) considered themselves to be "healthy," did not report health problems, and were not receiving ongoing medical or hospital care. For the purposes of the study, "health" was self-defined, but the sample excluded people with medical problems for which they were attending an outpatient clinic. No remuneration was offered to participants.

The participants ranged in age from 19 to 89 years with a mean age of 58 years (SD = 17.93); the sample tended to be older, consisting of 92 individuals over the age of 50 (64%) and 52 over the age of 70 (36%). Sixty-eight percent were married or living with a partner. Sixty-two of the participants (43.1%) were female, 73 were male (50.7%), and six (4.2%) did not identify their gender. The sample was well educated, with 52.1% reporting postsecondary education. Fourteen participants (7.2%) resided in a hospital or other medical-care facility.

Procedure

Permission was obtained from the ethics review committee of the hospital to seek participants through its outpatient clinics. In the case of the joint replacement clinics, the research assistant attended a preoperative teaching session, described the study, and invited participation. She also visited other clinics and the dialysis unit and left letters inviting participation. She was then contacted by individuals expressing an interest in completing the questionnaire. Healthy participants were recruited via notices posted at the university. Telephone follow-up was used for outpatients who had indicated an interest in participating but had not returned the questionnaire. Participants were also asked if they were willing to complete a second questionnaire 2 to 4 weeks after the first for the purpose of assessing test-retest reliability.

All participants completed a demographic data sheet, the *World Health Organization Quality of Life–100 Inventory* (WHOQOL-100), the SF-36 (McHorney, Ware, Lu, & Sherbourne, 1994), Cantril's (1965) Self-Anchoring Striving Scale (SASS), and a Visual Analogue Scale (VAS) for pain. Completed inventories, along with the form for collecting demographic data and the signed consent forms, were returned to the investigator in self-addressed stamped envelopes.

Instruments

WHOQOL-100. The WHOQOL-100 is an inventory designed to assess the perceptions of QOL in an individual. For this study, it was

scored according to the guidelines of the developers (Harper & Power, n.d.). The WHOQOL provides 24 individual facet scores and one additional facet describing the person's overall rating of his or her QOL. It does not support aggregating facet and/or domain scores to obtain a total QOL score. The facets include pain and discomfort; energy and fatigue; and sleep and rest (in the physical health domain); positive feelings; thinking, learning, memory, and concentration; self-esteem; body image and appearance; and negative feelings (in the psychological well-being domain); mobility; activities of daily living; dependence on medicinal substances and medical aids; and work capacity (in the level of independence domain); personal relationships; social support; and sexual activity (in the social relationships domain); physical safety and security; home environment; financial resources; accessibility and quality of health and social care; opportunities for acquiring new information and skills; participation in and opportunities for recreation/leisure activities; physical environment; and transport (in the environment domain); and spirituality and personal beliefs (in the spirituality domain). The instrument was not modified for use in Canada and no new items or national items were added. Response scales were on a five-point Likert scale from lowest to highest possible scores. Higher scores denote perceptions of better QOL. Three facet scores that are negatively framed — pain, negative feelings, and dependence on medicinal substances and medical aids — were recoded to calculate domain scores that reflect higher QOL with higher scores. The 24-facet scores are combined to provide six domain scores. Each facet score is calculated by taking an average of the ratings for the four items within that facet. Within each domain, the facets that contribute to the domain score are summed and an average is taken, to yield a domain score. The facet score for spirituality is also taken as the domain score for spirituality. Scores on the WHOQOL-100 can be transformed to a 0–100 scale if necessary (Harper & Power).

Demographics

The participants completed a brief sociodemographic questionnaire: gender, date of birth, age, education level, and marital status. They also rated their health as good or poor. They were presented with a list of common health conditions (e.g., diabetes, hypertension, heart disease, kidney disease) and asked to indicate those that were a concern for them.

SF-36. The SF-36 was developed for the Medical Outcomes Study (McHorney et al., 1994) and has been used extensively over the last 10 years. The scale measures subjective health and functional status in eight domains: limitations in physical activities, limitations in usual role activities because of physical health problems, limitations in social activities, bodily pain, general mental health, limitations in usual role activities

because of emotional problems, vitality, and general health perceptions. The reliability and validity of this scale have been demonstrated in numerous studies with various populations. Across patient groups, all subscales were found to be internally consistent (most with Cronbach's alpha above .8) and had satisfactory item-discriminant validity (McHorney et al.).

Cantril's Self-Anchoring Striving Scale, a measure of life satisfaction, has been used extensively with normal and patient populations (Cantril, 1965) in at least 12 countries, including Canada and the United States. It consists of a ladder with rungs that depict the best to worst possible life. Individuals indicate where they are on the ladder. Test-retest reliability of this scale was .79 after 6 weeks. Concurrent validity was demonstrated by correlations of .50 to .74 with the Index of Well-Being (Molzahn, 1989).

Visual Analogue Scale for Pain. All participants were asked to complete a visual analogue scale for pain. They were asked to rate their pain, on a 100 mm line, from no pain to extreme pain. Visual analogue scales have been shown to measure pain reliably and validly (Jensen, 2003).

Data Analysis

SPSS version 10 was used to conduct the data analysis. Frequencies and appropriate descriptive statistics were calculated for all variables. Spearman's rho was used to calculate correlations between variables. One-way analyses of variance were used to determine whether there were differences between the healthy group and the group with one or more health problems. These were confirmed using Bonferroni adjustments at $p < .05$.

Ethical Considerations

The study was approved by a university Human Ethics Review Committee and by a hospital Ethics Review Committee. Participants were asked to complete a consent form and return it with their questionnaire. The questionnaires were coded with a number to ensure confidentiality. Consent forms and inventories were labelled with participant identification codes and stored separately, under lock and key.

Findings

Stability (Test-Retest) Reliability

To test stability reliability, participants who volunteered to complete the instruments on a second occasion 2 to 4 weeks after the first administration were given a second questionnaire and asked to mail it back; 23

participants agreed to do so. The WHOQOL-100 was found to be relatively stable. Test-retest reliability ranged from .60 to .91 on the various domains, with the low of .60 for spirituality/personal beliefs. For the facets, test-retest reliability ranged from .61 to .94, with three exceptions: opportunities for new information and skills (r = .33), transportation (r = .42), and physical safety and security (r = .54) (see Table 1).

Table 1 ***Reliability of WHOQOL-100 Domains and Facets***

DOMAIN/Facet	r (n = 23)	α (n = 140)
Physical health	.86**	.77
Pain and discomfort	.86**	.81
Energy and fatigue	.84**	.88
Sleep and rest	.89**	.94
Psychological	.78**	.79
Positive feelings	.64**	.82
Thinking, learning, memory, and concentration	.74**	.85
Self-esteem	.74**	.80
Body image	.76**	.84
Negative feelings	.81**	.90
Independence	.91**	.89
Mobility	.87**	.91
Adl	.79**	.89
Medications	.94**	.94
Work capacity	.73**	.95
Social relationships	.87**	.71
Personal relationships	.61**	.65
Social support	.77**	.84
Sexual activity	.86**	.84
Environment	.77**	.80
Physical safety and security	.54**	.83
Home environment	.67**	.84
Financial resources	.85**	.91
Health and social services	.77**	.82
Opportunities for new information	.33	.79
Leisure	.68**	.82
Physical environment	.57**	.61
Transportation	.42*	.89
Spirituality/personal beliefs	.60**	.89

Consistency Reliability

Cronbach's alpha was used to determine the internal consistency of the facets and domains. For the various facets, internal consistency ranged from a low of .61 for physical environment to a high of .95 for work capacity. For the domains, internal consistency ranged from .71 for social relationships to .89 for independence (see Table 1).

Concurrent Validity

Although there is no "gold standard" for the measurement of QOL, we examined correlations of the WHOQOL-100 with other measures, including the SASS, the SF-36, and the VAS for pain. There were moderately large and statistically significant correlations between the SASS and the various domains of the WHOQOL-100, with correlations of $r = .41$ ($p < .001$) between the SASS and the overall QOL item on the WHOQOL-100, $r = .66$ ($p < .001$) for the social relationships domain with the SASS, and $r = .62$ ($p < .001$) for the correlations between both physical health and psychological well-being domains with the SASS.

The WHOQOL-100 "overall" facet was correlated with the eight domains of the SF-36 ($r = .30$–63). In the area of physical functioning, comparisons between the WHOQOL-100 facets and domains with those of the SF-36 revealed strong correlations between related scales. The WHOQOL-100 physical functioning domain scores were highly correlated with scores on the SF-36 physical functioning scale ($r = .54$; $p < .001$), the SF-36 bodily pain scale ($r = .72$, $p < .001$), and the SF-36 vitality scale ($r = .75$, $p < .001$). The WHOQOL-100 facet for pain was correlated with the SF-36 vitality scale ($r = .82$, $p < .001$) and the SF-36 bodily pain scale ($r = .72$, $p < .001$). The WHOQOL-100 facets of energy ($r = .84$, $p < .001$), sleep ($r = .48$, $p < .001$), and participation/opportunity for recreation and leisure ($r = .53$, $p < .001$) were also correlated with the SF-36 vitality scale. As might be expected, the WHOQOL-100 facets of work capacity ($r = .63$, $p < .001$) and participation/opportunity for recreation/leisure ($r = .29$, $p < .001$) were correlated with the SF-36 role-physical scale.

In the psychosocial areas thought to contribute to QOL measured by the WHOQOL-100 and SF-36, significant correlations between responses were also found. The WHOQOL-100 facet, personal relationships, and WHOQOL-100 domain, social relationships, correlated with the SF-36 social functioning scale ($r = .32$ and $.31$, respectively; $p < .001$). The SF-36 role-emotional scale was correlated with the WHOQOL-100 facets participation/opportunity for recreation and leisure ($r = .26$, $p < .01$), work capacity ($r = .30$, $p < .001$) and WHOQOL-100 domain psychological functioning ($r = .45$, $p < .001$) with SF-36 role-emotional

scale. The following WHOQOL-100 facets were found to be significantly correlated with the SF-36 mental health scale: positive feelings ($r = .55$, $p < .001$); negative feelings ($r = .55$, $p < .001$); thinking, learning, memory, and concentration ($r = .44$, $p < .001$); self-esteem ($r = .49$, $p < .001$); and the entire WHOQOL-100 psychological domain ($r = .67$, $p < .001$).

As predicted, there were negative correlations between the VAS for pain and the various domains of the WHOQOL-100, ranging from -.25 for the relationship between social relationships on the WHOQOL-100 with the VAS for pain and -.57 for physical health on the WHOQOL-100 with VAS for pain.

Discriminatory Power

It was important for us to determine whether QOL scores reflected differences in health status. It was predicted that people who considered themselves to be healthy would have higher QOL scores than those who reported one or more health problems. We examined QOL scores for statistically significant differences between the two groups. The healthy group consisted of healthy volunteers from the university. One-way ANOVA was used to compare the mean scores of respondents who reported no health problems and those who reported one or more problems. As predicted, the instrument was able to discriminate between the two groups on overall assessment of QOL and three of the domains — physical health, psychological well-being, and level of independence — that are most likely to be affected by illness. People who reported no health problems had a higher QOL than those with one or more health problems on the overall ($F = 7.10$; $p < .01$) physical ($F = 12.75$; $p < .01$), psychological ($F = 5.95$; $p < .05$), social ($F < 1.0$; ns), environment ($F < 1.0$; ns), and spirituality domains ($F = 2.25$; ns). As one might expect, there were no statistically significant differences between the groups on social relationships, environment, and spirituality/personal beliefs.

Discussion

The WHOQOL-100 was field-tested with a sample of 144 people on Vancouver Island, British Columbia, Canada. While this is not a representative sample, the findings are similar to those of other field tests of the instrument with English-speaking people (Bonomi, Patrick, Bushnell, & Martin, 2000; Skevington, 1999; WHOQOL Group, 1998a).

Stability reliability for the domains was generally good, with correlations ranging from .6 on the spiritual domain to .91 on the independence domain. Bonomi et al. (2000) report similar stability reliability of

the domains using an intraclass correlation coefficient; they found that reliability ranged from .83 on the physical domain to .96 on the independence domain in 64 healthy people completing the instruments 2 weeks apart in the absence of life-altering events. It is not clear why test-retest reliability for the facets "transportation" and "opportunities for new information" was low in this study. It is possible that something occurred in the community to change individuals' perceptions about these areas or that the wording of the items was unclear. Also, the reason for the modest stability reliability coefficient (.6) for the domain of personal beliefs/spirituality is not clear. It is unlikely that there would be a change in personal beliefs over the period between the two administrations of the questionnaire when the scores on the other domains remained fairly constant. Otherwise, test-retest reliability was acceptable for the various domains. It is interesting to note that few researchers have actually assessed stability reliability and that those who have assessed it report scores for the domains and not the facets.

In this study, the consistency reliabilities for the various domains were in the acceptable range, from .71 for the social domain to .89 for the independence domain. However, for the facets, the reliability for physical environment was lower (.61), followed by personal relationships (.65). The lower number of questions in the facets partially explains the lower consistency reliability. Even though accepted minimum standards for alpha coefficients are .7 for group comparisons, it is not uncommon to observe an alpha of less than .7 with subscales in some instruments (Bonomi et al., 2000). Other investigators report similar Cronbach's alphas for internal consistency of the domains but do not report findings for the facets, which makes it difficult to compare these findings (Bonomi et al, 2000; Power et al., 1999; Skevington, 1999; WHOQOL Group, 1998a).

Although all correlations between the WHOQOL-100 and other measures examined were statistically significant, some of the correlations were less than the range of 0.40 to 0.80 reflecting good criterion validity. In particular, we expected a higher correlation between the SASS and the overall QOL item on the WHOQOL-100 (0.41); this may be explained, at least in part, by the observation that QOL and life satisfaction are different but related concepts. The correlations between the SASS and the various domains of the WHOQOL-100 were moderately large, and there were strong correlations between the WHOQOL-100 domain scores with similar SF-36 domains. There were lower correlations between the WHOQOL-100 facets participation/ opportunity for recreation/leisure ($r = .29$, $p < .001$) and the SF-36 role-physical scale, but this might be explained by the differences between the concepts. In the WHOQOL-100 social domain, moderately low ($r = .32$) relationships

were found with the social functioning scale of the SF-36; this might be explained by the fact that the latter instrument focuses on function rather than satisfaction. Bonomi et al. (2000), in the United States, note that while 77% of their predicted correlations for the WHOQOL-100 domains were upheld for the SF-36, convergent validity was not consistently supported.

In relation to discriminatory power, in most domains healthy people had better QOL than those with one or more health problems, but it is worth noting that healthy people did not consistently rate QOL higher than people with multiple health problems, especially in the social, environmental, and spiritual domains. Similar findings are reported from Britain in relation to spiritual and psychological domains (Skevington, 1999). This raises questions regarding how illness impacts on QOL. We typically assume that illness affects all areas of QOL, but that is not necessarily the case; the effects may differ depending on the nature of the illness and on the individual. For instance, a life-threatening illness such as cancer may have a greater negative impact on the spiritual domain of QOL than a well-controlled chronic illness such as asthma or arthritis.

Suggestions for Future Research

At this time, we are not recommending any changes to the items or the overall structure of the instrument. While reliability and validity could have been better in a few areas, confirmatory factor analyses support the factor structure of the instrument (Power et al., 1999; WHOQOL Group, 1998a). We recommend use and testing of the instrument with a wider range of samples to ensure that the noted effects were not spurious.

Further examination of the reliability of these items with a larger, more representative sample is recommended. Assessing criterion validity with another multidimensional global QOL measure, such as Ferrans and Powers's (1992) Quality of Life Index or Flanagan's Quality of Life Scale (Burkhart, Anderson, Archenholtz, & Hagg, 2003), in addition to the SF-36 and the SASS, would be advisable. Little research has actually examined the stability or reproducibility reliability of the instrument, and this should be the focus of further research.

Further evaluation of the instrument with other samples and populations is advisable before it is used with special populations. Research is also needed to clarify whether the instrument is appropriate for all age groups; it has not been tested with children or adolescents, and it may be that some of the items need modification for older adults.

While the WHOQOL-100 has been used in many studies and is often employed in cross-national cross-cultural studies, its use in the Canadian context has been limited. Given Canada's cultural diversity,

further work should be done to examine the appropriateness of various language versions for the study of cultural and minority communities.

Potential Applications of the Instrument

While there is still no perfect QOL instrument, the WHOQOL-100 is a reasonable choice for the comprehensive measurement of QOL in health-care settings, particularly with young and middle-aged adults. Because this instrument measures QOL from the perspective of the individual across important domains, it is more useful for a wide range of health-care settings than many other QOL measures that focus on specific diseases.

The instrument has numerous potential applications in mental health settings, among people with depression, schizophrenia, and other disorders. Orsel, Akdemir, and Dag (2004) examined the reliability of the WHOQOL among 54 stabilized patients diagnosed with schizophrenia and 49 healthy matched subjects in Turkey. They report that it is "a reliable subjective QOL scale for patients diagnosed with schizophrenia" but recommend further testing in large follow-up studies to assess its clinical sensitivity. The WHOQOL has also been validated with people receiving antidepressant medications (Skevington & Wright, 2001) and people with chronic pain (Skevington, Carse, & Williams, 2001).

For some clinical applications, a shorter version of the instrument, the WHOQOL-BREF, may be more appropriate (WHOQOL Group, 1998b). The BREF consists of 26 items and four facets and is useful with clinical samples where fatigue is likely to be an issue or when multiple measures are planned. It may be useful to augment it with disease-specific measures in clinical trials assessing the efficacy of specific interventions. The factor structure of the BREF is somewhat different; the spirituality domain is combined with the psychological well-being domain and the physical health domain is combined with level of independence. It has also undergone considerable testing (WHOQOL Group, 1998b).

Further work is being done on the development of supplementary modules for people with HIV/AIDS (O'Connell, Skevington, Saxena, & WHOQOL-HIV Group, 2003), for older adults (Power, Quinn, Schmdt, & WHOQOL-OLD Group, 2005; World Health Organization Regional Office for Europe, 2005), and for people with developmental disabilities (M. Power, personal communication, October 5, 2004). In addition, a group has developed a module on spirituality (WHOQOL SRPB Group, 2006).

Limitations

The study has a number of limitations. The sample was a convenience sample and not representative of any wider population. There may have

been some biases in sampling by virtue of self-selection of the participants. The size of the sample was also limited, particularly for the assessment of test-retest reliability.

Summary and Conclusions

Overall, the WHOQOL-100 performed fairly well in this sample. There was evidence of test-retest reliability, consistency reliability, and concurrent validity. As well, it appears that the instrument was able to discriminate on the physical, psychological, and independence domains between people who considered themselves to be healthy and those with health concerns. QOL is an important health outcome and nurse researchers may wish to consider use of the WHOQOL-100 measure in their studies. Further testing is recommended with diverse populations.

References

Adler, M. J. (1970). *The time of our lives: The ethics of common sense.* New York: Holt Rinehart & Winston.

Adler, M. J. (1971). *The common sense of politics.* New York: Holt Rinehart & Winston.

American Thoracic Society. (2004). *Instruments.* Retrieved August 19, 2004, from http://www.atsqol.org/qinst.asp

Bonomi, A. E., Cella, D. F., Hahn, E. A., Bjordal, K., Sperner-Unterweger, B., Gangeri, L., et al. (1996). Multilingual translation of the Functional Assessment of Cancer Therapy (FACT) quality of life measurement system. *Quality of Life Research, 5,* 309–320.

Bonomi, A. E., Patrick, D. L., Bushnell, D. M., & Martin, M. (2000). Validation of the United States' version of the World Health Organization Quality Of Life (WHOQOL) instrument. *Journal of Clinical Epidemiology, 53,* 1–12.

Bowden, A., & Fox-Rushby, J. A. (2003). A 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, South America. *Social Science and Medicine, 57,* 1289–1306.

Bowling, A. (1995). *Measuring disease: A review of disease-specific quality of life measurement scales.* Philadelphia: Open University Press.

Bowling, A. (1997). *Measuring health: A review of life measurement scales* (2nd ed.). Philadelphia: Open University Press.

Burkhardt, C. S., Anderson, K. L., Archenholtz, B., & Hagg, O. (2003). The Flanagan Quality of Life Scale: Evidence of construct validity. *Health and Quality of Life Outcomes, 1,* 59. Retrieved March 7, 2005, from http://www.hqlo.com/contenet/1/1/59.l

Cantril, H. (1965). *The pattern of human concerns.* New Brunswick, NJ: Rutgers University Press.

Carver, D., Chapman, C., Salazar, V., Stadnyk, K., & Rockwood, K. (1999). Validity and reliability of the medical outcomes Short Form-20 question-

naire as a measure of quality of life in elderly people living at home. *Age and Aging, 28,* 169–174.

Fayers, P. M. (2000). *Quality of life: Assessment, analysis and interpretation*. Toronto: John Wiley.

Ferrans, C. E., & Powers, M. J. (1992). Psychometric assessment of the Quality of Life Index. *Research in Nursing and Health, 15,* 29–38.

Fry, P. S. (2000). Whose quality of life is it anyway? Why not ask seniors to tell us about it? *International Journal of Aging and Human Development, 50,* 361–383.

Harper, A., & Power, M. (n.d.). *WHOQOL user manual*. Unpublished document.

Harrison, M. B., Juniper, E. F., & Mitchell-DiCenso, A. (1996). Quality of life as an outcome measure in nursing research: "May you have a long and healthy life." *Canadian Journal of Nursing Research, 28*(3), 49–68.

Jensen, M. P. (2003). The validity and reliability of pain measures in adults with cancer. *Journal of Pain, 4,* 2–21.

Karnofsky, D. A., & Borchenal, J. H. (1949). The clinical evaluation of chemotherapeutoic agents in cancer. In C. M. Macleod (Ed.), *Evaluation of chemotherapeutic agents* (pp. 199–205). New York: Columbia University Press.

MAPI Research Institute. (2004). QOLID: The Quality of Life Instruments Database. Retrieved August 19, 2004, from http://www.qolid/org/

McDowell, I., & Newell, C. (1996). *Measuring health: A guide to rating scales and questionnaires.* New York: Oxford University Press.

McHorney, C. A., Ware, J. E., Lu, J. F. R., & Sherbourne, C. D. (1994). The MOS 36 item Short-Form health survey (SF-36). III: Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Medical Care, 32,* 40–66.

McPherson, C. J., & Addington-Hall, J. M. (2003). Judging the quality of care at the end of life: Can proxies provide reliable information? *Social Science and Medicine, 56,* 95–109.

Mitchell, J., & Kemp, B. (2000). Quality of life in assisted living homes: A multi-dimensional analysis. *Journals of Gerontology: Psychological Sciences, 558*(2), 117–127.

Molzahn, A. E. (1989). Perceptions of physicians, nurses and patients regarding quality of life of individuals with end stage renal disease. Unpublished doctoral dissertation, University of Alberta.

Molzahn, A. E., & Kikuchi, J. F. (1998). Children and adolescents of parents undergoing dialysis therapy: Their reported quality of life. *American Nephrology Nurses Association Journal, 25,* 411–418, 428.

Molzahn, A. E., Northcott, H. C., & Dossetor, J. B. (1997). Perceptions of quality of life of physicians, nurses and patients regarding people with ESRD. *American Nephrology Nurses Association Journal, 24,* 325–335.

O'Connell, K., Skevington, S. M., Saxena, S., & WHQOL-HIV Group. (2003). Preliminary development of the World Health Organization's Quality of Life HIV instrument (WHOQOL-HIV): Analysis of the pilot version. *Social Science and Medicine, 57,* 1259–1275.

Orsel, S., Akdemir, A., & Dag, I. (2004). The sensitivity of quality of life scale WHOQOL-100 to psychopathological measures in schizophrenia. *Comprehensive Psychiatry, 45,* 57–61.

Power, M., Harper, A., Bullinger, M., & WHOQOL Group. (1999). The World Health Organization WHOQOL-100: Tests of the universality of quality of life in 15 different cultural groups worldwide. *Health Psychology, 18,* 495–505.

Power, M., Quinn, K., Schmidt, S., & WHOQOL-OLD Group. (2005). The development of the WHOQOL-OLD module. *Quality of Life Research, 14,* 2197–2214.

Sartorius, N., & Kuyken, W. (1994). Translation of health status instruments. In J. Orley & W. Kuyken (Eds.), *Quality of life assessment: International perspectives* (pp. 3–18). Berlin: Springer-Verlag.

Skevington, S. M. (1999). Measuring quality of life in Britain: Introducing the WHOQOL-100. *Journal of Psychosomatic Research, 47,* 449–459.

Skevington, S. M., Carse, M. S., & Williams, A. C. (2001). Validation of the WHOQOL-100: Pain management improves quality of life for chronic pain patients. *Clinical Journal of Pain, 17,* 264–275.

Skevington, S. M., Sartorius, N., Amir, M., & WHOQOL Group. (2004). Developing methods for assigning quality of life in different cultural settings: The history of the WHOQOL instruments. *Social Psychiatry and Psychiatric Epidemiology, 39,* 1–8.

Skevington, S. M., & Wright, A. (2001). Changes in the quality of life of patients receiving antidepressant medication in primary care: Validation of the WHOQOL-100. *British Journal of Psychiatry, 178,* 261–267.

Spilker, B. (1996). *Quality of life and pharmaeconomics in clinical trials* (2nd ed.). Philadelphia: Lippincott-Raven.

Szabo, S. (1996). The World Health Organization Quality of Life (WHOQOL) assessment instrument. In B. Spilker, *Quality of life and pharmacoeconomics in clinical trials* (2nd ed.) (pp. 355–362). Philadelphia: Lippincott-Raven.

Ware, J. E., Gandek, B., & IQOLA Project Group. (1994). The SF-36 health survey: Development and use in mental health research and the IQOLA project. *International Journal of Mental Health, 23,* 49–73.

Ware. J. E., & Sherbourne, C. D. (1992). The MOS 36-item Short-Form health survey (SF-36). I: Conceptual framework and item selection. *Medical Care, 30,* 473–483.

Williams, J. I. (1999). The unanswered challenges in measuring quality of life. *Canadian Journal of Nursing Research, 30,* 295–298.

World Health Organization. (2004). *Measuring quality of life.* Retrieved August 19, 2004, from http://www.who.int/evidence/assessment-instruments/qol/

World Health Organization Quality of Life (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.

World Health Organization Quality of Life (WHOQOL) Group. (1996). What quality of life? World Health Organization Quality of Life assessment. *World Health Forum, 17*(4), 354–356.

World Health Organization Quality of Life (WHOQOL) Group. (1998a). The World Health Organization Quality of Life assessment (WHOQOL): Development and general psychometric properties. *Social Science and Medicine, 46,* 1569–1585.

World Health Organization Quality of Life (WHOQOL) Group. (1998b). Development of the World Health Organization WHOQOL-BREF Quality of Life assessment. *Psychological Medicine, 28,* 551–558.

World Health Organization Regional Office for Europe. (2005). *Quality of life.* Retrieved February 15, 2005, from http://www.euro.who.int/ageing/Quality/20020710_1

World Health Organization WHOQOL SRPB Group. (2006). A cross cultural study of spirituality, religion and personal beliefs as components of quality of life. *Social Science and Medicine, 62,* 1486–1497.

Authors' Note

The authors gratefully acknowledge funding from the University of Victoria Internal SSHRC competition as well as the dedicated assistance of Christine Upright, who served as a research assistant for the project, and Susan Larke, who assisted with the data analysis.

Comments or queries may be directed to Anita E. Molzahn, P.O. Box 1700, Victoria, British Columbia V8W 2Y2 Canada. Telephone: 250-721-7958. Fax: 250-721-6231. E-mail: amolzahn@uvic.ca

---

Anita E. Molzahn, RN, PhD, is Professor, School of Nursing, University of Victoria, British Columbia, Canada. Ginette Pagé, RN, PhD, Professor (Retired), School of Nursing, Université du Québec à Rimouski, resides in Dunham, Quebec, Canada.