Quality of Well-Being in Older People With Osteoarthritis

ERIK J. GROESSI,1 ROBERT M. KAPLAN,2 AND TERRY A. CRONAN2

Objective. To examine the sensitivity of the Quality of Well-Being Scale (QWB) as a measure of health-related quality of life (HRQOL) in people with osteoarthritis (OA).

Methods. The QWB was administered, along with the Arthritis Impact Measurement Scale (AIMS) and other health measures. Health care utilization data were also obtained.

Results. People with OA had a mean QWB score of 0.643. The QWB scores were significantly correlated with total AIMS scores, self-rated health status, health care costs, depression scores, and most AIMS subscales. In addition, changes in QWB scores after 1 year were significantly correlated with changes in total AIMS scores and some AIMS subscales.

Conclusion. The QWB appears to be a useful and sensitive generic, utility-based measure of HRQOL in people with OA.

KEY WORDS. Osteoarthritis; Quality of life; Quality of Well-Being Scale; Outcomes measurement; Health assessment.

INTRODUCTION

Arthritis is the most frequently occurring chronic condition among older Americans, affecting 49% of those older than 65 years (1). Osteoarthritis (OA) has been shown to affect the health status of older persons on 3 primary dimensions: physical disability, psychological disability, and pain (2). Social support and social activity are also affected by arthritis (3,4).

The Arthritis Impact Measurement Scale (AIMS) is a widely used, disease-specific measure that addresses the multiple impacts of arthritis severity (5,6). The AIMS is composed of 11 subscales relating to symptoms and functional impairment associated with arthritis. Although the AIMS may be a sensitive measure of arthritis symptomatology, it does not provide a scaled score that can be easily used in cost-effectiveness analysis. Therefore, many researchers recommend using both a generic and a disease-specific measure (7,8).

The Quality of Well-Being Scale (QWB) is a comprehensive, generic measure of health-related quality of life (HRQOL) that combines information about symptoms and functioning into a single-scaled score that is independent of diagnosis. This independence is useful for comparing quality of life across illnesses, treatments, and populations. The QWB is linked to the concept of quality-adjusted life years (QALYs). QALYs combine quality and quantity of life into a single index that adjusts survival time for reduced life quality. The cost of an intervention or treatment can be divided by the number of QALYs lost to a particular health condition or produced by an intervention to estimate the cost per QALY. This value can then be directly compared among different options that compete for health care resources.

The present study examines the sensitivity of a generic outcome measure (QWB) for assessing HRQOL in people with OA. Although the QWB has been validated as a general measure of HRQOL with several other specific diseases, its sensitivity has not been validated in people with OA. This article also provides an estimate of the impact of OA on HRQOL.

SUBJECTS AND METHODS

Subjects. Members of a Southern California health maintenance organization (HMO) agreed to participate in an intervention testing the effects of social support and education on health and health care utilization. The results of the intervention aspects of the study are reported elsewhere (9–12). To be eligible, HMO members needed a diagnosis of OA, which was defined as self-reported chronic pain and stiffness and being told by a physician...
that they had OA. Diagnoses were confirmed in 90% of participants through medical records. Informed consent was obtained from all 363 patients. The mean age of the participants was 69.2 years (SD 5.6) and 75% were retired. Other demographic characteristics of the sample are summarized in Table 1.

**Methods. Quality of Well-Being Scale.** The QWB is a generic measure of HRQOL that combines preference-weighted values for symptoms and functioning (13). Symptoms are assessed by questions about 27 different symptom complexes over the 6 days prior to the assessment. Functioning is assessed by a series of questions designed to record functional limitations over the previous 6 days, within 3 separate domains (mobility, physical activity, and social activity). The 4 domain scores (3 functioning + 1 symptom) are weighted by preference to create a total score that provides an expression of well-being that ranges from 0 for death to 1.0 for asymptomatic optimal functioning. The preference weights were obtained from the ratings of 856 people randomly sampled from the general population (13). In the General Health Policy Model, total QWB scores are integrated with the number of people affected and the duration of time affected to produce the output measure of QALYs. QALYs combine morbidity and mortality outcomes into a single number (14–16).

**Arthritis Measurement Impact Scale.** The self-administered AIMS consists of 57 questions categorized into 9 subscales: mobility, physical activity, dexterity, social role, social activity, activities of daily living, pain, depression, and anxiety. Two additional subscales have been subsequently derived (general health and illness susceptibility). Evaluations have shown the AIMS to be internally consistent with substantial test-retest (α > 0.80) and internal consistency reliabilities (α > 0.70) (5). Internal reliability ranged from 0.63 to 0.88 for the 9 subscales (7). The AIMS has been validated by comparing AIMS scores with the Health Assessment Questionnaire (2) and physician estimates of functional level and disease activity (6).

**The Center for Epidemiologic Studies-Depression Scale (CES-D).** The CES-D is a self-administered measure of depression used in general population surveys (17). The measure consists of 20 items that were previously included in a number of longer depression scales (18). A 4-point Likert scale (0 = rarely or none of the time, and 3 = most or all of the time), is used to assess the frequency of each symptom (item) over the past week. After 4 items are reverse coded, the scores for all items are summed for an overall score that ranges from 0 to 60. Radloff (17) reported high internal consistency (α = 0.85–0.90), moderate test-retest reliability (α = 0.51–0.67), and high concurrent and construct validity for the CES-D.

**Self-rated health.** In addition to the standardized scales, patients rated their own health using the categories excellent, very good, good, fair, and poor. The measure was scored as an ordinal scale with excellent assigned the value 5 and poor assigned 1. Self-rated health was measured at the 1-year assessment, but not at baseline.

**Health care costs.** Health care utilization data were obtained from the HMO for all participants 1 year before and 1 year after they entered the study. Health care costs were estimated by multiplying specific health care service usage by estimates of the 1992 national average cost for each service (19). The different health care services and corresponding costs were grouped as follows: hospital days ($1,436); emergency room visits ($650); primary care and urgent care physician visits ($100); nurse practitioner, physician’s assistant, and nurse visits ($50); phone contact ($25); home visit ($250); and other hospital visits ($100).
Analysis. This article focuses on baseline and 1-year followup data. However, the global self-rating of health question was not asked at the baseline assessment; only the 1-year post-assessment was available for analysis. As a result, change scores were not available for the self-rated health status question. Scores on the CES-D at the 1-year assessment were available for only half of the participants, and therefore, were not examined. All correlations use the Pearson product-moment calculation method. An $\alpha$ of 0.01 was used for testing the significance of the correlations.

RESULTS

The mean QWB score for the 363 patients at the baseline assessment was 0.643 (SD 0.090). Scores ranged from 0.42 to 1.0, the median score was 0.641, and skewness was 0.447. The mean score suggests that the quality of life for an older person with OA in our sample is lower than that of a community cohort similar in age (0.72) and lower than the mean for adults with chronic obstructive pulmonary disease (0.66) (20,21). Quality of life for the OA sample was similar to that of adult inpatients with depression (0.64) and to that of adults with advanced cancer (0.63), but higher than that of people with fibromyalgia (0.56), acquired immunodeficiency syndrome (0.61), and Alzheimer’s disease (0.51) (14,16,22-24).

The present sample did not appear to be particularly prone to depression, as measured by the CES-D (mean 8.61). Some researchers argue that scores as low as 16 on the CES-D suggest the possibility of clinical depression (17,25). Only about 15.4% of the patients were at or above this level, which is near the range of 9% to 14.7% that is expected for older people without chronic pain, but quite a bit lower than the range of 39% to 87% that has been observed in people with chronic pain (26-29).

The mean AIMS total score was 93.9 (SD 15.2) with scores ranging from 60 to 136. The median score was 92.0 and skewness was 0.57. The AIMS total scores were divided into quintiles to examine whether QWB scores were sensitive to different levels of OA severity. Each quintile was significantly different from all nonadjacent quintiles ($F[4,354] = 24.1, P < 0.001$; see Figure 1).

Table 2 presents the correlations between the baseline scores for the QWB, the AIMS total score and subscales, and the CES-D. The QWB was significantly correlated with 10 of the 11 AIMS subscales and the CES-D depression score. The negative correlations reflect the fact that higher scores on those measures indicate more impairment, whereas higher QWB scores are associated with fewer symptoms and less impairment.

The QWB was also negatively correlated with health care costs for the year prior to the baseline assessment, and with health care costs for the year following the assessment. Although the correlation is small (−0.16), the relationship appears consistent and linear ($t = 3.54, P < 0.001$).

The data in Table 2 indicate that the AIMS total score

<table>
<thead>
<tr>
<th>Variable</th>
<th>QWB</th>
<th>AIMS total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIMS total score</td>
<td>−0.477†</td>
<td>−0.65†</td>
</tr>
<tr>
<td>AIMS physical activity subscale</td>
<td>−0.61†</td>
<td>−0.67†</td>
</tr>
<tr>
<td>AIMS health subscale</td>
<td>−0.52†</td>
<td>−0.48†</td>
</tr>
<tr>
<td>AIMS social role subscale</td>
<td>−0.46†</td>
<td>−0.41†</td>
</tr>
<tr>
<td>AIMS pain subscale</td>
<td>−0.43†</td>
<td>−0.48†</td>
</tr>
<tr>
<td>AIMS mobility subscale</td>
<td>−0.36†</td>
<td>−0.48†</td>
</tr>
<tr>
<td>AIMS daily living subscale</td>
<td>−0.26†</td>
<td>−0.38†</td>
</tr>
<tr>
<td>AIMS depression subscale</td>
<td>−0.25†</td>
<td>−0.45†</td>
</tr>
<tr>
<td>AIMS anxiety subscale</td>
<td>−0.25†</td>
<td>−0.44†</td>
</tr>
<tr>
<td>AIMS dexterity subscale</td>
<td>−0.21†</td>
<td>−0.44†</td>
</tr>
<tr>
<td>AIMS social activity subscale</td>
<td>−0.14†</td>
<td>−0.39†</td>
</tr>
<tr>
<td>AIMS illness susceptibility subscale</td>
<td>−0.09†</td>
<td>−0.40†</td>
</tr>
<tr>
<td>Health care costs (1 year prior to baseline assessment)</td>
<td>−0.15†</td>
<td>0.107†</td>
</tr>
<tr>
<td>Health care costs (1 year after baseline assessment)</td>
<td>−0.15†</td>
<td>0.08†</td>
</tr>
<tr>
<td>CES-D</td>
<td>−0.20†</td>
<td>0.65††</td>
</tr>
<tr>
<td>Self-rated health status (1 year after assessment)</td>
<td>0.36†‡</td>
<td>−0.49‡</td>
</tr>
</tbody>
</table>

* QWB = Quality of Well-Being Scale; AIMS = Arthritis Impact Measurement Scale; CES-D = Center for Epidemiologic Studies-Depression Scale.
† Denotes significance at $P < 0.01$.
‡ Correlation based on 1-year assessment because self-rated health was not measured at baseline.
may also be related to health care costs, but the strength of the relationship is less than that observed for the QWB. Neither the AIMS nor the QWB were significantly correlated with changes in health care costs.

Table 3 presents the correlations between the changes in QWB, health care costs, and the AIMS total and subscale scores. The results indicate that, despite the increase in measurement error brought about by examining change, the relationship is less than that observed for the QWB. Although statistically significant, several of the correlations in Table 2 are relatively low. In particular, the results indicate that the QWB was designed to capture all sources of HRQOL, whereas the AIMS is more focused on the impact of arthritis. However, the impact of OA on health care costs may be more pronounced in those with severe conditions, and the attendance requirements of the intervention made it more difficult for those with severe OA to participate. Therefore, selection bias must be considered. Although the correlation between health care costs and the QWB is small, the finding supports the sensitivity of the QWB because health care costs have proven very difficult to predict and are often unrelated to quality of life measures.

The present study suggests that the QWB is sensitive to OA symptoms and severity. Figure 1 shows that mean QWB scores differ by as much as 0.12 between patients with AIMS scores below 81 and those above 107. Each quintile differs from those adjacent to it by about 0.030, which surpasses the minimal clinically important difference of 0.023 for the QWB (30).

The QWB provides some symptom data, but the AIMS may be more sensitive to finer gradations in arthritis-specific complaints. Despite the many advantages of disease-specific measures, they cannot be used for general policy analysis. Cost-utility analysis typically requires a generic measure that can be used to calculate QALYs. Evidence from this study supports the validity and responsiveness of the QWB as an outcome measure for studies of patients with OA. The data suggest that the QWB is correlated with AIMS. Despite these findings, many of the correlations between the QWB and AIMS are modest. Thus, the QWB is not a substitute for a disease-specific outcome measure. For studies that might be used for policy analysis, both types of measures should be used (8).

Although statistically significant, several of the correlations in Table 2 are relatively low. In particular, the rela-
relationships between the QWB and the AIMS social activities scales, and other aspects of emotional functioning, are quite modest. It might be argued that these correlations are of lesser clinical importance. However, the AIMS also has significant measurement error, and reductions in correlation due to attenuation are common. Relationships between the QWB and these scales are approximately equivalent to the relationship between physical functioning on the Medical Outcomes Study Short-Form 36 (SF-36) and the social and emotional variables from other arthritis measures (31). Although the QWB tends to show only modest correlations with the social and emotional scales of the SF-36 and the AIMS, it is substantially correlated with primary measures of depression (22) and other aspects of mental health (32).

Our results are limited by the manner in which patients were selected for the study. The patients were randomly selected to receive an invitation to participate, but then participation was voluntary. In addition, participation required travel of various distances (mean 14.0 miles, SD 12.4) to meetings on a weekly basis for 10 weeks and then on a monthly basis for 10 months. An earlier study (10) compared the volunteers for the study with a nonvolunteer group of patients with OA from the same HMO. The volunteers were more educated, had been members of the HMO longer, had similar but slightly fewer comorbid conditions, and had more health care contacts, although the average cost per contact was much lower than for nonvolunteers. There was also some evidence that the present sample may have less depression than the average chronic pain population. Further validation is needed in a representative cohort with a range of disease severity and duration, including those with significant comorbidities.

A second limitation is the sample size. The sample size was substantial for testing the effects of the intervention, but future QWB validation research with a larger sample of OA patients would be helpful.

A third limitation was the relatively unrefined measure of health care costs. Our estimates did not include costs associated with medications, radiology/imaging, laboratory tests, specialist visits, costs of non-HMO providers, and indirect costs. Unfortunately, these data were not available from the HMO, and we recognize the limited generalizability of our cost estimates. It is also likely that some patients had secondary health care coverage and utilization that we did not measure. However, secondary coverage rates at this HMO typically are low. Therefore, results from this study should be replicated using a more complete cost accounting system.

Although the QWB has been validated in a wide variety of samples with different primary disease states, administration and scoring of the instrument are more time consuming and costly (requiring a trained interviewer) than they are for some other widely used quality of life measures, including the SF-36 (33,34). However, this extra cost has been justified because of the usefulness of the QWB for facilitating cost-effectiveness analysis and public policy/medical decision making. In addition, it allows for comparisons across disease groups and intervention types.

One of the advantages of the QWB is that it links to a general health policy model. The model can be used to estimate QALYs lost to arthritis. For example, people with OA in the current study had a mean QWB score of 0.64. This suggests that they lose about 0.36 equivalents of a life year for each year they are in this condition instead of perfect health. However, evidence from a community cohort suggests that the average person in this age range has a QWB score of 0.72 (20). Thus, the loss attributable to OA is approximately 0.08 QALYs/year. These calculations suggest that over the course of 12 years, the equivalent of 1 year of life is lost for an arthritis patient. Collectively, 12 arthritis patients lose the equivalent of 1 QALY each year they have the condition.

The SF-36 is another generic measure of HRQOL (31) and a number of studies suggest that it is comparable in sensitivity to disease-specific measures (7,35). In addition, an arthritis-specific index for the SF-36 has been developed and validated (36,37). However, the output produced by the SF-36 and its derivatives consists of a series of descriptive profiles from which a single utility weighted score cannot be directly derived. Hence, the SF-36 is not useful for cost-effectiveness analyses.

Although the SF-36 cannot be directly used to estimate QALYs, Fryback et al (38) developed a method to estimate QWB scores from SF-36 data. The disadvantage is that approximately 40% of the variance in QWB scores is left unexplained and significant measurement error is introduced. In other words, the mapping methods lack precision and increase measurement error. Ultimately, this reduces the power of the analysis. Thus, it is preferable to have original data using the QWB or another utility-based measure.

Recently, a self-administered version of the QWB has been developed to reduce the cost of assessment and to improve symptom measurement (39). The self-administered version is highly correlated with the interviewer-administered QWB and has been validated in a number of disease-state samples (40,41).

In summary, the QWB is appropriate for measuring health outcomes in patients with OA. It is correlated with disease-specific measures and health care costs. Unlike disease-specific measures, it can be used for cost-effectiveness analysis.

ACKNOWLEDGMENT
We thank Kaiser Permanente for their cooperation and assistance in conducting this study, and all the research assistants and participants who made this study possible.

REFERENCES
5. Meenan RF, Gertman PM, Mason JH, Dunaf R. The arthritis


