Validity of the Quality of Well-Being Scale for Patients With Alzheimer's Disease

DAVID N. KERNER
San Diego State University
University of California, San Diego

THOMAS L. PATTERSON, PhD
IGOR GRANT, MD
ROBERT M. KAPLAN, PhD
University of California, San Diego

The Quality of Well-Being Scale (QWB) is a utility-weighted measure of health-related quality of life that can be used in clinical trials, population studies, and cost/utility analyses. This article reports evidence for the validity of the QWB in patients with Alzheimer's disease. The subjects were 211 patient/spouse dyads and control dyads recruited from the University of California, San Diego, Alzheimer's Disease Research Center (ADRC) and from community referrals. Among these, three quarters were patients, and one quarter were age- and gender-matched controls. Patient data were obtained by caregiver proxy. Analyses demonstrated that the QWB was strongly associated with dementia ratings and behavioral problems. Caretakers of patients with low QWB scores also reported using more respite time. We conclude that the general QWB score allows data from Alzheimer's disease studies to be used in comparative cost/utility analysis.

Alzheimer's disease is a degenerative brain disorder that results in gradual atrophy of higher cortical regions. It is marked by gradual

AUTHORS' NOTE: The authors wish to thank the members of the Alzheimer's Caregiver Project staff, who assisted in gathering data for this project. Research was supported by NIMH Grant #MH42840 (Igor Grant, principal investigator) and NIH Grant AG0 5131 to the UCSD Alzheimer's Disease Research Center (Leon Thal, principal investigator). Address Correspondence to: Robert M. Kaplan, Ph.D., Division of Health Care Sciences, 0622, Department of Family and Preventive Medicine, University of California, San Diego, La Jolla, CA 92093-0622.
onset and a deteriorating course. The first symptoms of Alzheimer’s disease include forgetfulness, anomia, irrationality, loss of initiative, and disorientation. These initial symptoms progress to widespread dementia, loss of functioning, and death. An amnesic syndrome is most prominent in many Alzheimer’s disease patients; in others, naming and spatial difficulties are primary (Schwartz, Baron, & Moscovitch, 1989). Alzheimer’s disease is more common than previously believed, affecting about 15% of adults over age 64 (Evans, 1990). In the general population, this risk rises to 25% at age 90. In first-degree relatives of Alzheimer’s disease patients, the risk at age 90 is 50% (Mohs, Breitner, Silverman, & Davis, 1987).

The purpose of this article is to quantify the impact of Alzheimer’s disease. In contrast to diseases that cause early death, Alzheimer’s disease causes people to lose function gradually over an extended period of time. An exclusive focus on mortality does not fully recognize the serious impact Alzheimer’s disease has on health-related quality of life. To understand the full impact of Alzheimer’s disease, it is necessary to develop models that consider the effects on both mortality and life quality. Similarly, new interventions developed for the treatment of Alzheimer’s disease need to be evaluated with measures that consider side effects as well as the benefits.

Questions about the cost and effectiveness of medical care have brought considerable attention to medical outcomes research. Investigators in both the public and private sectors have struggled to find appropriate methodologies to evaluate health care technologies. In 1993, the Department of Health and Human Services appointed a multidisciplinary group of methodologists to recommend standardized strategies for the evaluation of health care. The panel, which released its report in 1996 (Gold, Siegel, Russel, & Weinstein, 1996), suggested that standardized outcomes analyses be conducted to evaluate the cost/effectiveness of medical care. These analyses require preference-weighted measures of health-related quality of life, such as the one used in this study. Although there has been considerable interest in measuring the cost/effectiveness of treatments for Alzheimer’s disease, little is known about the validity of general outcomes measures for these patients.
The approach applied in this analysis conceptualizes outcomes in terms of Quality-Adjusted Life Years (QALYs). QALYs integrate mortality and morbidity to express health status in terms of equivalents of well years of life. A variety of methods have been used to estimate QALYs. In this report, we use a method known as the Quality of Well-being Scale (QWB). Although the scale is only one piece of the total model, it has particular importance for the measurement of clinical outcomes. Ultimately, we hope to compare the impacts of different diseases and the effects of their treatments. Traditionally, researchers have used measures that are unique to a given illness. These disease-specific measures have limited usefulness when making comparisons among different disease states. The QWB is a general measure that may conceivably be used in any disease population. However, this requires that the validity of the QWB be established for particular illnesses. The remainder of this article offers validity evidence for the use of the QWB for evaluating patients with Alzheimer's disease.

The ultimate goal of QWB assessment in Alzheimer’s disease is to develop a single number that represents the current impact of disease. Furthermore, when integrated over time, this number represents the combined effects of morbidity and mortality on patients with Alzheimer’s disease. Ultimately, this will allow an expression of the impact of Alzheimer’s disease on the number of QALYs lost to this disease. However, application of this model in Alzheimer’s disease requires specific validation for this patient group. Previous articles (Kaplan et al., 1989; Kaplan, Atkins, & Timms, 1984; Kaplan, Bush, & Berry, 1976; Kaplan, McCutchan, Navarro, & Anderson, 1994; Orenstein, Nixon, Ross, & Kaplan, 1989) have offered general validity evidence for the QWB system, as well as evidence for the validity in specific disease categories. Construct validity is established by showing systematic relationships between a variable and other measures that presumably represent the same construct. These associations define the meaning of the construct. The analyses presented below are used to evaluate the meaning of the QWB for studies in Alzheimer’s disease.
Method

SUBJECTS AND MEASURES

The subjects were 159 patients with the diagnosis of probable or possible Alzheimer’s disease and their spousal caregivers, along with 52 control nonpatient/spouse dyads (total = 211) recruited as part of a longitudinal study on Alzheimer’s disease caregiving. Couples had been married for an average of 47 years. About 55% of these subjects were recruited from the University of California, San Diego (UCSD), Alzheimer’s Disease Research Center (ADRC), where they were diagnosed by staff neurologists and neuropsychologists. Diagnosis of questionable, possible, or probable Alzheimer’s disease were made using NINCDS-ADRDA criteria (McKhann et al., 1984). The remainder of the subjects were obtained through community support groups or physician referrals, with the caregiver reporting a prior physician diagnosis of probable or possible Alzheimer’s disease for their spouse. Subjects also met the following eligibility requirements: (a) spousal caregivers were providing home care for the patient at the baseline assessment, (b) both patient and caregiver were willing to participate in a longitudinal study involving assessments at 6-month intervals, (c) both patient and caregiver could communicate in English, and (d) patients and caregivers resided within the San Diego, California, metropolitan area.

Control subjects were 52 age-matched married couples selected from neighborhoods sociodemographically matched to the caregivers. Controls were chosen from a group of 150 volunteers available through the ADRC and other geriatric research centers affiliated with UCSD. From each control couple, one spouse was randomly assigned as nonpatient and the other as noncaregiver.

The average age of patients was 73.7 years ($SD = 7.1$); the controls were somewhat younger, with an average age of 69.3 years ($SD = 7.3$). The majority (64.3%) of patients were male. The controls were evenly split by gender. The subjects were, on average, middle class (raw score on the Hollingshead Two Factor Index of Social Position = 36.3;
Table 1

Demographic Data for Patients and Controls

<table>
<thead>
<tr>
<th></th>
<th>Patients</th>
<th></th>
<th>Controls</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Age</td>
<td>73.7</td>
<td>7.1</td>
<td>69.3</td>
<td>7.3</td>
</tr>
<tr>
<td>Gender (percentage males)</td>
<td>64.3</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Socioeconomic status</td>
<td>38.6</td>
<td>16.5</td>
<td>29.4</td>
<td>15.7</td>
</tr>
<tr>
<td>Length of time since diagnosis (years)</td>
<td>8.4</td>
<td>3.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample size</td>
<td>159</td>
<td></td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>Total sample (used for analyses)</td>
<td>211</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Socioeconomic status = Hollingshead Two Factor Index of Social Position (Hollingshead & Redlich, 1958).

Table 2

Patients' and Controls' Clinical Dementia Rating (CDR) Scores

<table>
<thead>
<tr>
<th>CDR</th>
<th>N</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>44</td>
<td>20.8</td>
</tr>
<tr>
<td>1</td>
<td>19</td>
<td>9.0</td>
</tr>
<tr>
<td>2</td>
<td>41</td>
<td>19.4</td>
</tr>
<tr>
<td>3</td>
<td>40</td>
<td>18.9</td>
</tr>
<tr>
<td>4</td>
<td>45</td>
<td>21.3</td>
</tr>
<tr>
<td>Unavailable</td>
<td>22</td>
<td>10.4</td>
</tr>
<tr>
<td>Total N: 211</td>
<td></td>
<td>100.0</td>
</tr>
</tbody>
</table>

Note. Data were not available for 8 controls and 14 patients.

$SD = 16.8$) (Hollingshead & Redlich, 1958). Table 1 provides a summary of participant demographics. Data about each patient was obtained by proxy from each patient's spousal caregiver. Information was likewise gathered by proxy from spouses of nonpatients. All participants, including both patient/caregiver dyads and nonpatient/noncaregiver dyads, completed or were administered the QWB, along with a battery of disease specific-measures, including the Mattis Dementia Rating Scale (MDRS), the Memory and Behavior Problems Checklist (MBPC), and the Brief Symptom Inventory (BSI). Measures of respite time for caregivers were also obtained. Subject dementia ratings are given in Table 2. Because some dementia ratings were
made by the ADRC on a schedule separated in time from the caregiver data, dementia ratings were unavailable for 8 controls and 14 patients.

Quality of Well-Being Scale

The QWB is a comprehensive measure of health-related quality of life that includes several components. First, it obtains observable levels of functioning at a point in time (usually the previous 6 days) from three separate scales: Mobility, Physical Activity, and Social Activity. These scales each contain a variety of items. Second, each patient identifies his or her most undesirable symptom or problem from a list of 27 items. Then, the observed level of function and the subjective symptomatic complaint are weighted by preference, or the utility for the state, on a scale ranging from 0 (for dead) to 1.0 (for optimum function). The weights have been obtained from independent samples of 856 judges who rated the desirability of observable health states. Several studies have shown that the weights do not vary as a function of demographic variables, including race, income, and gender. Furthermore, most evidence indicates that the weights do not vary systematically as a function of prior experience with the rated health state (Kaplan, 1994). Using this system, it is possible to place the general health status of any individual on the continuum between death and optimal functioning for any point in time.

The QWB has been used in a wide variety of clinical and population studies (Anderson, Kaplan, Berry, Bush, & Rumbaut, 1989; Erickson, Kendall, Anderson, & Kaplan, 1989). In addition, the QWB has been used in clinical trials and studies to evaluate therapeutic interventions in a wide range of medical and surgical conditions. These include chronic obstructive pulmonary disease (Kaplan et al., 1984), AIDS (Kaplan et al., 1995), cystic fibrosis (Orenstein et al., 1989), diabetes mellitus (Kaplan, Hartwell, Wilson, & Wallace, 1987), atrial fibrillation (Ganiats, Palinkas, & Kaplan, 1992), lung transplantation (Squier et al., 1995), arthritis (Bombardier et al., 1986; Kaplan, Kozin, & Anderson, 1988), cancer (Kaplan, 1993a), sinus disease (Hodgkin, 1994), schizophrenia (Patterson et al., 1996), and a wide variety of other conditions (Kaplan, 1993b). In cystic fibrosis patients, for ex-
ample, QWB scores have been shown to be positively correlated with pulmonary function and exercise tolerance and negatively correlated with age. Furthermore, the QWB was able to track improvement over time in a 2-week intervention for treatment of pulmonary exacerbation, with changes in QWB statistically significantly correlated with changes in pulmonary function (Orenstein & Kaplan, 1991; Orenstein et al., 1989).

Validity of QWB items for older adults was reported in a study of 71 older adults who completed the QWB questionnaire and then were observed performing a variety of tasks, including tying shoes; buttoning; hooking; using safety pins, zippers, and Velcro; walking; and arising. There were significant linear associations between QWB scores and time to walk 30 feet ($p < .01$), manual dexterity ($p < .002$), grip strength ($p < .01$), need for assistance in standing ($p < .01$), and several other observed behaviors. Andresen and colleagues administered the QWB and several other measures to 200 older adults in the Group Health Cooperative in Seattle. They found that the QWB was significantly related to independent assessments of chronic disease (Andresen, Patrick, Carter, & Malmgren, 1995). A variety of studies demonstrate that the QWB has a reliability exceeding .90 for most populations (Anderson et al., 1989).

**MATTIS DEMENTIA RATING SCALE**

The MDRS (Coblentz et al., 1973) consists of a mental status examination that assesses basic cognitive functions such as recent and remote memory, attention, orientation, mental control, and language. It is administered by a trained observer, either a neuropsychologist or a neuropsychology technician. The test-retest reliability coefficients for its various subscales range from .61 (for attention) to .94 (for conceptualization). Concurrent validity has been established with the Wechsler Adult Intelligence Scale ($r = .86$).

**MEMORY AND BEHAVIOR PROBLEM CHECKLIST**

The MBPC (Zarit, Reever, & Bach-Peterson, 1980) consists of 29 common problems encountered in dementia. The spousal caregiver (or, in the case of controls, the spousal noncaregiver) completed these
items by proxy for the reference subject. This measure assesses both the frequency of occurrence of the dementia problems in the patient and the degree of stress that the caregivers experience. Each problem and related stress is rated on a 5-point scale. The measure has content and concurrent validity and a reliability coefficient alpha of .78.

**BRIEF SYMPTOM INVENTORY**

The BSI is a 53-item self-report inventory taken from the Hopkins Symptom Checklist (Derogatis, Lipman, Rickels, Uhlenhuth, & Kovi, 1974; Derogatis, Yevzeroff, & Wittelsberger, 1975). It measures five dimensions of psychiatric symptom distress that are commonly associated with Alzheimer’s disease: anxiety, depression, obsessive-compulsiveness, somatization, and interpersonal difficulty. Subjects (caregiver proxies for the patients and noncaregiver proxies for the controls) rate each item on how distressing it has been over the past 6 months. Sample items include nervousness, feelings easily hurt, numbness in body parts, and urge to break things. Internal consistency reliabilities of the five symptom dimensions have been measured by coefficients alpha (.84 to .87). Both criterion-related validity and construct validity have been evaluated in psychiatric outpatients, other outpatients, and normal controls.

**RESPITE TIME**

Measures of respite time taken by patient caregivers were also obtained. Respite time refers to the amount of relief that caregivers took from the burden of caring for the patients with Alzheimer’s disease. Respite time taken may increase as the disease progresses, making it an indirect measure of the severity of the disease. However, other factors, such as the health of the caregiver and the availability of resources, also may influence the amount of respite time taken. Not much is known about the reliability and validity of respite time as a marker of disease progression in Alzheimer’s disease. Therefore, results concerning a relationship between respite time and the QWB must be interpreted with caution.

To assess respite time, the caregivers were asked, “On average, how many hours per day do you spend away from your husband/wife?”
Table 3
Means and Standard Deviations for Study Measures

<table>
<thead>
<tr>
<th></th>
<th>Patients</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Quality of Well-Being Scale</td>
<td>0.51</td>
<td>0.06</td>
</tr>
<tr>
<td>Mattis Dementia Rating Scale</td>
<td>87.7</td>
<td>36.8</td>
</tr>
<tr>
<td>Memory/Behavior Problem Checklist</td>
<td>1.56</td>
<td>0.75</td>
</tr>
<tr>
<td>Brief Symptom Inventory</td>
<td>1.45</td>
<td>0.38</td>
</tr>
<tr>
<td>Respite timea</td>
<td>2.4</td>
<td>1.3</td>
</tr>
</tbody>
</table>

a. Respite time was assessed on a 0 to 4 Likert-type scale, with 0 = no hours/day required and 4 = 19-24 hours/day.

The options were 0 hours, 1 to 6 hours, 7 to 12 hours, 13 to 18 hours, and 19 to 24 hours per day.

Results

Parametric tests were used to evaluate the relationships between the QWB and the other measures. Subjects who completed each measure of interest were included in the analysis. The means and standard deviations for all measures are shown in Table 3. Scores on the QWB were found to be strongly associated with measures of different aspects of impairment resulting from Alzheimer’s disease.

Patients with poorer cognitive functioning in areas such as recent and remote memory, attention, orientation, mental control, and language, as measured by the MDRS, tended to have lower QWB scores ($r = .52, p < .01$).

Lower QWB scores were also associated with greater behavioral impairment, as measured by the MBPC ($r = .64, p < .01$). Strong relationships were found between the QWB and MBPC items. For each of these items, mean QWB scores were calculated for each of the five MBPC response options. The means were compared by analysis of variance. Differences were found for many items, such as Losing Things, $F(4, 123) = 15.49, p < .001$; Asking Repeatedly, $F(4, 122) = 36.43, p < .001$; Forgetting Day, $F(4, 125) = 11.32, p < .001$; Unable to Shop, $F(4, 124) = 10.54, p < .001$; Unable to Cook, $F(4, 123) = 10.23, p < .001$; and Unable to Do Simple Tasks, $F(4, 124) = 13.27$, 
Figure 1. Mean QWB score for patients who have problems never, sometimes, or everyday in losing things (upper left) asking repeatedly (upper right) forgetting day (middle left) shopping (middle right), cooking (lower left), and doing simple things (lower right).

$p < .001$. The correlations obtained were highly significant: Lower QWB scores were associated with poorer reported functioning (see Figure 1). Follow-up analyses indicated that in most cases, the significant results obtained were a reflection of large differences in QWB
between items scored as *never* or *not at all* and items indicating that the poorer functioning was present to at least some degree.

QWB scores were also associated with self-reported psychiatric distress: The relationship between the QWB and the BSI, although weaker than the relationship between the QWB and the measures of cognitive and behavioral functioning, was nevertheless still statistically significant ($r = -.26, p < .01$).

Caregivers of patients with lower QWB scores received a greater amount of respite time, $F(4, 123) = 25.26, p < .01$, and needed it more often, $F(4, 108) = 13.42, p < .01$, than did caregivers of patients with higher QWB scores (see Figure 2).

It might be argued that the analyses were biased because some nonaffected patients were included. Contrasts between patients and controls provide estimates of the impact of disease in comparison to
people roughly matched by age, sex, socioeconomic status, and living conditions. For a more stringent test, we compared QWB scores against MBPC scores with analysis limited to those carrying a diagnosis of Alzheimer's disease. For this analysis, there were significant linear contrasts for several MBPC scales, including Unable to Dress, \( F(1, 95) = 4.65, p < .03; \) Unable to Feed, \( F(1, 95) = 3.87, p < .05; \) Unable to Bathe, \( F(1, 95) = 3.63, p = .05; \) Unable to Shave, \( F(1, 95) = 6.16, p < .01; \) Incontinent, \( F(1, 95) = 9.23, p < .01; \) and Unable to Cook, \( F(1, 95) = 6.79, p < .01. \)

**Discussion**

In this article, we present evidence for the construct validity of a general QWB. Using data from the Alzheimer's Caregivers Research Project and the UCSD ADRC, we suggest that the general QWB is significantly associated with measures of dementia, memory and behavior problems, psychiatric symptoms, and respite time. These variables were selected for study because of their presumed relationship with quality of life.

As a general measure of behavioral status, the QWB generates a single number for each patient. This number can be used in a variety of ways. Data from the QWB are often used in a methodology known as quality-adjusted survival analysis. Using this method, survival time is adjusted by health-related quality of life. For example, in traditional survival analysis, an individual is scored 1.0 if alive and 0.0 if dead. Adjusted survival analysis assigns wellness scores between 0.0 and 1.0 based on health-related quality of life. In this study, people with Alzheimer's disease obtained scores of about .50. In quality-adjusted survival analysis, each year of survival in this state is counted as 0.50 QALYs.

A mathematical model integrates components of the model to express outcomes in a common measurement unit. Using information on current functioning and duration, it is possible to express the health outcomes in terms of QALYs. The model for point in time quality of well-being is
QWB = 1 - (observed morbidity × morbidity weight) 
- (observed physical activity × physical activity weight) 
- (observed social activity × social activity weight) 
- (observed symptom/problem × symptom/problem weight)

The net cost/utility ratio is defined as

\[
\frac{\text{net cost}}{\text{net QWB} \times \text{duration in years}} = \frac{\text{cost of treatment} - \text{cost of alternative}}{\text{[QWB}_{T} - \text{QWB}_{C}] \times \text{duration in years}}
\]

where \(\text{QWB}_{T}\) and \(\text{QWB}_{C}\) are measures of quality of well-being, adjusted for baseline values, for treatment and control groups respectively.

Consider, for example, an Alzheimer's disease patient who is in an objective state of functioning that is rated by community peers as 0.5 on a 0 to 1.0 scale. If he or she remained in that state for an entire year, the individual would lose the equivalent of one half year of life. Now suppose that a new drug can improve cognitive functioning slightly and would bring QWB scores from .50 to .60. Over the course of 1 year, the patient would gain .10 QALYs. If the benefit was maintained over 10 years, 1 QALY would be gained. Or, for each 10 patients who received a .10 benefit for 1 year, a QALY would be gained.

Because QWB scores are comparable across different patient groups, the QWB is valuable as a general outcome measure as well. For example, medical interventions in diverse disease states can be directly compared in cost/utility analyses. Such direct comparisons are difficult at best when only disease-specific measures are employed. By assessing outcomes with a general measure such as the QWB, it is possible to directly compare the public health impact of treatments for Alzheimer's disease with the impact of any other medical treatment. The QWB may therefore be especially useful in clinical trial assessment and estimates of a procedure's impact on a disease group. Much research is being conducted on treatments for Alzheimer's disease. Using the QWB in this way could help inform public policy by providing useful quality of life data on the impact of a given treatment in preventing or delaying the onset, or reducing the symptoms of, Alzheimer's disease.
This study is necessarily limited in several respects. One problem is the reliance on data obtained from a sample of convenience at only one research center. Because of the selection criteria, ADRC patients represent a more homogenous group than is found in most studies on Alzheimer's disease. Few other studies have examined only Alzheimer's disease patients who are living with a caregiver spouse. By limiting the study to spouses who had been married for, on average, almost 50 years, it was hoped that the quality of the proxy data would be high. This sample homogeneity, although a strength, nevertheless limits generalizability to Alzheimer's disease patients with spousal caregivers. In addition, ADRC subjects are not a random sample of all similar patients. Therefore, no infallible method of statistical control exists with which to ensure complete generalizability.

Another issue is the use of proxy respondents. Because patients in this study were cognitively impaired, proxy reports were used. The QWB has a separate proxy questionnaire that obtains information about those unable to report for themselves. The proxy questionnaire was validated in the initial QWB validation study (Kaplan et al., 1976). Validation of proxy responding among older cognitively impaired adults is particularly difficult, because there is no clear standard against which to validate proxy reports (see Rubinstein et al., 1989). Rothman et al. (1989) found that proxy reports for the Sickness Impact Profile (SIP) corresponded relatively well with patient reports. However, the correspondence was better for physical health than for psychosocial problems. Magaziner, Hebel, and Warren (1987) found good correspondence between reports by patients and proxies for questions about specific diseases. However, correspondence for reports of symptoms was lower. In this study, it is worth noting that the items for which there was greatest discrimination on the QWB were those for which proxy respondents would have the greatest opportunity to directly observe. These included inability to perform several basic activities of daily living.

Stewart, Sherbourne, and Brod (1996) offered an excellent review of methodological issues in the assessment of quality of life for older and demented patients. They noted that it is widely assumed that self-reports are not valid for patients with dementia because they are
cognitively impaired and cannot report on their own experiences. Thus, use of proxy respondents is often necessary. The review cites several papers showing that patients with early stage dementia can provide some reliable self-reports. Studies addressing these issues have produced fairly consistent results. Proxies tend to overestimate patient disability for instrumental activities but tend to give reports consistent with patients for self-care activities (Magaziner, Simonsick, Kashner, & Hebel, 1988). Several studies, reviewed by Stewart and colleagues (1996), show that proxies may estimate disability to be greater than when reports are based on patient descriptions. An evaluation of proxy reports for patients with Alzheimer’s disease using the SIP suggested that responses are most valid when given by proxies who are family members (Krenz, Larson, Buchner, & Canfield, 1988).

In summary, the issue of proxy responding is very difficult. Health-related quality of life measures depend on patient reports. Those patients with serious dementia are often unable to provide these reports, and proxies may be the only source of information. Validity studies are not possible because the validity criterion is typically patient reports. When the patients themselves are known to be unreliable, the assessment of validity becomes very difficult. One of the interesting findings in this study was that the QWB corresponded to proxy reports even when the analysis was restricted to patients who carried the diagnosis of Alzheimer’s disease. This is a stringent test of validity because the range of variability within Alzheimer’s disease patients was restricted. Nevertheless, there was a systematic relationship between QWB scores and proxy reports for several memory and behavior problems. The evaluation of proxy reporting will continue to be a theoretical and methodological challenge. We hope future work addresses these problems.

The QWB requires reports of both physical functioning and symptoms. DeBon, Pace, Kozin, and Kaplan (1995) demonstrated that there is reasonable validity for proxy reports of physical functioning in the elderly. This confirms reports relevant to other instruments (Magaziner et al., 1987; Rothman, Hedrick, & Inui, 1989). The difficulty has usually been in the evaluation of symptoms. The QWB scores only the symptom or problem with the lowest weight. For patients with Alzheimer’s disease, we would assume all have the symptom, “trouble thinking, learning, or remembering.” This symptom has one of the
lowest weights in the QWB system and, therefore, will almost always be the one used in scoring. In other words, the fact that patients and proxies may have disagreed on other symptoms or problems is moot because the only symptom used in scoring was assumed to be present in all Alzheimer’s disease patients.

We have chosen to test the validity of the QWB against several widely used dementia measures that quantify a broad range of difficulties associated with Alzheimer’s disease. No attempt was made to compare the QWB with all, or even a majority of possible measures. Also, we are not arguing that the QWB is a suitable replacement for disease-specific measures such as the MBPC. These measures make available necessary and useful diagnostic information; such information cannot be provided by the QWB. On the other hand, we do believe that the concept of a QALY and measures such as the QWB may be valuable for outcomes research in Alzheimer’s disease.

This study suggests several directions for future research. First, it will be important to replicate these results on a larger and more representative sample of Alzheimer’s disease patients. Another direction will be to use the QWB to evaluate the impact of both behavioral and pharmacological interventions for the treatment of Alzheimer’s disease. The cost/effectiveness of these interventions can be estimated using these methods. Longitudinal studies that track changes in quality of life over time are also needed. Finally, methodological studies evaluating issues, such as the validity of proxy reports, are also needed.

REFERENCES


Derogatis, L. R., Lipman, R. S., Rickels, K., Uhlenhuth, E. H., & Kovi, L. (1974). The Hopkins

Derogatis, L. R., Yevzeroff, H., & Wittelsberger, B. (1975). Social class, psychological disorder,


Kaplan, R. M., Anderson, J. P., Patterson, T. L., McCutchan, J. A., Weinrich, J. D., Heaton, R. H.,
Scale for persons with HIV infection. *Psychosomatic Medicine, 57*, 138-157.

The Quality of Well-being Scale: Applications in AIDS, cystic fibrosis, and arthritis. *Medical Care, 27*, S27-S43.


(including discussion of a general health-decision model). *Quality of Life and Cardiovascular Care, 4*, 131-139.


