

Exploration of the Value of Health-Related Quality-of-Life Information From Clinical Research and Into Clinical Practice

GORDON H. GUYATT, MD; CAROL ESTWING FERRANS, RN, PhD; MICHELE Y. HALYARD, MD; DENNIS A. REVICKI, PhD; TARA L. SYMONDS, PhD; CLAUDETTE G. VARRICCHIO, DSN, RN; ANNA KOTZEVA, MD; JOSE M. VALDERAS, PhD; JORDI L. ALONSO, MD; AND THE CLINICAL SIGNIFICANCE CONSENSUS MEETING GROUP

Quality-of-life (QOL) instruments used in clinical research can provide important evidence to inform decisions about alternative treatments. This is particularly true when patients, such as those with cancer who are contemplating toxic chemotherapy, face tradeoffs between quantity of life and QOL or when the primary goal of therapy is to improve how patients feel. Surrogate measures (cardiac function, exercise capacity, bone density, tumor size) are inadequate substitutes for direct measurement of QOL. Quality-of-life measures will be most valuable when they comprehensively measure aspects of QOL that are both important to patients and likely to be influenced by therapy, when the QOL measurement instruments are valid (measuring what is intended) and responsive (able to detect all important changes, even if small), and when the results are readily interpretable (determining whether treatment-related changes are trivial, small but important, or large). Researchers are finding new, imaginative ways to help clinicians understand the magnitude of treatment impact on QOL. Additionally, QOL measures may be useful in clinical practice. Recent results from well-designed randomized controlled trials suggest that information on patient QOL provided to clinicians might, in some circumstances, result in benefits for these patients. Further investigation is warranted to confirm these observations and to define the particular combination of methods and settings most likely to yield important benefits.

Mayo Clin Proc. 2007;82(10):1229-1239

CI = confidence interval; CRQ = Chronic Respiratory Questionnaire; HRQL = health-related quality of life; NNT = number needed to treat; QOL = quality of life; RCT = randomized controlled trial

CLINICAL SCENARIO

A 75-year-old man with severe chronic obstructive pulmonary disease is considering participation in a respiratory rehabilitation program. After many years of heavy smoking, the patient finally stopped but has severely impaired pulmonary function and dyspnea on minimal activity. He feels constantly tired and is often frustrated and anxious. Participation in the program would disrupt his routine and involve more than an hour of driving on most days for 4 weeks. He is wondering whether the program is worth the inconvenience and effort. "Doctor," he asks the chest physician responsible for his care, "how much benefit will I get from this rehabilitation program?"

The goal of this article is to explore the potential usefulness of health-related quality-of-life (HRQL) measures for clinicians. Clinicians and their patients may benefit from quality-of-life (QOL) measurement in 2 different ways. First, investigators who use QOL instruments in clinical

research can provide important evidence to inform clinician and patient decision making about alternative treatments. Because of the potential usefulness of including QOL measures among the outcomes of many randomized controlled trials (RCTs), we present examples that illustrate how information about QOL can facilitate optimal medical management of patients.

The second context in which QOL measures may provide added value arises from their use in day-to-day clinical practice. Asking patients to complete QOL questionnaires may prove an efficient way of gathering needed data about their functioning and well-being, alert clinicians to problems that require intervention, and ultimately improve patient outcomes. Only empirical data, ideally from RCTs, can determine whether QOL measures, when used in clinical practice, actually achieve any of these desirable outcomes. Some measures in some contexts will achieve all these goals, whereas other measures in other contexts may prove a waste of time and energy. Therefore, this article reviews the RCTs that have tested the impact of the use of QOL measures in clinical practice and presents the results. Before beginning an exploration of the potential usefulness of QOL measures, we specify exactly what we mean by QOL measurement information.

WHAT IS HRQL?

Quality of life is a broad concept that relates to all aspects of human life. When using the concept in health care, HRQL denotes a focus on the effects of illness and treat-

From the Department of Clinical Epidemiology and Biostatistics, McMaster University and Health Sciences Center, Hamilton, Ontario (G.H.G.); College of Nursing, University of Illinois at Chicago, Chicago (C.E.F.); Department of Radiation Oncology, Mayo Clinic, Scottsdale, AZ (M.Y.H.); Center for Health Outcomes Research, MEDTAP International, Inc, Bethesda, MD (D.A.R.); Outcomes Research Department, Pfizer, Sandwich, Kent, UK (T.L.S.); Varricchio Consulting, Rockville, MD (C.G.V.); Health Services Research Unit, Institut Municipal d'Investigació Mèdica (IMIM-IMAS), Barcelona, Spain (A.K., J.M.V., J.L.A.). A list of the members of the Clinical Significance Consensus Meeting Group appears in Frost et al.¹

Individual reprints of this article are not available. Address correspondence to Gordon H. Guyatt, MD, Department of Clinical Epidemiology and Biostatistics, McMaster University and Health Sciences Center, Hamilton, ON L8N 3Z5, Canada (guyatt@mcmaster.ca).

© 2007 Mayo Foundation for Medical Education and Research

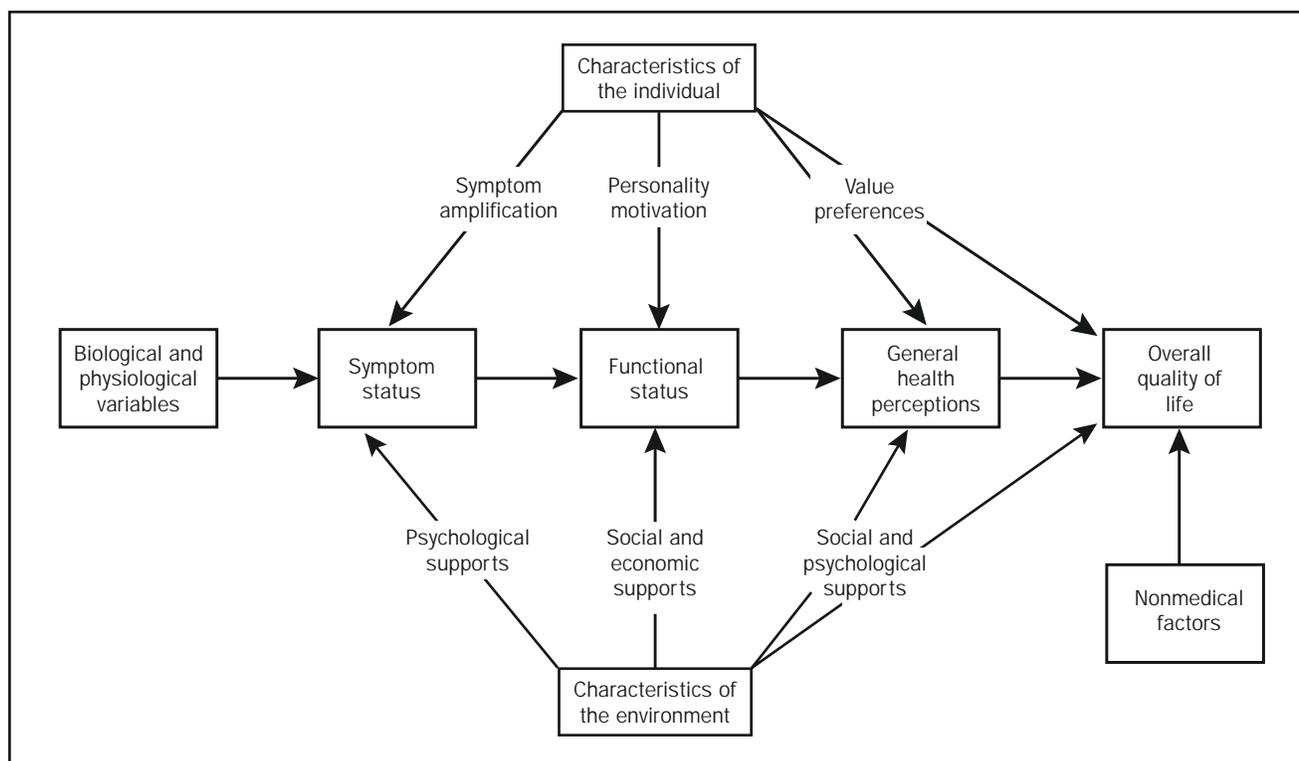


FIGURE 1. Relationship among measures of patient outcomes in a health-related quality-of-life conceptual model. From *JAMA*,² with permission. Copyright 1995, American Medical Association, all rights reserved.

ment on QOL. Under the rubric of QOL, authors have used many different terms as if they were all the same thing, which has resulted in some confusion.

There are a number of conceptual models to characterize QOL.² Wilson and Cleary³ developed one particularly helpful model (Figure 1) that focuses on 5 types of patient outcomes. The model begins with the most fundamental determinants of health status: biological and physiological variables. The next box is symptom status, which includes emotional, cognitive, and physical symptoms. The third is functional status, which addresses physical, social, role, and psychological functioning. The fourth box, general health perceptions, is a subjective evaluation that integrates all the preceding components. The fifth level is overall QOL, described in this model as subjective well-being, which means how happy or satisfied a person is with life as a whole.

The model highlights 3 important points. First, although the model begins with biological and physiological variables, the subsequent components are not limited to physical aspects of health. Experts generally agree that QOL depends on more than physical health status alone: illness can have a pervasive effect that seeps into all areas of life.⁴ The second point is that many of the components of this model require patient input. One cannot measure QOL

without taking into account the patient's direct experience⁵ because it is important in the decision-making process as it relates to disease management. Third, it is worthwhile to distinguish among the various components of the model. In the literature, these components are often lumped together under the general heading of QOL, even though important differences exist among them. Different components of QOL produce different information by which to judge the efficacy of treatment.²

This report focuses on a broad conceptualization of QOL, which may include symptoms, functional status, health perceptions, and overall QOL. We distinguish among these components when separation of the component is important.

CONCEPTUALIZATION AND ADDED VALUE OF QOL IN CLINICAL RESEARCH

Clinicians offer treatment to their patients for 3 reasons. Clinicians encourage their patients to adhere to a treatment when they believe it (1) increases longevity, (2) prevents future morbidity, or (3) makes patients feel better. Feeling better includes avoiding discomfort (eg, pain, nausea, breathlessness), disability (ie, loss of function), and distress (ie, emotional problems).⁶ For many years clinicians were

willing to substitute physiological or laboratory tests for the direct measurement of the third end point, in part because of difficulty in measurement. However, during the past 20 years, clinicians have recognized the importance of direct measurement of how people are feeling and the extent to which they are able to function in daily activities, that is, QOL.

Our discussion is designed for clinicians asking the question, "To what extent will alternative management strategies result in my patient feeling better or worse?" Other questions clinicians may ask before considering treatment include "How is my patient feeling now?" or "What issues or concerns are on my patient's mind?" Quality-of-life studies can help clinicians address these questions through an understanding of what average patients may be experiencing or the average patient's worries and preoccupations. For instance, what are the concerns of women who are receiving chemotherapy or hormonal therapy for breast cancer⁷⁻⁹ or of men with prostate cancer?¹⁰ Attention to such studies may increase clinicians' understanding of patients' perspective and thus improve their communication with their patients.

Still, the pragmatic clinician is primarily interested in ensuring that patients make the best decisions regarding their management. Keeping this goal in mind, when should the clinician be concerned about whether and how a trial included QOL measurement?

DO WE NEED TO BE CONCERNED ABOUT QOL OUTCOMES?

Early clinical trials did not include measurement of QOL. When should a clinician be concerned that investigators have not paid adequate attention to how patients feel?

Most patients agree that, in most circumstances, prolonging their lives is a sufficient reason to accept a course of treatment. Previously, investigators demonstrated that 24-hour oxygen administration in patients with severe chronic airflow limitation reduced mortality.¹¹ Since the intervention prolongs life, a subsequent report suggesting that more intensive oxygen therapy had little or no impact on QOL does not diminish enthusiasm for continuous oxygen administration.¹² Although feeling better is important to patients with heart failure, if angiotensin-converting enzyme inhibitors extend life to an appreciable degree, clinicians may be ready to recommend the treatment to patients even in the absence of known benefit to QOL.¹³ Similarly, since milrinone shortens¹⁴ life span, QOL assessments become irrelevant if clinicians are unwilling to prescribe a treatment that leads to premature death even if it improves QOL in the interval.

Many life-prolonging interventions improve QOL. When treatment is associated with adverse effects, however, it may lead to important decrements in QOL. For

example, the patient considering whether to undergo toxic cancer chemotherapy that will provide marginal gains in longevity faces a QOL benefit-cost tradeoff. At the extreme, an intervention such as mechanical ventilation may prolong the life of a patient in a vegetative state, but the patient's family may wonder if keeping their loved one alive is the best option.

When the primary goal of treatment is to improve how people are feeling and physiological measures that correlate with patients' experiences are unavailable, measurement of some aspect of QOL becomes imperative. For example, clinicians would pay little attention to studies of antidepressant medications that failed to measure patients' mood or to RCTs of antimigraine medication that failed to measure pain.

Decisions are difficult when the relationship between physiological or laboratory measures and QOL outcomes is uncertain. Physicians have substituted physiological measures for QOL measurements (substitute end points) not because they have been uninterested in making patients feel better but because they assumed a strong link between physiological measurements and patients' functioning and well-being. Unfortunately, substitute end points such as bone density for fractures, homocysteine for coronary artery disease, and laboratory exercise capacity for ability to undertake day-to-day activities have often proved misleading.

Changes in conventional measures of clinical status show only weak to moderate correlations with changes in QOL^{15,16} and often fail to detect changes in QOL that are important to patients.¹⁷ In contrast, focus on physiological measures may lead clinicians to believe that a treatment is beneficial when in fact it does not change the way patients feel. For example, trials in patients with chronic lung disease have shown treatment effects on peak flow rates without improvement in QOL.^{18,19}

If the primary goal of therapy is to improve the way patients feel, direct measurement of QOL and use of the results in clinical decision making become imperative. Clinicians can use the following litmus test: If the end points measured by the investigators were the only thing that changed, would patients be willing to take the treatment? In addition to change in clinical or physiological variables, patients require that they feel better or live longer. Consider the patient in the aforementioned "Clinical Scenario" section of this article. The focus of the patient's question is QOL, and only trials that measure the outcomes in which he is interested—dyspnea in daily life, fatigue, and emotional function—will be relevant to addressing his concerns. The following vignette provides a further example of this issue.

A 70-year-old postmenopausal woman living on a limited income in Ontario, Canada, has, during the past 5 years, lost 5.08 cm (2 in) in height and experienced

TABLE 1. Issues to Consider in Evaluative Studies of QOL*

General issues	Specific points to consider
How can clinicians be sure that investigators have measured aspects of life that patients value?	Did the investigators present a detailed account of the content of their instruments? Did the investigators ask the patients directly? Did the investigators cite prior research about their instruments?
Did the HRQL instruments work in the intended way?	Ensure the HRQL measures are Valid—measure what they are intended to measure Reliable—consistently measure HRQL in the absence of change Responsive—instrument’s ability to detect change
Are there important aspects of HRQL that have been omitted?	Check if Investigators have used generic instruments or disease-specific or condition-specific measures Generic instruments are, on average, less powerful in detecting treatment effects than specific instruments but broadly assess different aspects of HRQL Specific instruments are, on average, more powerful in detecting treatment effects but sometimes focus too narrowly to detect unanticipated treatment effects

*HRQL = health-related quality of life; QOL = quality of life.

moderate, intermittent back pain. A radiograph of her vertebral spine shows 2 vertebral compression fractures. The physician informs the patient that she has osteoporosis, is likely to have repeated vertebral fractures, and is at higher risk than other women her age for long-bone fractures of the arm or hip. The physician suggests to the patient that she take cyclical etidronate, a drug that the Ontario drug benefit plan will largely pay for in persons older than 65 years.

“What will the drug do for me?” asks the patient.

“Well,” replies the physician, “a rigorous systematic review of randomized trials shows that etidronate will improve measurements of your bone density.”²⁰

“I’m not impressed,” replies the patient, “unless those stronger bones improve my quality of life.”

“Well,” the clinician points out, “the drug decreases your risk of having any more of those spine fractures by about a third. Now, I’m not sure how much of your back pain is actually due to the fractures, but there is a good chance that the drug can prevent some future back pain.”

“Hmmm, that sounds promising. But will it do anything to prevent those more serious fractures you’ve warned me about, particularly hip fractures, which I’ve heard would have really serious consequences for my quality of life.”

“Unfortunately, etidronate probably doesn’t reduce long-bone fractures. I could offer you 2 other drugs that rigorous systematic reviews of randomized trials have shown do reduce those fractures, alendronate and risidronate, but they aren’t covered by the drug plan.”²¹

What follows is a detailed discussion of exactly how large a risk of vertebral and long-bone fractures the patient is facing, the consequences of those fractures for the patient’s QOL, exactly how much the drugs will reduce those risks, and how much the drugs would cost the patient. Bone density plays no further part in the discussion, and the functional and QOL consequences of fractures play a major part.

DID THE INVESTIGATORS USE ADEQUATE QOL MEASURES?

The added value of QOL outcomes depends on their measurement properties (Table 1). Measures with limitations in reliability and validity may produce misleading results. Clinicians who use QOL results to guide their clinical care may benefit from a basic awareness of the key issues in determining QOL data credibility. The key points from a more extended clinician-oriented discussion are briefly summarized subsequently.²²

How Can Clinicians Be Sure That Investigators Have Measured Aspects of Life That Patients Value?

Investigators may show that the outcomes they have measured are important to patients by asking them directly. For example, in a study that examined QOL in patients with chronic airflow limitation, the investigators used a literature review and interviews with clinicians and patients to identify 123 items that reflected possible ways that patients’ illness might affect their QOL.²³ Then they asked 100 patients to identify the items that were problems for them and to indicate their importance. They found that the most important problem areas for patients were dyspnea during day-to-day activities and chronic fatigue. An additional area of difficulty was emotional function, including having feelings of frustration and impatience. Reference to other literature and inquiries of health care workers may also be important, particularly in pediatric or cognitively impaired populations.

If the authors do not present direct evidence that their outcome measures are relevant and important to patients, they may cite prior research or provide detailed information on the content of their outcome measures. For example, the authors of an article that described an RCT of surgery vs watchful waiting for benign prostatic hyperplasia “assessed the degree to which urinary difficulties bothered the patients or interfered with their activities of daily living, sexual function, social activities, and general

well-being.”²⁴ Few would doubt the importance of these items and, since patients in primary care often are untroubled by minor symptoms of benign prostatic hyperplasia, the importance of including them in the results of the trial. As a result, because the RCT showed that surgery was associated with improvement in symptoms and reduced interference with activities of daily living, clinicians should have little difficulty recommending surgery to a patient whose high priority is improving his or her QOL.

Did the QOL Instruments Work in the Intended Way? Evidence concerning the performance of QOL measures is necessary for understanding the measurement properties of these instruments for clinical research. Investigators use QOL instruments in 2 distinct ways. One is to help clinicians distinguish between people who have a better or worse level of QOL (ie, discriminative instruments), and the other is to measure whether people feel better or worse over time (ie, evaluative instruments).²⁵

To be useful, discriminative instruments must be reliable. Differences among patients must be high in relationship to the variability within patients over time. By contrast, instruments used to evaluate change over time must be able to detect any important changes in the way patients are feeling, even if those changes are small. The term we use for the ability to detect change is *responsiveness*.

An unresponsive instrument can result in a false-negative result in which the intervention improves how patients feel, but the instrument fails to detect the improvement. For example, researchers who conducted an RCT of a diabetes education program reported no changes in 2 measures of well-being, attributing the result to, among other factors, lack of integration of the program with standard therapy.²⁶ However, those involved in the educational program, in comparison to a control group, showed an improvement in knowledge and self-care, along with a decrease in feelings of dependence on physicians. Given these changes, another explanation for the investigators' finding of no difference between treatments in well-being might be inadequate responsiveness of the well-being measures.

To be useful, QOL measures must be *valid*, in that they measure what they are intended to measure.²⁷ Evaluating validity involves specifying and testing hypotheses about the relationship between the QOL instrument and other health-related measures, such as disease stage or severity, other QOL measures, or other clinical status measures. For example, we would expect that patients with lower treadmill exercise capacity generally will have more dyspnea in daily life than those with higher exercise capacity. When we are interested in evaluating change over time, we examine correlations of change scores. Patients who experience deterioration in their treadmill exercise

capacity should, in general, show increases in dyspnea, whereas those whose exercise capacity improves should experience less dyspnea. The technical term for this process is testing an instrument's *construct validity*. Although investigators have identified other types of validity, clinicians should focus on evidence of construct validity when assessing the credibility of QOL measures used in clinical studies.

Are There Important Aspects of HRQL That Have Been Omitted? Although investigators may have addressed QOL issues, they may not have done so comprehensively. Disease-specific (ie, instruments tailored to a specific disease, such as asthma, breast cancer, or arthritis), function-specific (eg, instruments that examine sleep or sexual function), or problem-specific (eg, pain or fatigue) QOL measures may comprehensively sample all aspects of QOL relevant to a specific illness and also be responsive, but they are unlikely to deal comprehensively with adverse effects.

Consider, for instance, an RCT of methotrexate in 141 patients with chronically active Crohn disease despite at least 3 prior months of prednisone therapy.²⁸ Patients who received methotrexate were twice as likely to be in clinical remission after 16 weeks of treatment as those who received placebo (39.4% vs 19.1%; $P=.025$), and actively treated patients received less prednisone and showed less disease activity. However, the decision to give methotrexate depends on the balance between the benefits and risks. Patients considering methotrexate treatment will likely wonder about how much better or worse they will feel while taking this medication. Therefore, without information about the effect of the medication on QOL, neither the clinician nor the patient can make a fully informed choice.

In that study, the investigators administered the Inflammatory Bowel Disease Questionnaire, a valid and responsive measure of QOL tailored to the problems of patients with Crohn disease and ulcerative colitis.²⁹ Although the Inflammatory Bowel Disease Questionnaire measures some adverse effects of methotrexate (ie, nausea, lethargy), it does not assess other toxicity such as rash or mouth ulcers. What could the investigators have done to ensure they had captured relevant adverse effects?

One option to capture the impact of adverse effects is to administer generic measures such as *health profiles* that yield scores for all domains of QOL (including, for example, mobility, self-care, and physical, emotional, and social function). However, generic instruments may only cover each QOL domain superficially and may also fail to detect the impact of adverse effects. In addition, the relatively superficial coverage of generic instruments limits their responsiveness. Generic instruments are, on

average, less powerful in detecting treatment effects than specific instruments.³⁰

For instance, in an RCT highly relevant to the dilemma of the patient with chronic obstructive pulmonary disease in the “Clinical Scenario” section, investigators conducted an RCT of respiratory rehabilitation vs conventional community care.³¹ A number of generic measures of QOL, the Sickness Impact Profile, Quality of Well-Being Index, and the Standard Gamble, failed to distinguish between patients in the treatment and control groups. A condition-specific measure that focuses on the problems of patients with respiratory disability, the Chronic Respiratory Questionnaire (CRQ), demonstrated a decrease in dyspnea and fatigue and an improvement in emotional function in patients who participated in rehabilitation.³² If the patient in the “Clinical Scenario” section had only the results of the generic instruments, he would conclude that rehabilitation would not be worth the inconvenience and time. However, with the CRQ results, the question of rehabilitation remains open.

HOW CAN WE INTERPRET THE MAGNITUDE OF EFFECT ON HRQL?

Simply knowing the treatment-associated changes in scores in QOL questionnaires will not, without further information, be useful to clinicians. For instance, one clinical trial showed that patients with acute back pain who were prescribed bed rest had a mean score 3.9 points worse than those of control patients on the Oswestry Back-Disability Index, a measure that focuses on disease-specific functional status.³³ Patients with severe rheumatoid arthritis allocated to treatment with cyclosporine had a mean disability score that was 0.28 unit better than that of control patients.³⁴ This information alone leaves clinicians at a loss as to whether these differences are unimportant, small but important, of moderate magnitude, or reflect large and extremely important differences in efficacy among treatments.

When reading the literature, clinicians must arrive at some general estimates of how a given therapy will affect QOL.³⁵ A number of methods are available for understanding the magnitude of QOL effects. Investigators may relate changes in QOL questionnaire scores to well-known functional measures (eg, the New York Heart Association functional classification), to clinical diagnosis (eg, the change in score associated with a diagnosis of major depressive disorder), to patients’ global ratings of the magnitude of change they have experienced,³⁶ or to the extent that patients rate themselves as feeling better or worse than other patients.³⁷

These comparisons lead to estimates of change in QOL measures that, either for individual patients or for a group of patients, constitute trivial, small, medium, and large

differences. For instance, the chest physician caring for the patient in the “Clinical Scenario” section will be interested in knowing that, by examining the relationship between global ratings of change and change in questionnaire score, investigators have established that an improvement or deterioration of 0.5 approximates a small but important difference in CRQ domains of dyspnea, fatigue, and emotional function, all measured on a 7-point scale.³⁸

This type of information still leaves a problem in the interpretation of results from clinical trials. For instance, if the mean difference between treatment and control on the CRQ is only 0.4, does this imply that we can dismiss the difference as unimportant to patients? Investigators have gained insight into this issue by examining the distribution of change in QOL in individual patients and by calculating the proportion of patients who achieved important QOL benefits from treatment and the associated numbers needed to treat (NNTs). For instance, in the RCT of respiratory rehabilitation mentioned earlier,³¹ the mean differences between rehabilitation and standard community care were 0.60 in CRQ dyspnea, 0.45 in fatigue, and 0.40 in emotional function.

The mean difference of 0.60, however, does not imply that every patient achieved a small but important difference. Similarly, differences below the threshold of 0.5, such as those seen in the rehabilitation study in fatigue and emotional function, do not imply that none of the patients experienced significant improvement. Rather, one would expect a distribution of benefits across patients, some achieving large improvements in QOL and others showing trivial or no gains.

Investigators can examine the distribution of change in treatment and control groups to estimate the proportion of patients who benefited from respiratory rehabilitation. Examining their data, the investigators were able to estimate that 44% of the rehabilitation patients improved in day-to-day shortness of breath, whereas the comparable figure for the control group was 20%. Thus, the net proportion of patients who benefited from rehabilitation was 24%. The proportions of patients who benefited in terms of fatigue and emotional function were 23% and 30%, respectively.³⁹

One can take the inverse of these percentages to determine the NNT to achieve a small but important improvement in a single patient. For instance, one needs to treat 1/0.24 or approximately 4 patients to achieve an important reduction in dyspnea for a single patient. The comparable NNTs for fatigue and emotional function are 4 and 3, respectively. These numbers represent an impressive benefit compared with many other NNTs. For instance, we need to treat 100 average patients with myocardial infarction with tissue-type plasminogen activator rather than

streptokinase to prevent 1 premature death, and we need to screen approximately 300 women for 10 years with mammography to prevent 1 premature breast cancer death.

HOW CAN QOL DATA HELP PATIENTS MAKE INFORMED DECISIONS ABOUT TREATMENT?

Even the most common problems of a chronic disease do not affect all those who are comparably afflicted. For instance, 92% of patients with inflammatory bowel disease experience frequent bowel movements, 82% experience abdominal cramps, 78% feel frustrated, and 76% feel depressed.⁴⁰ Descriptive data about patient experience may help clinicians determine what questions to ask about the patient's health status.

Quality of life can provide unique information that leads to the choice of effective treatments, rejection of ineffective interventions, and clarification of the tradeoffs between management strategies with major benefits and health-related outcome costs. In the remainder of this section, we provide an example of each of these situations.

After conducting a systematic search and applying explicit eligibility criteria, investigators can identify all relevant QOL studies that address a therapeutic controversy. They can then apply meta-analytic methods to QOL data to arrive at a robust estimate of treatment effects. For instance, a group of investigators have summarized the RCTs that compared structured respiratory rehabilitation to conventional community care for patients with severe chronic obstructive pulmonary disease.⁴¹ Their pooled estimate of effect showed an overall average effect slightly greater than the minimal important difference in dyspnea, fatigue, and emotional function. Through application of methods similar to those aforementioned for a single rehabilitation trial, results show that patients can anticipate that, on average, they will achieve small but important reductions in dyspnea and fatigue and improvement in emotional function from the program; chances of improving are approximately 50%.⁴² Although few studies provide information that is so easily interpreted, QOL investigators are working hard to improve the transparency of their presentation.

Knowing that the likelihood of an important benefit from the program is approximately 50% is exactly what the patient in the "Clinical Scenario" requires. He now has to consider whether a 50-50 chance of achieving an important improvement in at least 1 of his problem areas—dyspnea, fatigue, anxiety, and depression—is worth the disruption in his life and the travel to the rehabilitation program. This decision will depend on his values and preferences.

Another example of the application of QOL methods to improve clinical care is that orthopedic surgeons frequently recommend arthroscopic surgery with lavage and debride-

ment for patients with painful osteoarthritis of the knee. Indeed, more than 650,000 of these procedures are performed each year in the United States,⁴³ at a cost of approximately \$5000 per procedure. Investigators reported an RCT in which 180 patients underwent arthroscopy with debridement, arthroscopy with lavage, or placebo arthroscopy.⁴⁴ In the placebo arm, the knee was prepared and draped, 3 incisions were made in the skin, the surgeon asked for the instruments and manipulated the knee, saline was splashed to simulate lavage, and the patient was kept in the operating room for the amount of time required for debridement. Study personnel blinded to treatment group collected outcome data at 2 and 6 weeks and 3, 6, 12, 18, and 24 months postoperatively with a knee-specific pain scale, the pain and walking-bending domains of the Arthritis Impact Measurement Scales,⁴⁵ and the 36-Item Short-Form Health Survey bodily pain and physical function domains.⁴⁶ The investigators found no significant differences between treatment and placebo groups at any follow-up assessment. Patient reports of pain, physical functioning, or walking-bending were virtually identical in the 3 groups (Figure 2). The investigators specified minimal important differences for each outcome measure and noted that the confidence intervals (CIs) around the mean differences between groups excluded the minimal important difference. As previously noted, this approach does not absolutely exclude a subgroup of patients with an important difference, but it certainly makes important differences far less likely. The results of this trial, impossible to perform without use of QOL measures, should have an enormous impact on the use of this procedure and save many patients substantial cost and morbidity.

The optimal treatment of patients with newly diagnosed prostate cancer remains uncertain. Until recently, clinicians relied on observational studies to guide their management. The first RCT, recently completed, compared radical prostatectomy with watchful waiting in 695 men with newly diagnosed early-stage prostatic carcinoma. The study included rigorous safeguards against bias.⁴⁷ Investigators established centralized and concealed randomization, conducted a blinded adjudication of cause of death, and achieved 100% follow-up for death and 87% for measurement of QOL⁴⁸ through a mean follow-up of 6.2 years for mortality and 4 years for QOL. They found a trend toward decreased all-cause mortality (18% vs 15%; relative risk, 0.83; 95% CI, 0.57-1.20; $P=.31$) and a decrease in prostate-specific death rates (9% vs 5%; relative risk, 0.50; 95% CI, 0.27-0.91; $P=.02$), favoring the prostatectomy group. The QOL results were mixed. Sexual dysfunction occurred in 45% of those in the watchful waiting group and in 80% of those who underwent radical

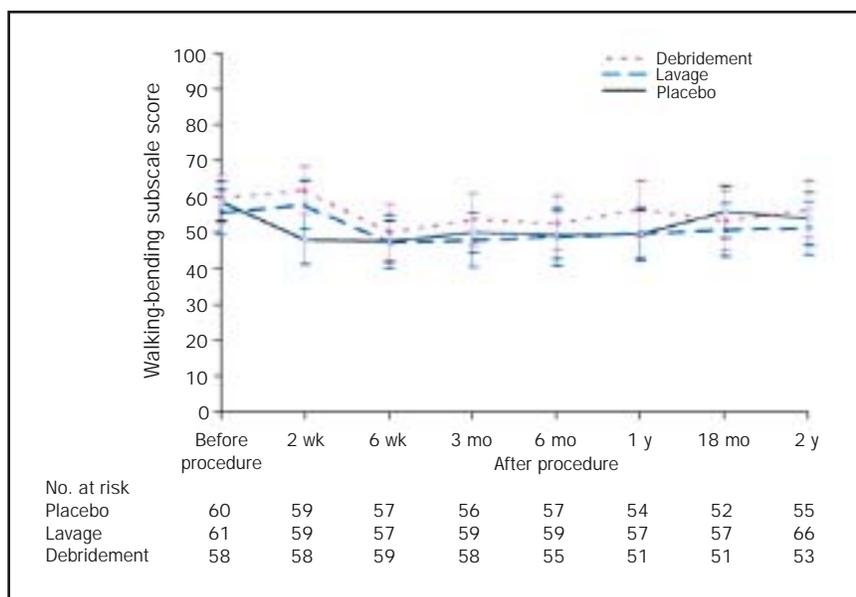


FIGURE 2. Patient scores on the walking-bending subscale of the Arthritis Impact Measurement Scales from a randomized trial of surgical intervention for osteoarthritis of the knee. Error bars indicate 95% confidence intervals. From *N Engl J Med*,⁴⁴ ©2002, Massachusetts Medical Society, all rights reserved.

prostatectomy. Of those in the watchful waiting group, 21% experienced urinary leakage, as did 49% of those in the prostatectomy group. Urinary obstruction, defined as a weak urinary stream, occurred more often in those allocated to watchful waiting (44%) than those who underwent prostatectomy (28%). Bowel function, anxiety, depression, and well-being did not differ between the groups.

These results suggest that men with different values will make different choices when presented with the dilemma of how to deal with newly diagnosed prostate cancer. The investigators put it well, “We cannot say that radical prostatectomy is better than watchful waiting for all men with localized prostate cancer. These alternatives are associated with complex and incommensurable outcomes, and each man must judge for himself which treatment is preferable.” In that trial, detailed QOL measurement provided critical evidence for men facing a difficult and important choice.

ADDED VALUE OF QOL IN CLINICAL PRACTICE

Using QOL in daily clinical practice is another way clinicians and their patients may benefit from measuring HRQL. The extent to which provision to health care professionals of formal information concerning the QOL of their patients improves the process and outcome of medical care is best assessed in RCTs in which clinicians are

randomized to receive or not receive systematic feedback. We identified 2 previous systematic reviews that addressed this general issue,^{49,50} and 2 systematic reviews with a specific focus on routine feedback of psychiatric questionnaires to clinicians in nonpsychiatric settings.^{51,52} Of these, the review by Espallargues et al⁵⁰ was far more comprehensive (it included all the RCTs identified in the other 3 reviews) and still met essential quality criteria for systematic reviews, including explicit eligibility criteria and selective inclusion of high-quality studies.

Espallargues et al⁵⁰ identified 21 RCTs in which patients or practitioners were randomly allocated to an intervention group that received provision of information about patients’ reports of health status or a control group that received or practiced routine care.⁵³⁻⁷⁷ These trials focused on a wide variety of outcomes, including physicians’ and patients’ satisfaction, process of care (diagnosis and treatment), and outcomes (in particular, QOL). We updated and extended this review by systematically searching for all relevant studies published in the 1998-2003 period (not covered in the aforementioned review). We identified 8 additional relevant RCTs,⁷⁸⁻⁸⁶ and after excluding 2 of the studies previously included because of pseudorandomization,^{61,77} we included 27 studies.

Six of these studies randomized individual physicians, 8 practices, and 13 patients. Quality scores on a scale of 0 to 7, in which higher numbers indicated better quality, ranged from 2.0 to 5.0, with a mean of 3.4. One serious quality

problem is that most studies that randomized physicians or groups of physicians (and should therefore have used physicians or groups of physicians as the unit of analysis or at least accounted for clustering of outcomes across physicians or groups) had patients as the unit of analysis. This analytical error substantially increases the likelihood of false-positive findings. Another major problem is that most studies assessed multiple measures of outcome, even within category (ie, multiple measures of satisfaction, process, and outcome), while nevertheless retaining nominal threshold *P* values of .05. This also increases the likelihood of false-positive results.

Of 12 RCTs that examined patient satisfaction, 5 found differences that favored the intervention group. Of 22 RCTs that assessed process of care, 14 detected some difference in process of care between intervention and control groups. Seven reported some impact on the diagnostic process, 2 on health services use, and 3 on advice, education, and counseling. Of the 16 studies that assessed patient outcomes, 7 found at least 1 outcome that was superior in the intervention compared with the control group.

We have conducted a search for RCTs published since 2003 and identified 7 additional studies.⁸⁷⁻⁹³ Three of these RCTs focused on enhanced feedback to psychotherapists⁹⁰⁻⁹² and were performed by the same group of investigators, consistently observing changes in both process and outcome-of-care variables.

A study by Velikova et al⁸⁸ that included 28 oncologists and 286 cancer patients yielded positive results. In this study, information about the patient's HRQL was associated with more frequent discussion about symptoms (without prolonging encounters) and with a better HRQL and emotional functioning at 6 months compared with the control patients. Positive findings were also observed in the studies by MacArthur et al⁹⁰ and more recently by Brodey et al⁸⁷ but not in the large trial performed in the United States that included 985 clinicians and 15,346 patients with feedback of both generic and specific HRQL instruments for 2 years.⁸⁸

If measurement of QOL with feedback to clinicians were resource free, even these mixed results might lead one to recommend the intervention with enthusiasm. Patients and clinicians often appreciate the information, and on occasion the information may influence the process and even the outcome of care. Unfortunately, the process involves resource use, and the constrained resources of the current health care environment suggest we should demand more from QOL feedback than high approval ratings from clinicians and patients. The less the increase in resources required, the less stringent would be the requirements for evidence of benefit of QOL measures in clinical practice: use of computer-based data collection may lower the re-

source burden and therefore encourage more widespread use.

Effects of QOL feedback to clinicians on process and outcome have been inconsistent, and because of methodological problems, including incorrect unit of analysis and multiple comparisons, the estimates provided are likely biased in favor of the intervention. What we can say with confidence is that policymakers, administrators, or physician groups thinking of investing resources to collect QOL data from patients and provide summaries to physicians, at least using methods tested up to now, should not routinely anticipate important improvement in either the process of care or outcome. At the same time, there may be particular approaches and particular settings that would be more likely to produce success.

Why did provision of information fail to improve management or outcome in so many studies? Routine patient interviews may provide physicians with all the information they need about patients' QOL to make appropriate management decisions. Alternatively, time constraints may prevent clinicians from conducting the exploration that, optimally, QOL information should trigger. Perhaps given adequate time, the additional information would ultimately lead to improved outcome. If either of these explanations is accurate, then changes in either the intervention or study design are unlikely to lead to subsequent positive trials.

Other explanations for failure of the interventions include wrong choice of QOL instruments, inadequate time in educating clinicians about how to use the measures, suboptimal presentation of the information, necessity for clinicians having longer duration of experience with the instruments, or inadequate length of follow-up. These explanations, if accurate, would suggest that modifications to the interventions or to the study designs might yield more positive results. A deeper understanding coming from a theory-driven approach, combining knowledge of whether and how an intervention works, might best inform modifications in the interventions or study designs.⁹⁴

Whatever the explanation, the currently available interventions may, in some situations, have important positive health effects. Research to discover the combination of patients, clinicians, and provision of information that one requires for important benefits remains worthwhile.

CONCLUSION

Recent findings from well-designed RCTs suggest that information provided to clinicians on their patients' QOL might, in some circumstances, result in benefits for these patients. Further investigation is warranted to confirm these observations and to define the particular combination

of methods and settings most likely to yield important benefits. Since the major goal of most of our therapeutic interventions is to make patients feel better, QOL measurement is likely to become even more important in the future.

REFERENCES

1. Frost MH, Bonomi AE, Cappelleri JC, Schünemann HJ, Moynihan TJ, Aaronson AK, Clinical Significance Consensus Meeting Group. Applying quality-of-life data formally and systematically into clinical practice. *Mayo Clin Proc.* 2007;82(10):1214-1228.
2. Ferrans CE. Definitions and conceptual models of quality of life. In: Lipscomb J, Gotay CC, Snyder C, eds. *Outcomes Assessment in Cancer*. Cambridge, NY: Cambridge University Press; 2005.
3. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life: a conceptual model of patient outcomes. *JAMA.* 1995;273(1):59-65.
4. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med.* 1993;118:622-629.
5. Osoba D. Lessons learned from measuring health-related quality of life in oncology. *J Clin Oncol.* 1994;12(3):608-616.
6. Fletcher RH, Fletcher SW, Wagner EH. *Clinical Epidemiology: The Essentials*. 3rd ed. Baltimore, MD: Williams & Wilkins; 1996:4-6.
7. Sutherland HJ, Lockwood GA, Boyd NF. Ratings of the importance of quality of life variables: therapeutic implications for patients with breast cancer. *J Clin Epidemiol.* 1990;43(7):661-666.
8. Ganz, PA. Impact of tamoxifen adjuvant therapy on symptoms, functioning and quality of life. *J Natl Cancer Inst Monogr.* 2001;(30):130-134.
9. Hürny C, Bernhard J, Coates AS, et al. International Breast Cancer Study Group. Impact of adjuvant therapy on quality of life in women with node-positive operable breast cancer [published correction appears in *Lancet* 1997; 350(9073):298]. *Lancet.* 1996;347(9011):1279-1284.
10. Kornblith AB, Herr HW, Ofman US, Scher HI, Holland JC. Quality of life of patients with prostate cancer and their spouses: the value of a database in clinical care. *Cancer.* 1994;73(11):2791-2802.
11. Nocturnal Oxygen Therapy Trial Group. Continuous or nocturnal oxygen therapy in hypoxemic chronic obstructive lung disease: a clinical trial. *Ann Intern Med.* 1980;93(3):391-398.
12. Heaton RK, Grant I, McSweeney AJ, Adams KM, Petty TL. Psychologic effects of continuous and nocturnal oxygen therapy in hypoxemic chronic obstructive pulmonary disease. *Arch Intern Med.* 1983;143(10):1941-1947.
13. Mulrow CD, Mulrow JP, Linn WD, Aguilar C, Ramirez G. Relative efficacy of vasodilator therapy in chronic congestive heart failure: implications of randomized trials. *JAMA.* 1988;259(23):3422-3426.
14. Packer M, Carver JR, Rodeheffer RJ, et al. PROMISE Study Research Group. Effect of oral milrinone on mortality in severe chronic heart failure. *N Engl J Med.* 1991;325(21):1468-1475.
15. Juniper EF, Svensson K, O'Byrne PM, et al. Asthma quality of life during 1 year of treatment with budesonide with or without formoterol. *Eur Respir J.* 1999;14(5):1038-1043.
16. Juniper EF, Norman GR, Cox FM, Roberts JN. Comparison of the standard gamble, rating scale, AQLQ and SF-36 for measuring quality of life in asthma. *Eur Respir J.* 2001;18(1):38-44.
17. Juniper EF, Price DB, Stampone PA, Creemers JP, Mol SJ, Fireman P. Clinically important improvements in asthma-specific quality of life, but no difference in conventional clinical indexes in patients changed from conventional beclomethasone dipropionate to approximately half the dose of extrafine beclomethasone dipropionate. *Chest.* 2002;121(6):1824-1832.
18. COMBIVENT Inhalation Solution Study Group. Routine nebulized ipratropium and albuterol together are better than either alone in COPD. *Chest.* 1997;112(6):1514-1521.
19. Jaeschke R, Guyatt GH, Willan A, et al. Effect of increasing doses of beta agonists on spirometric parameters, exercise capacity, and quality of life in patients with chronic airflow limitation. *Thorax.* 1994;49(5):479-484.
20. Cranney A, Guyatt GH, Krolicki N, et al. Osteoporosis Research Advisory Group (ORAG). A meta-analysis of etidronate for the treatment of postmenopausal osteoporosis. *Osteoporos Int.* 2001;12(2):140-151.
21. Cranney A, Guyatt GH, Griffith L, Wells G, Tugwell P, Rosen C, Osteoporosis Methodology Group, Osteoporosis Research Advisory Group. IX: summary of meta-analyses of therapies for postmenopausal osteoporosis. *Endocr Rev.* 2002;23(4):570-578.
22. Guyatt G, Naylor CD, Juniper E, Heyland D, Jaeschke R, Cook D. Quality of life. In: Guyatt G, Rennie D, Evidence-Based Medicine Working Group, eds. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. Chicago, IL: AMA Press; 2002.
23. Guyatt GH, Berman LB, Townsend M, Pugsley SO, Chambers LW. A measure of quality of life for clinical trials in chronic lung disease. *Thorax.* 1987;42(10):773-778.
24. Wasson JH, Reda DJ, Bruskewitz RC, Elinson J, Keller AM, Henderson WG, Veterans Affairs Cooperative Study Group on Transurethral Resection of the Prostate. A comparison of transurethral surgery with watchful waiting for moderate symptoms of benign prostatic hyperplasia. *N Engl J Med.* 1995; 332(2):75-79.
25. Guyatt GH, Kirshner B, Jaeschke R. Measuring health status: what are the necessary measurement properties? *J Clin Epidemiol.* 1992;45(12):1341-1345.
26. de Weerd I, Visser AP, Kok GJ, de Weerd O, van der Veen EA. Randomized controlled multicentre evaluation of an education programme for insulin-treated diabetic patients: effects on metabolic control, quality of life, and costs of therapy. *Diabet Med.* 1991;8(4):338-345.
27. Hays RD, Revicki DA. Reliability and validity, including responsiveness. In: Fayers P, Hays RD, eds. *Assessing Quality of Life in Clinical Trials*. 2nd ed. New York, NY: Oxford University Press; 2005.
28. Feagan BG, Rochon J, Fedorak RN, et al. The North American Crohn's Study Group Investigators. Methotrexate for the treatment of Crohn's disease. *N Engl J Med.* 1995;332(5):292-297.
29. Guyatt G, Mitchell A, Irvine EJ, et al. A new measure of health status for clinical trials in inflammatory bowel disease. *Gastroenterology.* 1989;96(3): 804-810.
30. Wiebe S, Guyatt GH, Weaver B, Matijevic S, Sidwell C. Comparative responsiveness of generic and specific quality-of-life instruments. *J Clin Epidemiol.* 2003;56(1):52-60.
31. Goldstein RS, Gort EH, Stubbing D, Avendano MA, Guyatt GH. Randomised controlled trial of respiratory rehabilitation. *Lancet.* 1994; 344(8934):1394-1397.
32. Guyatt GH, King DR, Feeny DH, Stubbing D, Goldstein RS. Generic and specific measurement of health-related quality of life in a clinical trial of respiratory rehabilitation. *J Clin Epidemiol.* 1999;52(3):187-192.
33. Malmivaara A, Hakkinen U, Aro T, et al. The treatment of acute low back pain—bed rest, exercises, or ordinary activity? *N Engl J Med.* 1995; 332(6):351-355.
34. Tugwell P, Pincus T, Yocum D, et al. Methotrexate-Cyclosporine Combination Study Group. Combination therapy with cyclosporine and methotrexate in severe rheumatoid arthritis. *N Engl J Med.* 1995;333(3):137-141.
35. Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR, Clinical Significance Consensus Meeting Group. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc.* 2002;77(4):371-383.
36. Juniper EF, Guyatt GH, Willan A, Griffith LE. Determining a minimal important change in a disease-specific Quality of Life Questionnaire. *J Clin Epidemiol.* 1994;47(1):81-87.
37. Redelmeier DA, Guyatt GH, Goldstein RS. Assessing the minimal important difference in symptoms: a comparison of two techniques. *J Clin Epidemiol.* 1996;49(11):1215-1219.
38. Jaeschke R, Singer J, Guyatt GH. Measurement of health status: ascertaining the minimal clinically important difference. *Control Clin Trials.* 1989;10(4):407-415.
39. Guyatt GH, Juniper EF, Walter SD, Griffith LE, Goldstein RS. Interpreting treatment effects in randomised trials. *BMJ.* 1998;316(7132):690-693.
40. Mitchell A, Guyatt G, Singer J, et al. Quality of life in patients with inflammatory bowel disease. *J Clin Gastroenterol.* 1988;10(3):306-310.
41. Lacasse Y, Brosseau L, Milne S, et al. Pulmonary rehabilitation for chronic obstructive pulmonary disease. *Cochrane Database Syst Rev.* 2002; (3):CD003793.
42. Norman GR, Gwadry Sridhar F, Guyatt GH, Walter SD. Relation of distribution- and anchor-based approaches in interpretation of changes in health-related quality of life. *Med Care.* 2001;39(10):1039-1047.
43. Owings MF, Kozak LJ. Ambulatory and inpatient procedures in the United States, 1996. *Vital Health Stat 13.* 1998 Nov;(139):1-119.
44. Moseley JB, O'Malley K, Petersen NJ, et al. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med.* 2001; 347(2):81-88.
45. Meenan RF, Mason JH, Anderson JJ, Guccione AA, Kazis LE. AIMS2: the content and properties of a revised and expanded Arthritis Impact Measurement Scales Health Status Questionnaire. *Arthritis Rheum.* 1992;35(1):1-10.
46. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I: conceptual framework and item selection. *Med Care.* 1992;30(6): 473-483.
47. Holmberg L, Bill-Axelsson A, Helgesen F, et al. Scandinavian Prostatic Cancer Group Study Number 4. A randomized trial comparing radical prostatectomy with watchful waiting in early prostate cancer. *N Engl J Med.* 2002; 347(11):781-789.

48. Steineck G, Helgesen F, Adolfsson J, et al. Scandinavian Prostatic Cancer Group Study Number 4. Quality of life after radical prostatectomy or watchful waiting. *N Engl J Med*. 2002;347(11):790-796.
49. Greenhalgh J, Meadows K. The effectiveness of the use of patient-based measures of health in routine practice in improving the process and outcomes of patient care: a literature review. *J Eval Clin Pract*. 1999;5(4):401-416.
50. Espallargues M, Valderas JM, Alonso J. Provision of feedback on perceived health status to health care professionals: a systematic review. *Med Care*. 2000;38(2):175-186.
51. Gilbody SM, House AO, Sheldon TA. Routinely administered questionnaires for depression and anxiety: systematic review. *BMJ*. 2001;322(7283):406-409.
52. Gilbody SM, House AO, Sheldon T. Routine administration of Health Related Quality of Life (HRQoL) and needs assessment instruments to improve psychological outcome—a systematic review. *Psychol Med*. 2002;32(8):1345-1356.
53. Rubenstein LV, Calkins DR, Young RT, et al. Improving patient function: a randomized trial of functional disability screening. *Ann Intern Med*. 1989;111(10):836-842.
54. Kazis LE, Callahan LF, Meenan RF, Pincus T. Health status reports in the care of patients with rheumatoid arthritis. *J Clin Epidemiol*. 1990;43(11):1243-1253.
55. White P, Atherton A, Hewett G, Howells K. Using information from asthma patients: a trial of information feedback in primary care. *BMJ*. 1995;311(7012):1065-1069.
56. Linn LS, Yager J. The effect of screening, sensitization, and feedback on notation of depression. *J Med Educ*. 1980;55(11):942-949.
57. Yager J, Linn LS. Physician-patient agreement about depression: notation in medical records. *Gen Hosp Psychiatry*. 1981;3(4):271-276.
58. Rand EH, Badger LW, Coggins DR. Toward a resolution of contradictions: utility of feedback from the GHQ. *Gen Hosp Psychiatry*. 1988;10(3):189-196.
59. Reifler DR, Kessler HS, Bernhard EJ, Leon AC, Martin GJ. Impact of screening for mental health concerns on health service utilization and functional status in primary care patients. *Arch Intern Med*. 1996;156(22):2593-2599.
60. Dowrick C, Buchan I. Twelve month outcome of depression in general practice: does detection or disclosure make a difference? *BMJ*. 1995;311(7015):1274-1276.
61. Dowrick C. Does testing for depression influence diagnosis or management by general practitioners? *Fam Pract*. 1995;12(4):461-465.
62. Hoepfer EW, Nycz GR, Kessler LG, Burke JD Jr, Pierce WE. The usefulness of screening for mental illness. *Lancet*. 1984;1(8367):33-35.
63. Zung WWK, Magill M, Moore JT, George DT. Recognition and treatment of depression in a family medicine practice. *J Clin Psychiatry*. 1983;44(1):3-6.
64. German PS, Shapiro S, Skinner EA, et al. Detection and management of mental health problems of older patients by primary care providers. *JAMA*. 1987;257(4):489-493.
65. Shapiro S, German PS, Skinner EA, et al. An experiment to change detection and management of mental morbidity in primary care. *Med Care*. 1987;25(4):327-339.
66. Calkins DR, Rubenstein LV, Cleary PD, et al. Functional disability screening of ambulatory patients: a randomized controlled trial in a hospital-based group practice. *J Gen Intern Med*. 1994;9(10):590-592.
67. Mazonson PD, Mathias SD, Fifer SK, Buesching DP, Malek P, Patrick DL. The mental health patient profile: does it change primary care physicians' practice patterns? *J Am Board Fam Pract*. 1996;9(5):336-345.
68. Mathias SD, Fifer SK, Mazonson PD, Lubeck DP, Buesching DP, Patrick DL. Necessary but not sufficient: the effect of screening and feedback on outcomes of primary care patients with untreated anxiety. *J Gen Intern Med*. 1994;9(11):606-615.
69. Brody DS, Lerman CE, Wolfson HG, Caputo GC. Improvement in physicians' counseling of patients with mental health problems. *Arch Intern Med*. 1990;150(5):993-998.
70. Magruder-Habib K, Zung WWK, Feussner JR. Improving physicians' recognition and treatment of depression in general medical care: results from a randomized clinical trial. *Med Care*. 1990;28(3):239-250.
71. Goldsmith G, Brodwick M. Assessing the functional status of older patients with chronic illness. *Fam Med*. 1989;21(1):38-41.
72. Wasson J, Hays R, Rubenstein L, et al. The short-term effect of patient health status assessment in a health maintenance organization. *Qual Life Res*. 1992;1(2):99-106.
73. Rubenstein LV, McCoy JM, Cope DW, et al. Improving patient quality of life with feedback to physicians about functional status. *J Gen Intern Med*. 1995;10(11):607-614.
74. Moore JT, Silimperi DR, Bobula JA. Recognition of depression by family medicine residents: the impact of screening. *J Fam Pract*. 1978;7(3):509-513.
75. Johnstone A, Goldberg D. Psychiatric screening in general practice: a controlled trial. *Lancet*. 1976;1(7960):605-608.
76. Wagner AK, Ehrenberg BL, Tran TA, Bungay KM, Cynn DJ, Rogers WH. Patient-based health status measurement in clinical practice: a study of its impact on epilepsy patients' care. *Qual Life Res*. 1997;6(4):329-341.
77. Callahan CM, Dittus RS, Tierney WM. Primary care physicians' medical decision making for late-life depression. *J Gen Intern Med*. 1996;11(4):218-225.
78. Detmar SB, Muller MJ, Schornagel JH, Wever LDV, Aaronson NK. Health-related quality-of-life assessments and patient-physician communication: a randomized controlled trial [published correction appears in *JAMA* 2003;289(8):987]. *JAMA*. 2002;288(23):3027-3034.
79. Williams JW Jr, Mulrow CD, Kroenke K, et al. Case-finding for depression in primary care: a randomized trial. *Am J Med*. 1999;106(1):36-43.
80. Wasson JH, Stukel TA, Weiss JE, Hays RD, Jette AM, Nelson EC. A randomized trial of the use of patient self-assessment data to improve community practices. *Eff Clin Pract*. 1999;2(1):1-10.
81. Saitz R, Horton NJ, Sullivan LM, Moskowitz MA, Samet JH. Addressing alcohol problems in primary care: a cluster randomized, controlled trial of a systems intervention: the screening and intervention in primary care (SIP) study. *Ann Intern Med*. 2003;138(5):372-382.
82. Smith P. The role of the general health questionnaire in general practice consultations. *Br J Gen Pract*. 1998;48(434):1565-1569.
83. Dailey YM, Humphris GM, Lennon MA. Reducing patients' state anxiety in general dental practice: a randomized controlled trial. *J Dent Res*. 2002;81(5):319-322.
84. Whooley MA, Stone B, Soghikian K. Randomized trial of case-finding for depression in elderly primary care patients. *J Gen Intern Med*. 2000;15(5):293-300.
85. McCusker J, Verdon J, Tousignant P, de Courval LP, Dendukuri N, Belzile E. Rapid emergency department intervention for older people reduces risk of functional decline: results of a multicenter randomized trial. *J Am Geriatr Soc*. 2001;49(10):1272-1281.
86. McCusker J, Dendukuri N, Tousignant P, Verdon J, Poulin de Courval L, Belzile E. Rapid two-stage emergency department intervention for seniors: impact on continuity of care. *Acad Emerg Med*. 2003;10(3):233-243.
87. Brodey BB, Cuffel B, McCulloch J, et al. The acceptability and effectiveness of patient-reported assessments and feedback in a managed behavioral healthcare setting. *Am J Manag Care*. 2005;11(12):774-780.
88. Velikova G, Booth L, Smith AB, et al. Measuring quality of life in routine oncology practice improves communication and patient well-being: a randomized controlled trial. *J Clin Oncol*. 2004;22(4):714-724.
89. Fihn SD, McDonnell MB, Diehr P, et al. Effects of sustained audit/feedback on self-reported health status of primary care patients. *Am J Med*. 2004;116(4):241-248.
90. MacArthur C, Winter HR, Bick DE, et al. Redesigning postnatal care: a randomised controlled trial of protocol-based midwifery-led care focused on individual women's physical and psychological health needs. *Health Technol Assess*. 2003;7(37):1-98.
91. Hawkins EJ, Lambert MJ, Vermeersch DA, Slade KL, Tuttle KC. The therapeutic effects of providing patient progress information to therapists and patients. *Psychother Res*. 2004;14(3):308-327.
92. Lambert MJ, Whipple JL, Smart DW, Vermeersch DA, Nielsen SL, Hawkins EJ. The effects of providing psychotherapists with feedback on patient progress during psychotherapy: are outcomes enhanced? *Psychother Res*. 2001;11(1):49-68.
93. Whipple JL, Lambert MJ, Vermeersch DA, Smart DW, Nielsen SL, Hawkins EJ. Improving the effects of psychotherapy: the use of early identification of treatment failure and problem solving strategies in routine practice. *J Counsel Psychol*. 2003;50(1):59-68.
94. Greenhalgh J, Long AF, Flynn R. The use of patient reported outcome measures in routine clinical practice: lack of impact or lack of theory? *Soc Sci Med*. 2005;60(4):833-843.