

---

## SPECIAL ARTICLE

---

# Health-related quality of life

Nares Sukcharoen MD.

*Department of Obstetrics and Gynecology, Faculty of Medicine, King Chulalongkorn Memorial Hospital, Chulalongkorn University, Bangkok 10330, Thailand*

For many years the outcomes of medical treatments have been measured primarily in terms of death, disability or cure but it is increasingly acknowledged that health-related quality of life (HRQoL) is an important additional outcome measure. Over the past 30 years several HRQoL instruments have been developed and most of them applied to studies of cancer or chronic diseases, generally to compare treatment outcomes in the context of clinical trials. A wide variety of instruments is currently available, which leaves the researcher with a confusing array of options. The main objective of this review is to present to researchers and practitioners in the field of Obstetrics and Gynecology a brief overview of the different aspects of HRQoL, how HRQoL is measured, and the different instruments developed to measure it.

### Quality of life measurement

Apart from health-related factors, quality of life also includes non-health-related elements, such as personal income, location of residence and leisure time. However, in medicine there is a need to measure HRQoL, comprising all the areas that can be affected by aspects of health. There is a consensus that HRQoL is a multidimensional construct. Generally, there are three principal aspects of HRQoL: physical, psychological and social.<sup>(1)</sup> These aspects are called domains. The physical domain refers to

perceived and observed bodily function or its disruption. Examples are pain, nausea and mobility. Psychological functioning includes both positive and negative effects, for example a patient's concerns, distress and mood. The social domain refers to the ability to perform activities related to a society role. These are activities of daily life with respect to responsibilities in and outside of the home, including those associated with family, friends and colleagues. Besides the measurement of HRQoL by means of registration of the actual conditions measured on these three main domains, a global HRQoL assessment is recommended. Such a global HRQoL refers to the subjective appraisal of one's own health as a whole. Therefore, HRQoL was defined as the "physical, psychological and social domains of health, seen as distinct areas that are influenced by a person's experiences, beliefs, expectations and perceptions".<sup>(2)</sup>

### Types of quality of life instrument

Several evaluative HRQoL instruments used in clinical trials have been developed. It can be classified into two types : generic instruments and specific instruments.(Table 1)



**Table 1.** Taxonomy of quality-of-life instruments.<sup>(3)</sup>

Generic instruments
Health profiles
Preference-based measures
Specific instruments
Disease-specific (e.g. diabetes)
Population-specific (e.g. frail elderly)
Function-specific (e.g. sexual functioning)
Condition- or problem-specific (e.g. pain)

## Generic instruments

Generic instruments are designed to be used for all kinds of diseases and medical treatments. These allow comparisons to be made across conditions and interventions. However, generic instruments may fail to capture those aspects of patients' experience that are of clinical interest in a specific clinical setting. They cannot be used to evaluate HRQoL changes in detail because they are not sensitive enough and may lack particular relevant subdomains. Evaluative HRQoL instruments encompass multiple domains and are therefore commonly indicated as health profile measures. Examples of health profile generic instruments are the Short Form 36-item Health Survey (SF-36)<sup>(4)</sup>, Sickness Impact Profile (SIP)<sup>(5)</sup>, the Nottingham Health Profile<sup>(6)</sup> and the EuroQol<sup>(7)</sup>. The SF-36 is also known as health profile instrument, because it provides individual quality-of-life scores over a range of independent dimensions. As generic instruments are able to compare patient quality of life across different illnesses over a wide range of health states, they are often used in health economic studies. Apart from health profile generic instruments, health economists have developed so-called preference-based HRQoL instruments. A distinguishing feature of a preference-based HRQoL instrument is carried out by repeated examinations measuring an unchanged characteristic.

## Specific instruments

Specific instruments are designed to assess

HRQoL in specific diagnostic groups or patient populations (e.g. individuals with ovarian cancer). The domains and items chosen for these instruments are directly related to the impairments caused by the disease process or its intervention. Therefore, specific instruments are likely to be more responsive to change. Condition- or problem- specific instruments focus solely on the conditions or problems produced by a given disease process without addressing other HRQoL issues. Examples of a symptom severity instrument is the McGill Pain Questionnaire.<sup>(8)</sup>

Choosing an HRQoL instrument is based on the goals of the proposed research. If the aim is to determine whether a particular therapy improves symptoms, a short symptom-specific instrument will be sufficient. If broader effects of a specific disease process are to be studied, a disease-specific instrument may be more appropriate. To be able to compare the results of one study with another, a generic instrument may be more appropriate. Most often, a combination of a generic and disease-specific instrument is used.

## Reliability and validity in the measurement of HRQoL

Reliability and validity of quality of life questionnaires are crucial to interpretation of the data they generate. Reliability refers to the lack of measurement error in the questionnaire, and validity indicates how well a questionnaire measures a particular construct.<sup>(9)</sup>

## Reliability

Types of reliability often reported for quality of life questionnaires include internal consistency and test-retest reliability. Internal consistency indicates whether the questionnaire items measure the same concept. (Table 2) This can be determined by item-total correlation, interitem reliability, split-half reliability, Kuder-Richardson coefficient, or Cronbach's alpha.<sup>(10)</sup> With multi-item scales, indices of internal consistency measure the extent of inter-item correlation among all items in a test. Measuring reliability by using the same test at two different times and correlating the scores is called test-retest reliability. Theoretically, reliability coefficients can range from 0 (the measure consists of error only) to 1.0 (no error exists in the measure).<sup>(11)</sup> In general, classification guidelines for reliability are as follows: if the correlation is higher than 0.7, it is considered to be adequate reliability; whereas if the correlation is 0.6 or less, it is considered marginal. Correlations of 0.8 or higher are excellent.<sup>(12)</sup>

## Validity

Validity of an HRQoL instrument is an important aspect, perhaps the most crucial. Validity is the degree to which an instrument's scores reflect what is intended to be measured. Procedures for determining test validity generally compare scores on a measure with other related, externally observable facts about the construct being measured.<sup>(5)</sup> (Table 2) Validity of a measure cannot be considered an absolute, but rather is relative to a particular situation or use. Questionnaires that have been found to be valid for a particular use in a specific population are not necessarily valid for other uses among other groups. For example, an HRQoL questionnaire that has been shown to be sensitive to changes in health status over time among stroke patients is not necessarily valid for a clinical trial of salvage chemotherapy for women with cervical cancer.

However, as evidence of the validity accumulates from multiple studies in a range of populations, one can be more confident that the extension of its use to a different population or

situation will be valid. Evidence for validity is classified in three principal categories: content, criterion, and construct validity.<sup>(9)</sup>

Content validity pertains to the issue of whether the scale items are sampled from the domain the instrument intends to measure. Some instruments are expected to cover several domains of content, for example, physical, emotional, and social functioning. To have content validity, an instrument must sample items from each of the intended content areas. Content validity is generally addressed in the development of an instrument, such as by having experts in the domain of interest rate the items as to their relevance to the domain.

Criterion validity refers to the relationship of the instrument scores to some aspect of a person's present state or future behavior. Criterion validity includes concurrent and predictive validity. Concurrent validity refers to the strength of the relationship between scores on a measure and a conceptually related variable measured at the same point in time.<sup>(11)</sup> Predictive validity refers to how well the instrument is able to predict future behavior or health state. For example, a high score on a quality of life questionnaire might predict a better tolerance to chemotherapy.

Construct validity is the extent to which scores on the measure of concern correlate with scores on other measures in predicted ways; it is usually expressed as convergent and divergent validity. Convergent validity is demonstrated when scores on an instrument are correlated with scores on a conceptually related variable, such as quality of life and ability to perform activities of daily living. Divergent validity is demonstrated by a lack of correlation with variables that are expected to be unrelated, such as quality of life and reading ability.

**Table 2.** Measurement properties of quality of life instruments<sup>(13)</sup>

	<b>Description/comment</b>
<b>Methodological properties</b>	
Reliability	Extent to which an instrument is free of measurement error (random error) Test-retest, internal consistency
Validity	Extent to which an instrument measures what it is supposed to measure (systemic measurement error)
Responsiveness	Content, criterion and construct validity Extent to which an instrument is able to detect and measure changes over time or after an intervention
<b>Practical issues</b>	
Length	Limited, because otherwise unanswered items are encountered Containing no more than 30 items
Appropriateness	The type of instrument being used should be related to the research question

Questionnaires developed in one language and culture must be cross culturally adapted and validated to ensure that use of a translated version in another country can still give meaningful results. Compliance with completing HRQoL instruments may form an obstacle to their successful implementation. If possible, the instruments should have the following characteristics: (1) be easy to understand, (2) take less than 10 min to complete, (3) contain not more than 30 items, and (4) be self administered. Moreover, attention should be given to the feasibility and completeness of data collection.<sup>(14)</sup>

When measured, HRQoL has been used to distinguish different patients or groups of patients, to predict patient outcomes, and to evaluate therapeutic interventions. Despite the proliferation of instruments and the burgeoning theoretical literature devoted to the measurement of HRQoL, no unified approach has been devised for its measurement, and little agreement has been attained on what it means. It is generally recognized as a subjective, multidimensional concept, which places emphasis on the self-perception and subjective experience but little agreement as to what items are subsumed in the concept. The variety of definitions used by researchers has made comparisons between populations and between instruments

problematic. However, lack of a uniform and explicit definition does not preclude researchers from gathering valid and useful HRQoL data. It merely serves as a reminder that the concept is implicit, subject to interpretation, and very subjective.<sup>(15)</sup>

In conclusion, HRQoL has become a more recognized treatment outcome in clinical trials. It is increasingly used to supplement objective clinical or biological measures of disease to assess the quality of service, the need for health care, the effectiveness of interventions, and in cost utility analyses. Their use reflects a growing appreciation of the importance of how patients feel and how satisfied they are with treatment in addition to the traditional focus on disease outcomes. In this respect, quality of life measures capture patients' perspectives of their disease and treatment, their perceived need for health care, and their preferences for treatment and outcomes. They are hailed as being patient centered.

Many health-related quality of life instruments have been developed, including those generic instruments and specific instruments. The generic health-related quality of life instruments are often used in conjunction with those specific instrument systems. The decision to use a particular health-related quality of life questionnaire should be based

on the validity and reliability of the instrument, especially if the questionnaire has been used and psychometric properties established in the study population. Another important factor to consider in choosing health-related quality of life instruments is to assess if the questionnaire appears to be sensitive to the differences expected between clinical trial arms. Furthermore, multinational collaborative research is needed to evaluate the reliability and validity of such questionnaires in different languages and cultures. Such research will enable patients and their physicians to have greater access to information about the quality of life implications of new treatments. Thailand has recently begun to use quality-of-life measurements to assess medical care. HRQoL instruments provide a means of collecting good quality health outcomes information. However, western scoring systems have proved not to be applicable in eastern countries, because of cultural differences between the populations.

## References

1. Donovan K, Sanson-Fisher RW, Redman S. Measuring quality of life in cancer patients. *J Clin Oncol* 1989; 7: 959-68.
2. Testa MA, Simonson DC. Assessment of quality-of-life outcomes. *New Engl J Med* 1996;28:835-40.
3. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med* 1993;118:622-9.
4. Ware JE, Sherbourne CD. The MOS36-item short-form health survey (SF-36). *Med Care* 1992;30:473-83.
5. Bergner M, Bobbitt RA, Carter WB, Gilson BS. The Sickness Impact Profile: development and final revision of a health status measure. *Med Care* 1981;19:787-805.
6. Hunt SM, McKenna SP, McEwen J, Williams J, Papp E. The Nottingham Health Profile: subjective health status and medical consultations. *Soc Sci Med* 1981;15: 221-9.
7. Brooks R, with the EuroQol group. EuroQol: the current state of play. *Health Policy* 1996;37:53-72.
8. Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. *Pain* 1975; 1: 277-99.
9. Anastasi A. *Psychological Testing*, 5th edition. New York: Macmillan, 1982.
10. Sax G. *Principles of educational and Psychological Measurement and Evaluation*, 4th edition. Belmont, CA: Wadsworth, 1997.
11. Windsor R, Baranowski T, Clark N, Cutter G. *Evaluation of Health Promotion, Health Education, and Disease Prevention Programs*, 2nd edition. Mountain View: Mayfield, 1994.
12. Minium EW. *Statistical Reasoning in Psychology and Education*, 2nd edition. New York, NY: Wiley, 1978.
13. Lagenhoff BS, Krabbe PFM, Wobbes T, Ruers TJM. Quality of life as an outcome measure in surgical oncology. *Br J Surg* 2001;88:643-52.
14. Fraser SC. Quality-of-life measurement in surgical practice. *Br J Surg* 1993; 80: 163-9.
15. Leplege A, Hunt S. The problem of quality of life in medicine. *JAMA* 1997;278:47-50.