

Dyspepsia

Health-Related Quality of Life as an Outcome in Research

H. Jorn Bovenschen, MD, Leo G. M. van Rossum, Martijn G. H. van Oijen, MSc, Robert J. F. Laheij, PhD, Paul F. M. Krabbe, PhD

Health-related quality of life (HRQOL) has become an important part of medical outcomes in research. Various aspects of HRQOL, including the different instruments that have been developed to measure it, how it is measured, and which methodologic issues are taken into account when applying it as an outcome measure, are discussed, and practical examples of how they relate to dyspepsia are provided. In one example, results showed that patients with dyspepsia had statistically significantly impaired HRQOL in all domains of the SF-36 in comparison with the general population. In another example, the presence of upper GI tract symptoms was found to further impair the already diminished HRQOL of patients with cardiovascular disease. When HRQOL is used along with conventional medical outcome measures, medical decision making can improve in both clinical practice and MCOs. (Drug Benefit Trends. 2004;16:544-556)

Key words: Outcomes measures • Quality of life • Dyspepsia

Until recently, it was common to measure the outcomes of medical treatments only in terms of mortality, morbidity, or cure. It has become clear that this view is insufficient to describe all relevant aspects of health. Quality of life is now recognized as an important additional outcome measure and is used frequently in research.¹⁻⁵ In medicine, interest in quality of life is focused mainly on areas that can be affected by aspects of health, or health-related quality of life (HRQOL).

Numerous HRQOL instruments have been developed and often are used to compare treatment outcomes in clinical trials. Instruments to measure HRQOL can be applied to any disease setting, but they have not been routinely used in clinical prac-

tice for evaluating treatment outcomes in patients with dyspepsia,³ perhaps because of the difficulty in choosing an adequate instrument from a range of options.^{6,7} Thus, in this overview, we will focus on various aspects of HRQOL as they relate to dyspepsia, including the different instruments that have been developed to measure it, how it is measured, and which methodologic issues are taken into account when applying it as an outcome measure.

When HRQOL is used in addition to conventional medical outcome measures, medical decision making can be improved not only in clinical practice but also in MCOs.

HRQOL as an Outcome Measure

Over the years, consensus has been established that HRQOL is a mul-

tidimensional concept. As such, HRQOL is generally divided into 3 domains: physical, social, and psychological.^{1,8} In the physical domain, perception and observation of normal or disrupted corporal functioning, such as mobility, pain, and nausea, are evaluated. In the social domain, the performance of societal functions is studied; these include activities of daily living and responsibilities in and out of the home, such as those associated with family, friends, and colleagues. In the psychological domain, mental and emotional functioning—for example, patients' concerns, distress, and mood—are examined.

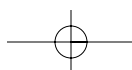
All HRQOL instruments should comprise these 3 core domains, and some may be further divided into several subdomains. Each domain is made explicit through a certain number of enquiries, or items. In addition to measuring separate domains, it is advisable to measure a global overall HRQOL.⁹⁻¹¹ This broader measurement generally refers to a patient's subjective evaluation of his or her own health, when regarding all aspects of the different domains as a whole.

Types of HRQOL Instruments

In clinical trials, 3 types of instruments (ie, questionnaires)—symptom-specific, disease-specific, and generic—typically are used to evaluate HRQOL.^{1,2,8,9,11}

Symptom-specific instruments focus exclusively on symptoms produced by a particular disease or condition. When these instruments are used, other aspects of HRQOL are

Dr Bovenschen is a physician, Mr van Rossum and Mr van Oijen are epidemiologists, and Dr Laheij is a clinical epidemiologist, all in the department of gastroenterology and hepatology at the University Medical Center, Nijmegen, the Netherlands. Dr Krabbe is a psychologist-methodologist in the department of medical technology assessment at the same institution.





Dyspepsia

excluded. There is some debate about whether such instruments satisfy the formal HRQOL definition. However, symptom-specific instruments are well suited to determine symptom improvement. In addition, these instruments are usually brief. An example of a symptom-specific instrument is the 15-item Gastrointestinal Symptom Rating Scale.^{12,13}

The domains and items selected for disease-specific instruments are directly associated with the impairments caused by a specific disease process or its therapeutic interventions. Disease-specific instruments are designed to assess HRQOL in investigative groups or patient populations. Therefore, disease-specific instruments, such as the Functional

Living Index of Cancer,¹⁴ are likely to be more reactive than generic instruments to treatment-related changes.

Generic instruments are developed for use in general populations for a broad assessment of a wide range of diseases and medical treatments. These instruments allow for comparisons of HRQOL to be made across different medical conditions and treatment interventions. However, generic instruments, such as the Medical Outcomes Study 36-Item Short Form (SF-36),¹⁵⁻¹⁷ may not pick up important specific clinical occurrences in a patient. Also, these instruments might not be sensitive enough to evaluate detailed changes and often lack relevant subdomains. Disease-specific or, possibly, symp-

tom-specific instruments are best suited to evaluate detailed HRQOL changes.

The items measured in each domain can consist of categorical scales, such as in the SF-36, or visual analog scales (VAS). In categorical scales, a number of options are offered, each with a unique numerical score. The categorical scale consists of at least 2 response categories, but usually, there will be more. Five categories per item, for example, is considered more responsive than 2 categories. In the SF-36, there are 8 domains of functional health, with the scales of each domain ranging from zero to 100; zero represents the lowest HRQOL, and 100 represents the highest possible health status in that domain (Table 1).

A VAS consists of a straight horizontal line of standard length (usually 10 or 20 cm), with the extremes of response indicated at each end. The participant is supposed to mark the line with a single tick at the point that best indicates his situation or preference between the given extremes. The score of the VAS is measured as the distance from the mark to the end of the line. A VAS apparently provides an infinite number of responses and can be somewhat complicated to evaluate. Also, VASs require more knowledge of and insight from the participant and therefore may necessitate more explanation than the easier-to-follow categorical form.

Choosing a Suitable HRQOL Instrument

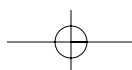
When deciding which HRQOL instrument is most suitable, one needs to keep in mind the goals of the proposed research. If the aim is to determine whether a specific treatment strategy improves symptoms, use of a short symptom-specific instrument is sufficient. For studying extensive effects on a specific disease process, a

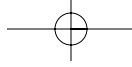
Table 1. The 8 Domains of the SF-36

Domain	Description
Physical functioning	Assesses limitations on normal physical activities; estimates the severity of limitation
Role-physical	Assesses limitations on a person's work function that are caused by physical health problems
Social functioning	Assesses the quantity and quality of interactions with others
Bodily pain	Assesses both the severity of physical pain and the extent to which it interferes with normal activities
General health	Assesses physical health status; is a good predictor of health care expenditures
Vitality	Assesses a subjective feeling of well-being, including energy and fatigue
Role-emotional	Assesses limitations on a person's work functions that are directly caused by emotional and mental problems
Mental health	Assesses the 4 major mental health dimensions of anxiety, depression, loss of behavioral or emotional control, and psychological well-being

SF-36, Medical Outcomes Study 36-Item Short Form.

Data from SF-36.org Web site.³⁷





disease-specific instrument may be more appropriate. To compare the results of 2 completely different studies, a generic HRQOL instrument may be most suitable. Often, a combination of a disease-specific instrument and a generic instrument is used, bringing together precision and more general research measures (Table 2).

Most HRQOL instruments cover multiple domains, each of them measured and presented separately and commonly indicated as profile measures. In contrast and in addition to these profile measures, health economists have developed preference-based HRQOL instruments, the strength of which is in the instruments' potential to allocate a single numerical value, a "preference value" or "utility," to the health state of a specific patient.

Using a predefined algorithm, a preference value can be estimated with descriptive health-status classifications. Such classifications consist of small sets of selected health domains (HRQOL concept) with varying levels, generally presented as questions with 3 to 7 response categories. Preference values are ex-

pressed using a scale from zero to 1.0, where zero means death and 1.0 indicates optimal health.

Combinations of the duration of health states with their corresponding preference value are called quality-adjusted life-years (QALYs). The QALY theory is attractive because it allows researchers to compare interventions with different types of medical outcomes. For economic evaluations, the costs of different treatment regimens can be related to the benefits of the strategies by performing a cost-effectiveness analysis. In this case, calculated QALYs are combined with the costs of each treatment to arrive at cost per QALY gained.¹⁸⁻²⁰

The QALY approach is frequently applied in comparisons of various medical interventions, such as treatments, medications, screening programs, and others, and is often one of the key considerations in allocation and reimbursement decision making.

Considerations for HRQOL Instruments

It is widely accepted that instruments used to assess HRQOL should be both methodologically sound and

appropriate. Surprisingly, though, many operational and even published HRQOL instruments have not been evaluated properly; these instruments should be viewed with caution. Furthermore, investigators should know in advance how data will be scored and analyzed in order to incorporate HRQOL instruments into their studies appropriately.

It is difficult to define the concept of HRQOL in explicit terms, and it is not always easy to choose the optimal HRQOL instrument in every case. Nevertheless, some methodologic, practical, and general properties of an HRQOL instrument have been defined. In addition to needing to be both practical and methodologically sound, an HRQOL instrument must be suitable to investigate the health problem in question and its probable range of effects, as well as be related to the research question.

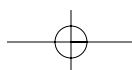
Methodologic properties. Reliability, validity, and responsiveness are 3 essential methodologic properties that an HRQOL instrument should have. Reliability is ensured when an instrument is free of measurement or random error and can be determined by studying test-

Table 2. Types of HRQOL Instruments

Instrument Type (Example)	Focus	Purpose	Construct	Property
Generic* (SF-36)	Broad perspective	Compares different conditions and treatment interventions	Often lacks subdomains	Does not detect subtle changes
Disease-specific* (FLIC)	Specific diagnostic groups and patient populations	Detects changes in disease course	More specific	More sensitive, responsive
Symptom-specific (GSRS)	Only symptoms	Registers symptom improvement	Does not fit HRQOL definition	More sensitive, responsive

HRQOL, health-related quality of life; SF-36, Medical Outcomes Study 36-Item Short Form; FLIC, Functional Living Index of Cancer; GSRS, Gastrointestinal Symptom Rating Scale.

*Most often, a combination of a generic and a disease-specific instrument is used.





Dyspepsia

retest and internal consistency capabilities of HRQOL instruments. Test-retest capabilities are evaluated by means of repeated, otherwise unchanged measurements. Internal consistency is determined by analyzing the level of association between items on the instrument within the same domain. A reliability coefficient greater than 0.80 is accepted as sufficient by most researchers.²¹

Perhaps the most crucial methodologic aspect of an HRQOL instrument is its validity—the absence of systematic measurement error—defined by the capability of the instrument to measure what it is supposed to measure. The validity concept has been divided into a number of subcategories, of which content, construct, and criterion validity are important for HRQOL instruments. Content validity is used to examine whether an instrument measures what it is supposed to measure and can be determined without scientific analysis. Conversely, criterion validity (also called predictive validity) is determined with scientific analysis but can only be applied if another instrument or method exists that is considered superior, or the gold standard. When no gold standard exists, an instrument can be analyzed by construct validity, which is used to establish whether the correct components or dimensions hypothesized to be principal to the concept of the instrument are being assessed.²²⁻²⁴

Responsiveness or sensitivity to change refers to the extent to which an instrument is able to detect and measure changes in HRQOL over time (eg, before and after an intervention). The outcome of the instrument should remain stable, unless HRQOL improves or deteriorates. Detection of changes in HRQOL is supported by 2 general features: (1) the more items a particular domain of an HRQOL instrument has, the more reactive to change it will be; (2)

the more response alternatives for each item are made available to respondents, the more responsive to changes in HRQOL the instrument will be. Five items per domain and 5 response categories per item are considered to be acceptable for detecting changes. More is possible, but the above-mentioned practical issues may make this difficult. Responsiveness is a quite recently acknowledged feature of HRQOL instruments and is well defined conceptually, but it still has some analytic shortcomings.²⁵⁻²⁸

Practical properties. To ensure that HRQOL instruments are practical, they must be straightforward and not too long; otherwise, some questions may go unanswered. Compliance with completing HRQOL instruments might influence the performance of the instrument. The instrument also should be self-administered, take no more than 10 minutes to complete, be easy to understand, and not contain too many items. Furthermore, attention should be given to the practicability and completeness of data collection.

Other considerations. Besides practical and methodologic properties, some other properties should be kept in mind when using HRQOL as an outcome measure.

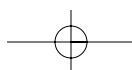
Time frames. A property that often needs careful consideration is the time frame a participant is asked to consider with each item. The problem lies in determining the balance of distinguishing between detecting possible differences and short-range fluctuations that do constitute real change. Usually, time frames are “the past week” or “the past month,” but the time frame should be considered according to the problem being assessed.

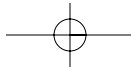
Repeated measurements. Repeated HRQOL measurements are required to reveal relevant changes in outcomes in a clinical trial. HRQOL

measurements should be scheduled at least 3 times and planned carefully. Ideally, the baseline assessment should be done before the onset of the disease under study so that full recovery of HRQOL state can be evaluated after therapy. However, it is almost never possible to assess HRQOL before a diagnosis is made. Therefore, in practice, the somewhat inappropriately called baseline assessment takes place before the start of treatment and before the randomization procedure, and can be predictive of health status during treatment and of survival. The second measurement takes place during treatment to fully assess the expected and unexpected influence of treatment, especially side effects, on HRQOL. The third evaluation is performed at the end of the treatment and constitutes treatment result on HRQOL. Certainly for chronic diseases, the third assessment may not fully address all of the HRQOL aspects because of possible relapses, ineffective therapy, or long-term results of treatment and adverse effects. Because it is almost always impossible to rule out these issues, application of follow-up instruments is recommended. Usually, a 6- or 12-month follow-up assessment is suggested.

Planning of a time frame and repeated measurements depends on the disease and has implications for the selection of study design and statistical analyses and vice versa. Because the course of a particular disease varies enormously among patients or may be unpredictable for individual patients, the measurement of HRQOL may require a special study design. Repeated measurements at fixed intervals or measurements when particular events occur, in combination with specific analytical techniques, may be performed for this purpose.

Missing data. In connection with



**Table 3. HRQOL Outcomes in Various Diseases and Comparison to Reference Values**

SF-36 Domain	HRQOL Score (SD) as Reported on the SF-36 by Controls and Patients With Various Diseases*							
	US General Population, n = 2474	Dyspepsia, n = 121	IBS, n = 140	Asthma, n = 246	Migraine, n = 303	PD, n = 73	RA, n = 693	GERD, n = 516
Physical functioning	84.2 (23.3)	69.0 (30.3)	78.2 (26.0)	71.0 (23.0)	88.9 (15.3)	70.2 (21.9)	42.0 (24.2)	79.7 (22.6)
Role-physical	81.0 (34.0)	44.1 (42.0)	52.3 (48.5)	61.0 (41.0)	56.0 (42.5)	38.8 (42.8)	28.1 (37.2)	71.6 (37.6)
Bodily pain	75.2 (23.7)	45.4 (22.0)	49.3 (26.0)	63.0 (27.0)	49.8 (23.3)	60.8 (24.5)	37.5 (17.7)	58.1 (21.0)
General health	72.0 (20.3)	54.0 (24.3)	56.5 (28.4)	60.0 (22.0)	72.1 (18.3)	62.6 (23.8)	55.2 (20.9)	67.7 (20.5)
Vitality	60.9 (21.0)	38.9 (21.9)	44.0 (26.0)	51.0 (22.0)	57.4 (19.7)	41.3 (23.9)	38.6 (22.0)	57.4 (19.9)
Social functioning	83.3 (22.7)	62.2 (28.3)	63.5 (30.8)	73.0 (21.0)	74.4 (20.8)	55.1 (26.0)	64.8 (27.1)	79.1 (23.2)
Role-emotional	81.3 (33.0)	49.7 (42.7)	60.3 (47.3)	75.0 (37.0)	81.5 (32.9)	38.4 (43.3)	57.3 (43.6)	77.8 (35.2)
Mental health	74.7 (18.1)	59.9 (21.4)	59.9 (21.4)	74.0 (18.0)	73.8 (15.6)	45.7 (19.7)	70.9 (18.9)	71.2 (18.5)
Mean score	76.6	52.9	58.0	66.0	69.2	51.6	49.3	70.3

HRQOL, health-related quality of life; SF-36, Medical Outcomes Study 36-Item Short Form; IBS, irritable bowel syndrome; PD, panic disorder; RA, rheumatoid arthritis; GERD, gastroesophageal reflux disease.

*The SF-36 domains are scored on a 0 - 100 scale, with higher scores indicating better health status.

Data from Frank L et al. *Clin Ther*. 2002.³⁴

repeated measurements and compliance, the problem of missing data is encountered frequently in longitudinal studies. Obviously, it is best to avoid the occurrence of missing data if possible, but special analytic procedures exist and are under development to at least partially correct for missing data.²⁹⁻³¹

Adjustments for multiple testing. Because most HRQOL instruments comprise multiple domains and often several HRQOL instruments are used together, adjustments for multiple testing should be made by reducing the significance level of statistical testing to overcome the possibility of detecting an apparently statistically significant effect by chance. Various corrections, for example, Bonferroni, are available to perform statistical testing in an appropriate way.³² In many cases, it will probably be advisable to consult a statistician

with special expertise in this field of correction for multiple testing.

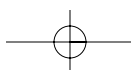
Translating instruments. Standardized translation procedures are required when an HRQOL instrument is used internationally. It can be difficult to overcome semantic differences in trials, especially when personal domains of HRQOL are involved. An international guideline has been developed that has become the standard for translating HRQOL instruments. However, the process is laborious and may discourage those considering translating an HRQOL instrument. (Translation consists of 2 forward and 2 back translations, a lay assessment, and a full report describing each step.) Cultural adaptations of existing official versions need to be reviewed by a native-speaking translation consultant, and again a lay assessment and a full report need to be conducted. Any new

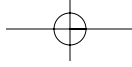
translation should be undertaken in close cooperation with the developers of the instrument.³³

Measuring HRQOL in Patients With Dyspepsia

Patients with symptoms of dyspepsia, such as gnawing or burning stomach pain, bloating, heartburn, nausea, vomiting, and belching, report worse HRQOL than persons without these symptoms.^{34,35} Overall and domain-specific differences in HRQOL between patients with dyspepsia and patients with other chronic diseases illustrate how using HRQOL as an outcome measure can help assess the impact of dyspepsia.

Two practical examples of HRQOL as related to dyspepsia are provided. In the first example, the HRQOL among patients with irritable bowel syndrome was compared with that of patients with dyspepsia





Dyspepsia

and various other chronic GI disease or non-GI disease states, as well as with reference values for HRQOL on the SF-36.³⁴

Results of the first example showed that patients with dyspepsia

had statistically significantly impaired HRQOL in all domains of the SF-36 in comparison with the general population. The greatest differences between the dyspepsia and the general population scores were for

the role-physical (44.1 vs 81, respectively) and for the role-emotional (49.7 vs 81.3) domains. Of patients with any GI disorder, those with dyspepsia had the lowest mean HRQOL score (52.9) (Table 3).

The impact of dyspepsia on HRQOL is relatively consistent for all domains when compared with the general population, whereas asthma and rheumatoid arthritis (RA) have more domain-specific impairments that explain the decline in HRQOL. Patients with asthma, for example, score nearly the same as patients with dyspepsia in the physical functioning domain but have the same mental health scores as the general population. However, patients with dyspepsia have lower mental health scores than patients with asthma. General health, vitality, and social functioning are similar between patients with RA and patients with dyspepsia. However, HRQOL scores related to physical functioning, role-physical, and bodily pain for patients with RA were much lower than scores for patients with dyspepsia; in the role-emotional and mental health domains, patients with dyspepsia were more impaired (Figure).

By comparing the general mean differences in HRQOL between disease states and by examining the scores for the 8 domains of the SF-36, the elements of HRQOL that are responsible for HRQOL differences can be determined.³³

The second example looks at the impact of GI symptoms on HRQOL in patients with cardiovascular disease and illustrates how the effect of different variables on HRQOL can be examined using a general scale and domain-specific information.³⁵ Occasionally, the HRQOL concept is used to measure the impact of domains on existing disease states. The influence of GI symptoms on the health status of patients with cardiovascular dis-

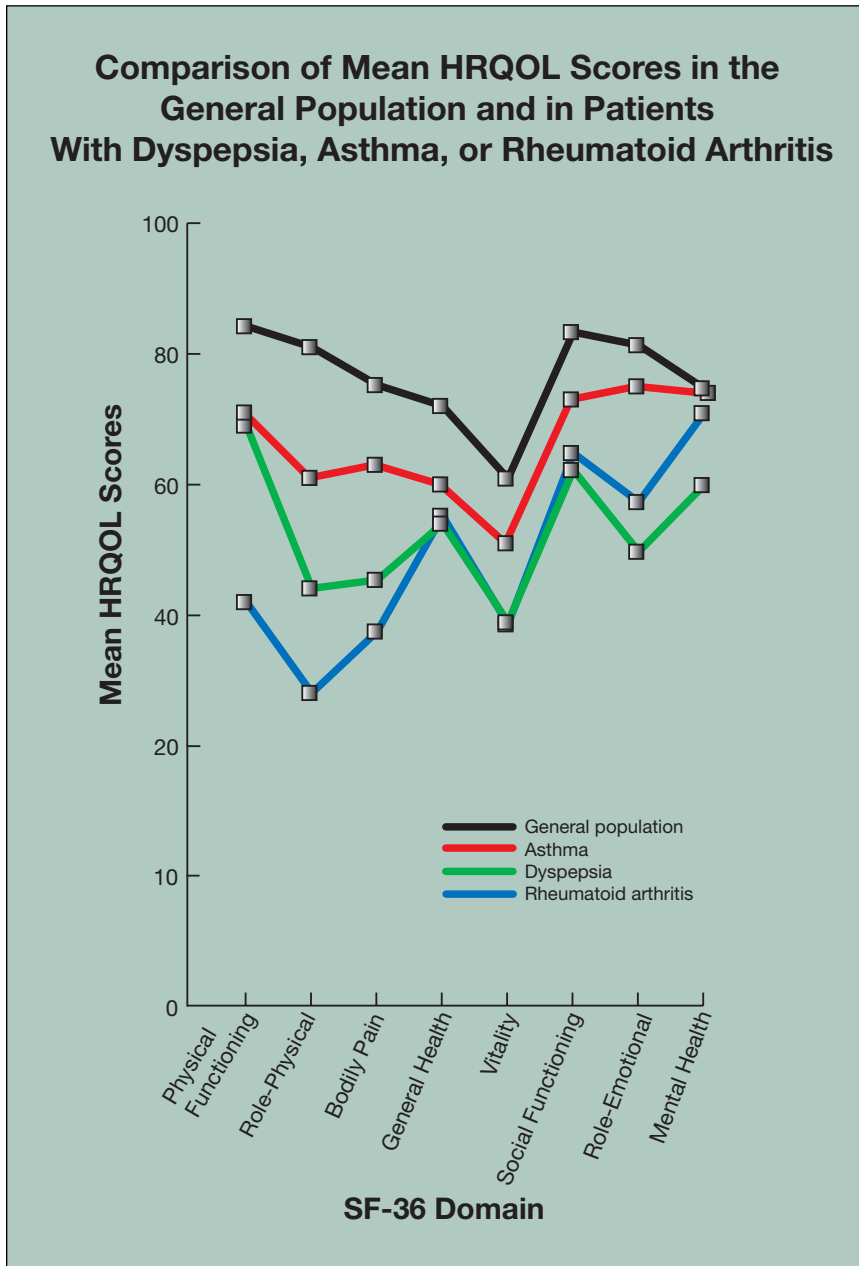
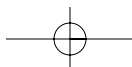
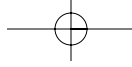


Figure. Patients with dyspepsia reported lower health-related quality-of-life (HRQOL) scores than patients with asthma or the general population. In addition, patients with dyspepsia also reported lower general health, social functioning, role-emotional, and mental health scores than any other comparison group. (SF-36, Medical Outcomes Study 36-Item Short Form.) (Data from Frank L et al. Clin Ther. 2002.³⁴)





ease, determined by the EuroQol-5D instrument, was studied.³⁵ The EuroQol-5D consists of a VAS, which ranges from zero to 100 and reflects overall health status. Five domains make up the EuroQol-5D—mobility, self-care, usual activities, pain/discomfort, and anxiety/depression.

The overall VAS score (95% confidence interval) was 70 (range, 68 to 72) for patients with cardiovascular disease but without upper GI tract symptoms. The VAS score was 61 (range, 59 to 63) for cardiology patients with upper GI tract symptoms. The differences in scores were statistically significant. Analyses showed that in all 5 domains, patients with upper GI tract symptoms experienced significantly more problems than patients without GI symptoms: mobility, 50% vs 37%, respectively; self-care, 19% vs 10%; usual activities, 70% vs 52%; pain/discomfort, 96% vs 38%; and anxiety/depression, 41% vs 20%. These results showed that the existence of upper GI tract symptoms strongly impairs the already diminished HRQOL of patients with cardiovascular disease in general, as well as in all separate domains.³⁵

Final Considerations

HRQOL is increasingly being recognized as an important outcome measure in addition to the conventional end points.¹⁻⁵ However, aspects of HRQOL measurement discussed here should be kept in mind when applying the HRQOL concept. It is possible to establish criteria against which HRQOL instruments can be evaluated.

HRQOL is a subjective outcome measure, so its measurement should include the perspective of participants and be sensitive to change over time. Those instruments should be reliable, valid, and responsive. Practical issues, such as timing of assessment, responsibilities for qual-

ity control, and data management, should be considered in order to assemble appropriate data. Statistical analysis should be performed carefully, without focusing too much on statistical significance.

Different instruments can be chosen to measure HRQOL, but the goal and other specific clinical aspects of the study should be kept in mind when choosing the instrument. HRQOL instruments have limitations, but by carefully choosing complementary instruments, an acceptable determination of the HRQOL state is possible. Both scientists and physicians are often reluctant to accept subjective evaluations such as HRQOL. However, HRQOL should be considered as an addition to objective measurement. As such, HRQOL can contribute to clinical outcomes as evaluations of survival, complications, and adverse effects related to use of therapies. HRQOL data may help guide physicians in selecting the most suitable therapy for a particular disease. Treatment evaluation can be measured in a meaningful way for both patients and physicians.

Although the 2 dyspepsia examples showed that HRQOL could be measured to determine how a disease and its related symptoms impacts patients' daily lives, it is not the only application for HRQOL measurement. It also can be useful in decision making in MCOs and economic evaluations, such as cost-benefit or cost-effectiveness analyses. When HRQOL is considered alongside conventional outcomes (mortality, morbidity, cure) by using preference values and QALY calculations, greater insight into how a disease impacts a patient and society is achieved.

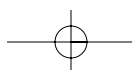
Future developments in HRQOL include methods to reduce the number of items needed, in order to improve these instruments' reliability and respondents' compliance. New

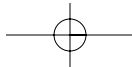
approaches that use multimedia devices in an interactive approach are being developed. This interactive concept, which was first used in intelligence testing, is called adaptive testing. The numbers of items on the instrument are limited to those that are useful for the evaluation of the specific situation for a specific person. Use of media such as the Internet can probably play an essential role in these methods of HRQOL evaluation.³⁶

HRQOL assessments are meaningful in clinical research as well as in MCOs. Adequate knowledge of the concept and how it can be used now and in the future is necessary for appropriate application in clinical practice, research settings, and management. ■

References

- Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med.* 1993;118:622-629.
- Anderson CS, Carter KN, Brownlee WJ, et al. Very long-term outcome after stroke in Auckland, New Zealand. *Stroke.* 2004;35L:1920-1924.
- Wiklund I. Aspects of quality of life in gastrointestinal disease: some methodological issues. *Scand J Gastroenterol.* 1995;30(suppl 208):129-132.
- Dean BB, Gerner D, Gerner RH. A systematic review evaluating health-related quality of life, work impairment, and healthcare costs and utilization in bipolar disorder. *Curr Med Res Opin.* 2004;20:139-154.
- Vigano A, Donaldson N, Higginson IJ, et al. Quality of life and survival prediction in terminal cancer patients. *Cancer.* 2004;101:1090-1098.
- Essink-Bot ML, Krabbe PF, Bonsel GJ, Aaronson NK. An empirical comparison of four generic health status measures. The Nottingham Health Profile, the Medical Outcomes Study 36-Item Short-Form Health Survey, the COOP/WONCA charts, and the EuroQol instrument. *Med Care.* 1997;35:522-537.
- Dougherty CM, Dewhurst T, Nichol WP, Spertus J. Comparison of three quality of life instruments in stable angina pectoris: Seattle Angina Questionnaire, Short Form Health Survey (SF-36), and Quality of Life Index—Cardiac Version III. *J Clin Epidemiol.* 1998;51:569-575.
- Testa MA, Simonson DC. Assessment of quality-of-life outcomes. *N Engl J Med.* 1996;334:835-840.
- MacKeigan LD, Pathak DS. Overview of health-related quality of life measures. *Am J Hosp Pharm.* 1992;49:2236-2245.
- Mishoe SC, Maclean JR. Assessment of health-related quality of life. *Respir Care.* 2001;46:1236-1257.
- Dedhiya S, Kong SX. Quality of life: an overview of the concepts and measures. *Pharm World Sci.* 1995;17:141-148.
- Svedlund J, Sjodin I, Dotevall G. GRS—a clinical rating scale for gastrointestinal symptoms in patients with irritable bowel syndrome and pep-





Dyspepsia

- tic ulcer disease. *Dig Dis Sci*. 1988;33:129-134.
13. Revicki DA, Wood M, Wiklund I, Crawley J. Reliability and validity of the Gastrointestinal Symptom Rating Scale in patients with gastroesophageal reflux disease. *Qual Life Res*. 1998;7:75-83.
 14. Elliott BA, Renier CM, Haller IV, et al. Health-related quality of life (HRQoL) in patients with cancer and other concurrent illnesses. *Qual Life Res*. 2004;13:457-462.
 15. Walters SJ. Sample size and power estimation for studies with health related quality of life outcomes: a comparison of four methods using the SF-36. *Health Qual Life Outcomes*. 2004;25:26.
 16. Ware JE Jr, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36), I: conceptual framework and item selection. *Med Care*. 1992;30:473-483.
 17. Brazier JE, Harper R, Jones NM, et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ*. 1992;305:160-164.
 18. Torrance GW, Feeny DH. Utilities and quality-adjusted life years. *Int J Technol Assess Health Care*. 1989;5:559-575.
 19. Williams A. Economics of coronary artery bypass grafting. *Br Med J*. 1985;291:326-329.
 20. Krabbe PF, Bonsel GJ. Sequence effects, health profiles and the QALY model: in search of realistic modeling. *Med Decis Making*. 1998;18:178-186.
 21. Dawson-Saunders B, Trapp RG. Ways to measure reliability. In: *Basic and Clinical Biostatistics*. 2nd ed. Norwalk, Conn: Appleton and Lange; 1994:57-58.
 22. Krabbe PFM, Essink-Bot ML, Bonsel GJ. The comparability and reliability of five health-state valuation methods. *Soc Sci Med*. 1997;45:1641-1652.
 23. Brazier J, Jones N, Kind P. Testing the validity of the Euroqol and comparing it with the SF-36 health survey questionnaire. *Qual Life Res*. 1993;2:169-180.
 24. Hays RD, Anderson RT, Revicki D. Assessing reliability and validity of measurements in clinical trials. In: Staquet MJ, Hays RD, Fayers PM, eds. *Quality of Life Assessment in Clinical Trials, Methods and Practices*. Oxford, England: Oxford University Press; 1998:167-182.
 25. Österberg A, Graf W, Karlbom U, Pahlman L. Evaluation of a questionnaire in the assessment of patients with faecal incontinence and constipation. *Scand J Gastroenterol*. 1996;31:575-580.
 26. Terwee CB, Dekker FW, Wiersinga WM, et al. On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. *Qual Life Res*. 2003;12:349-362.
 27. Wright JG, Young NL. A comparison of different indices of responsiveness. *J Clin Epidemiol*. 1997;50:239-246.
 28. Streiner DL, Norman GR. *Health Measurement Scales: A Practical Guide to Their Development and Use*. Oxford, England: Oxford University Press; 2003.
 29. Raboud JM, Singer J, Thorne A, et al. Estimating the effect of treatment on quality of life in the presence of missing data due to drop-out and death. *Qual Life Res*. 1998;7:487-494.
 30. Zwiderman AH. Statistical analysis of longitudinal quality of life data with missing measurements. *Qual Life Res*. 1992;1:219-224.
 31. Brown CH. Protecting against nonrandomly missing data in longitudinal studies. *Biometrics*. 1990;46:143-155.
 32. Rose MS, Koshman ML, Spreng S, Sheldon R. Statistical issues encountered in the comparison of health-related quality of life in diseased patients to published general population norms: problems and solutions. *J Clin Epidemiol*. 1999;52:405-412.
 33. da Mota Falcao D, Ciconelli RM, Ferraz MB. Translation and cultural adaptation of quality of life questionnaires: an evaluation of methodology. *J Rheumatol*. 2003;30:379-385.
 34. Frank L, Kleinman L, Rentz A, et al. Health-related quality of life associated with irritable bowel syndrome: comparison with other chronic diseases. *Clin Ther*. 2002;24:675-689.
 35. Laheij RJ, van Rossum LG, Krabbe PF, et al. The impact of gastrointestinal symptoms on health status in patients with cardiovascular disease. *Aliment Pharmacol Ther*. 2003;17:881-885.
 36. Revicki CA, Cella DF. Health status assessment for the twenty-first century: item response theory, item banking and computer adaptive testing. *Qual Life Res*. 1997;6:595-600.
 37. SF-36.org Web site. Available at: www.SF-36.org. Accessed October 7, 2004.

