Validity of the Quality of Well-Being Scale for Persons With Human Immunodeficiency Virus Infection

ROBERT M. KAPLAN, PHD, JOHN P. ANDERSON, PHD, THOMAS L. PATTERSON, PHD, J. ALLEN MCCUTCHAN, MD, JAMES D. WEINRICH, PHD, ROBERT K. HEATON, PHD, J. HAMPTON ATKINSON, MD, LEON THAL, MD, JAMES CHANDLER, MD, IGOR GRANT, MD, AND THE HNRC GROUP¹

To evaluate the validity of the Quality of Well-Being Scale (QWB) for studies of patients with human immunodeficiency virus (HIV) disease, 514 men were studied who were divided among four categories: Centers for Disease Control and Prevention (CDC) Group A (N = 272), CDC-B (N = 81), CDC-C (N = 47), and uninfected male controls ($N \approx 114$). The QWB and a variety of medical, neuropsychological, and biochemical measures were administered to all participants. When QWB scores were broken down by HIV grouping, the CDC-C group was significantly lower (.614) than the CDC-B (.679), CDC-A (.754), or control group (.801). The difference between Groups CDC-C and CDC-A was about .14 units of well-being, which suggests that individuals lose ¹/₇ equivalents of 1 well year of life for each year they are in Group CDC-C in comparison to the asymptomatic group (Group CDC-A). In comparison to the controls, this would equal a 1-year of life loss for each seven infected individuals. The QWB was shown to be significantly associated with CD4+ lymphocytes (p < .001), clinician ratings of neuropsychological impairment (p < .04), neurologists ratings of dysfunction (p < .001), and all subscales of the Profile of Mood States. Baseline QWB scores were significant prospective predictors of death over a median follow-up time of 30 months. Multivariate models demonstrated high covariation between predictors of QWB. It was concluded that the QWB is a significant correlate of biological, neuropsychological, neurological, psychiatric, and mortality outcomes for male HIV-infected patients.

Key words: HIV infection, quality of life, validity, outcomes research.

INTRODUCTION

Acquired immunodeficiency syndrome (AIDS), which results from infection with the human immunodeficiency virus (HIV), represents one of the most important threats to the world population in the 1990s. The World Health Organization estimates that more than 13 million people had been infected with HIV by 1993, and projections suggest that the pandemic continues to be out of control (1).

Beyond the numerous complications of AIDS, HIV infection may be associated with a broad range of milder conditions, including asymptomatic lymphadenopathy, thrombocytopenia, minor infections, wasting, fever, diarrhea, and a variety neurological disorders. The impact of HIV infection and AIDS on functioning is equally diverse. For example, disability may range from none to severe fatigue, dementia, blindness, or paraplegia. Available treatment for HIV infection appears to reduce mortality and overall morbidity rates temporarily (2) but contributes addi-

From the Naval Hospital, San Diego, and University of California, San Diego, La Jolla, California.

Address reprint requests to: Robert M. Kaplan, Division of Health Care Sciences 0622, Department of Family and Preventive Medicine, University of California, San Diego, La Jolla, CA 92093– 0622.

Received for publication October 18, 1993; revision received July 18,1994.

¹ The San Diego HIV Neurobehavioral Research Center (HNRC) group is affiliated with the University of California, San Diego, the Naval Hospital, San Diego, and the San Diego VA Medical Center, and includes: Igor Grant, MD, Director; J. Hampton Atkinson, MD, Co-Director; Robert A. Velin, PhD, Center Manager; Edward C. Oldfield III, MD, James L. Chandler, MD, Mark R. Wallace, MD, and Joseph Malone, MD, Co-Investigators Naval Hospital San Diego; J. Allen McCutchan, MD, Principal Investigator Medical Core; Stephen A. Spector, MD, Principal Investigator Virology Core; Leon Thal, MD, Principal Investigator Neurology Core; Robert K. Heaton, PhD, Principal Investigator Neuropsychology Core; John Hesselink, MD, and Terry Jernigan, PhD, Co-Principal Investigators Imaging Core; J. Hampton Atkinson, MD, Principal Investigator Psychiatry Core; Clayton A. Wiley, MD, PhD, Principal Investigator Neuropathology Core; Richard Olshen, PhD, and Ian Abramson, PhD, Co-Principal Investigators Biostatistics Core; Nelson Butters, PhD, Principal Investigator Memory Project; Renée Dupont, MD, Principal Investigator SPECT Project; Thomas Patterson, PhD, Principal Investigator Life Events Project: Sidney Zisook, MD, Principal Investigator Mood Project; Dilip Jeste, MD, Principal Investigator Psychosis Project; Hans Sieburg, PhD, Principal Investigator Dynamical Systems Project; and James D. Weinrich, PhD, Principal Investigator Sexology Project.

tional morbidity in some patients. The diverse impacts of both HIV disease and its treatment require a general approach to program and outcome assessment. Previous attempts to evaluate quality of life in HIV-infected patients have focused on psychological outcomes. Studies that apply general health-related quality-of-life scales in HIV have recently been reviewed (3–7). However, only a few of the studies have offered comprehensive evidence for validity (5–7).

There are several different approaches to the assessment of health-related quality of life. The most common methods develop psychometric profiles that represent health outcome along several defined dimensions (5–7). Decision theory offers a different approach to the measurement of health outcomes. Based on the argument that health outcomes are very complex, many sources of information must be integrated into health evaluations. For example, we must decide whether the patient is getting better or worse, whether the treatment is useful or not useful, or which among many alternative programs to support. These decisions require the placement of complex arrays of information onto simpler scales that can be used in the evaluation process. For instance, some medications produce benefits but also cause side effects. The side effects are typically measured in units that are unrelated to the benefits. The decision approach attempts to put this information together so that a single evaluative judgment can be reached.

THE QUALITY OF WELL-BEING SCALE

Within the last few years, there has been growing interest in using quality-of-life data to help evaluate the cost-utility or cost-effectiveness of health care programs. Cost studies have gained in popularity because health care costs have grown rapidly. Not all health care interventions are equally efficient in returning benefit for the expended dollar. Objective cost studies might guide policy makers toward an optimal and equitable distribution of scarce resources. In cost-effectiveness analysis, the benefits of a health care intervention are typically expressed in terms of some measure of health status. For example, an analyst might consider the cost to achieve an increase of 100 in the number of CD4+ cells. Costutility is a special use of cost-effectiveness that weights observable health states by preferences or utility judgments of quality (8). In cost-utility analysis, the benefits of medical care, behavioral interventions, or preventive programs are expressed in terms of well years produced. These outcomes have

also been described as quality-adjusted life years (QALYs) (9), discounted life years (10), or healthy years of life (11). Because the term QALYs has become most popular, we use it in this presentation.

QALYs integrate mortality and morbidity rates to express health status in terms of equivalents of well years of life. For example, if a man dies of AIDSassociated lymphoma at age 40 and we would have expected him to live to age 75, it might be concluded that the disease was associated with 35 lost years of life. If 100 men died at age 40 (and also had a life expectancy of 75 years), we might conclude that 3500 (100 men \times 35 years) life years had been lost. Yet, death is not the only outcome of concern in HIV disease. Many adults have AIDS-associated conditions, which leave them somewhat disabled over long periods of time. Although they are still alive. the quality of their lives has diminished. QALYs take into consideration the quality-of-life consequences of these illnesses. For example, a disease that reduces quality of life by one half will take away 0.5 QALYs over the course of each year. If it affects two people, it will take away 1.0 years (equal to 2×0.5) over each year period. A medical treatment that improves quality of life by 0.2 for each of five individuals will result in the equivalent of one QALY if the benefit is maintained over a 1-year period. This system has the advantage of considering both benefits and side effects of programs in terms of the common QALY units. One of the major issues in many studies is the assessment of the considerable toxic effects of treatments for conditions like AIDSassociated lymphoma, which generally offer only brief remissions in patients with AIDS. The general measurement system can also quantify the side effects. Furthermore, it can be used to evaluate the relative importance of side effects so that a net assessment of the treatment can subtract side effects from benefits. In addition, the cost of treatment can also be estimated, and calculations of the cost to produce a QALY can be calculated.

Although there are several different approaches for obtaining QALYs, most of them are similar (12). The approach that our group prefers involves several steps. First, patients are classified according to observable levels of functioning. These levels are represented by scales of Mobility, Physical Activity, and Social Activity. The dimensions and steps for these levels of functioning are shown in Table 1. The reader is cautioned that these steps are not actually the scale, only listings of labels representing the scale steps. Standardized questionnaires have been developed to classify individuals into one of each of these scale steps (13, 14). In addition to classification

Table 1.	Quality of Well-being/General Health Policy Model: Elements and Calculating Formulas (Function Scales, with Step
	Definitions and Calculating Weights)

Step No.	Step Definition	Weight
	Mobility Scale (MOB)	
5	No limitations for health reasons	-0.000
4	Did not drive a car, health related; did not ride in a car as usual for age (younger than 15 yr), health related, and/or did not use public transportation, health related; or had or would have used more help than usual for age to use public transportation, health related	-0.062
2	In hospital, health related	-0.090
	Physical Activity Scale (PAC)	
4	No limitations for health reasons	-0.000
3	In wheelchair, moved or controlled movement of wheelchair without help from someone else; or had trouble or did not try to lift, stoop, bend over, or use stairs or inclines, health related; and/or limped, used a cane, crutches, or walker, health related; and/or had any other physical limitation in walking, or did not try to walk as far as or as fast as other the same age are able, health related	-0.060
1	In wheelchair, did not move or control the movement of wheelchair without help from someone else, or in bed, chair, or couch for most or all of the day, health related	-0.077
	Social Activity Scale (SAC)	
5	No limitations for health reasons	-0.000
4	Limited in other (e.g., recreational) role activity, health related	-0.061
3	Limited in major (primary) role activity, health related	-0.061
2	Performed no major role activity, health related, but did perform self-care activities	-0.061
1	Performed no major role activity, health related, and did not perform or had more help than usual in performance of one or more self-care activities, health related	-0.106
	Calculating Formulas	
	Formula 1. Point-in-time well-being score for an individual (W): W = 1 + (CPXwt) + (MOBwt) + (PACwt + (SACwt))	
	where "wt" is the preference-weighted measure for each factor and CPX is Symptom/Problem complex. For example, the W score for a person with the following description profile may be calculated for 1 day as:	
CPX-11	Cough, wheezing or shortness of breath, with or without fever, chills, or aching all over	-0.257
MOB-5	No limitations	-0.000
PAC-1	In bed, chair, or couch for most or all of the day, health related	-0.077
SAC-2	Performed no major role activity, health related, but did perform self-care	-0.061
	W = 1 (-0.257) + (-0.000 + (-0.077) + (-0.061) = 0.605	
	Formula 2. Well years (WY) as an output measure:	
	$WY = [No. of persons \times (CPXwt + MOBwt + PACwt + SACwt) \times Time]$	

into these observable levels of function, individuals are also classified by the one symptom or problem that was most undesirable (Table 2). Nearly 80% of the population reports at least one symptom during a 6-day interval. Symptoms may be severe, such as serious chest pain, or minor, such as the inconvenience of taking medication or a prescribed diet for health reasons. The rationale for choosing only one symptom/problem is that the number of symptoms is not necessarily related to the degree of dysfunction. A person with four minor symptoms (a sore throat, runny nose, itchy eyes, and mild fatigue) may still go to work and show high functioning. Conversely, a person with a single symptom of severe fatigue may remain in bed, stay away from work, and need assistance with self-care. The Quality of Well-Being Scale (QWB) identifies the symptom that is most undesirable and grades it by the degree to which it affects everyday activities. Fatigue, for example, may have little effect, or it may have a profound effect, depending on how it affects the functional scales.

The most undesirable symptom or problem typically produces the greatest variation from wellness (1.0) and is inclusive of less severe symptoms (10). The functional classification (Table 1) and the accompanying list of symptoms or problems (Table 2) was created after extensive reviews of the medical and public health literature (10). Over the last decade, the methods for classification of function and symptoms were repeatedly condensed until we arrived at the current versions. Various methodological studies on the questionnaire have been conducted (13, 14). With structured questionnaires, an interviewer can generally obtain classifications on these four dimensions in 11 to 16 minutes. With a newly available form, this time has been reduced to about 7 minutes.

Once observable-behavioral levels of functioning have been classified, a second step is required to place each individual on the 0 to 1.0 scale of wellness. To accomplish this, the observable health states are weighted by "quality" ratings for the desirability of these conditions. Human value stud-

CPX No. Weights	CPX Description	
1	Death (not on respondent's card)	-0.727
2	Loss of consciousness such as seizure (fits), fainting, or coma (out cold or knocked out)	-0.407
3	Burn over large areas of face, body, arms, or legs	-0.387
4	Pain, bleeding, itching, or discharge (drainage) from sexual organs—does not include normal menstrual (monthly) bleeding	-0.349
5	Trouble learning, remembering, or thinking clearly	-0.340
6	Any combination of one or more hands, feet, arms, or legs either missing, deformed (crooked), paralyzed (unable to move), or broken—includes wearing artificial limbs or braces	-0.333
7	Pain, stiffness, weakness, numbness, or other discomfort in chest, stomach (including hernia or rupture), side, neck, back, hips, or any joints or hands, feet, arms, or legs	-0.299
8	Pain, burning, bleeding, itching, or other difficulty with rectum, bowel movements, or urination (passing water)	0.292
9	Sick or upset stomach, vomiting or loose bowel movement, with or without chills, or aching all over	-0.290
10	General tiredness, weakness, or weight loss	-0.259
11	Cough, wheezing, or shortness of breath, with or without fever, chills, or aching all over	-0.257
12	Spells of feeling upset, being depressed, or of crying	-0.257
13	Headache, or dizziness, or ringing in ears, or spells of feeling hot, nervous, or shaky	-0.244
14	Burning or itching rash on large areas of face, body, arms, or legs	-0.240
15	Trouble talking, such as lisp, stuttering, hoarseness, or being unable to speak	-0.237
16	Pain or discomfort in one or both eyes (such as burning or itching) or any trouble seeing after correction	-0.230
17	Overweight for age and height or skin defect of face, body, arms, or legs, such as scars, pimples, warts, bruises, or changes in color	-0.188
18	Pain in ear, tooth, jaw throat, lips, or tongue; several missing or crooked permanent teeth—includes wear- ing bridges or false teeth; stuffy, runny nose; or any trouble hearing—includes wearing a hearing aid	-0.170
19	Taking medication or staying on a prescribed diet for health reasons	-0.144
20	Wore eyeglasses or contact lenses	-0.101
21	Breathing smog or unpleasant air	-0.101
22	No symptoms or problem (not on respondent's card)	-0.000
23	Standard symptom/problem	-0.257
X24	Trouble sleeping	-0.257
X25	Intoxication	-0.257
X26	Problems with sexual interest or performance	-0.257
X27	Excessive worry or anxiety	-0.257

Table 2. Quality of Well-being/General Health Policy Model: Symptom/Problem Complexes (CPX) with Calculating Weights

Note: -0.727 for death becomes 0 when adjustments for mobility, physical activity, and social activity for death are included. Symptoms marked by X (X24 to X27) have not been scaled at this time. The weights for these symptoms represent an average case.

ies have been conducted to place the observable states onto a preference continuum with an anchor of 0 for death and 1.0 for completely well. In several studies, random samples of citizens from a metropolitan community evaluated the desirability of more than 400 case descriptions. With these ratings, a preference structure that assigned the weights to each combination of an observable state and a symptom/problem has been developed (10). Cross-validation studies have shown that the model can be used to assign weights to other states of functioning with a high degree of accuracy $(R^2 = .96)$ (15). The regression weights obtained in these studies are given in Tables 1 and 2. Studies have shown that the weights are highly stable over a 1-year period and that they are consistent across diverse groups of raters (13). Finally, it is necessary to consider the duration that individuals remain in various health states. For example, 1 year in a state that has been assigned the weight of 0.5, is equivalent to 0.5 of a

OALY. Table 1 provides an illustrative example of a calculation of a QALY outcome for an HIV-infected patient. On the day he was assessed, he coughed, wheezed, or was short of breath. He had no limitations in mobility because he drove his car to the clinic. However, he was in a bed or chair most of the day and performed no major social role activity. The preference weights associated with the observable state suggest that peers evaluate the state to be about 0.6 on a 0 to 1.0 scale. If patients remain in this state for an entire year, they lose 0.4 well years. If this situation were maintained over the course of a decade, the person would lose the equivalent of 4 well years of life. Both reliability (12) and validity studies have been previously published. These studies show that the QWB has internal consistency, is repeatable, has stable preference weights, and correlates with a wide variety of medical and psychosocial variables (10, 16).

The ultimate goal of QWB assessment in HIV

disease is to develop a single number that represents the current impact of disease. Furthermore, this number represents the combined effects of morbidity and mortality on HIV-infected patients when integrated over time. Ultimately, this will allow an expression of the impact of HIV disease on of the number of well years of life lost to this disease. However, application of this model in HIV disease requires specific validation for this patient group. Evidence for the validity of the QWB for studies of HIV-infected patients is offered in this article.

METHODS

HIV Neurobehavioral Research Center (HNRC) of the University of California, San Diego

The subjects were 400 HIV-positive men and 114 HIV-negative uninfected men who serve as controls. These subjects were assessed in a standardized protocol developed by the HNRC, a collaborative investigation of the University of California, San Diego, Naval Hospital in San Diego, and the San Diego Veterans Affairs Medical Center. The demographic characteristics of the participants are summarized in Table 3. The controls were slightly better educated, and there were fewer African Americans in the control group. No member of the control group had an HIV-related diagnosis. Among the HIV-positive group, the frequencies of medical diagnoses were herpes zoster (N = 30), Pneumocystis pneumonia (N = 23), Mycobacterium avium intracellulare (N = 16), cyptococcal meningitis (N = 3), and lymphoma (N = 3).

The protocol for evaluating patients in the HNRC is very extensive and cannot be described in detail here. However, we provide a brief overview of the various cores from which data were utilized in the present analysis. The complete protocol is available (17).

Medical core. The Medical Core provides complete medical,

drug, and sexual histories studies of immunologic status, including levels of CD4+ cells, serum β_2 -microglobulin, and P²⁴ antigens; physical examination; and classification of state of HIV infection with the Centers for Disease Control and Prevention (CDC) system, as revised in 1993. The medical core also administers the QWB questionnaire and the Karnofsky performance status measure.

Neurology core. This core provides a detailed neurological examination, which assesses sensation, motor function, reflexes, alertness, and concentration. Quantitative sensory testing is performed by evaluating responses to vibration changes. Evokedpotential testing evaluates speed and strength of nerve conduction in multiple pathways in the brain, spinal cord, and peripheral nerves. These include brain stem auditory-evoked potentials, visual-evoked potentials, somatosensory-evoked potentials, and so forth. Nerve conduction velocity studies assess functioning of peripheral nerves. The information from the neurological examinations was grouped into summary clinical ratings for central nervous system and peripheral nervous system functioning. The values ranged from 1 for unimpaired to 5 for high levels of impairment (18).

Neuropsychological core. A neuropsychological evaluation involves about 6 to 10 hours of standardized tests designed to evaluate the functioning of the central nervous system. These tests include subevaluations for memory, problem solving, concentration, language, sensory, and motor skills. The tests include the Wechsler Adult Intelligence Scale-Revised, the expanded Halstead-Reitan Battery, the Paced Serial Addition Test, and the Multilingual Aphasia Examination (see Ref. 17 for details). On the basis of these tests, a neuropsychologist provides a blinded summary rating on a 9-point scale. Scores below 5 are considered unimpaired, 5 is considered mild impairment, and 6 and above is termed greater than mild impairment (17).

Magnetic resonance imaging (MRI) core. This core detects structural changes in the brain with qualitative and quantitative imaging techniques. The method uses pixel images in two-dimensional planes and voxel estimates of volumes in three-dimensional planes. Computer image analysis techniques are used to establish landmark points and structural boundaries, and images are compared relative to a common anatomical coordinate system. A simple summary score gives the number of parenchymal abnormalities. The method was described by Jernigan et al. (19).

Table 3. Summary of Demographic Characteristics for HIV-Positive Participants and Controls

	Controls ($N = 114$)		HIV+(N = 400)		- 1/-1
	Mean	(SD)	Mean	(SD)	<i>p</i> value
SES (5 point scale, 1 lowest, 5 highest)	3.7	(1.1)	3.8	(1.0)	<i>p</i> < .6
Education (years)	14.8	(2.1)	14.1	(2.1)	p < .01
Income (in US dollars)	22,560	(20,785)	22,049	(16,047)	p < .9
Ethnicity	N	(%)	N	(%)	
American Indian	2	(1.8)	1	(0.3)	
African American	4	(3.5)	49	(12.3)	
Cuban			2	(0.5)	
 Filipino 			5	(1.3)	
Hispanic	6	(5.3)	21	(5.3)	
Latin American			1	(0.3)	
Mexican American			8	(2.0)	
Asian	2	(1.8)	1	(0.3)	
Pacific Islander		,	5	(1.3)	
Puerto Rican	2	(1.8)	1	(0.3)	
White	98	(86.0)	306	(76.5)	<i>p</i> < .01

Psychiatry core. A psychiatric evaluation assesses emotional well-being, both historically and currently, with a variety of standardized tests including the Profile of Mood States (POMS) (20) and the Beck Depression Inventory (BDI) (21). The POMS is a questionnaire measure of six emotional states: tension/anxiety, anger/hostility, depression/dejection, vigor/activity, fatigue/inertia, and confusion/bewilderment. The POMS was chosen because the items that define these states are based on emotional rather than somatic correlates of mood. The BDI is a standard self-report measure of somatic and nonsomatic symptoms of depression and has established validity in medically ill populations (22).

Analysis

This article presents data from 514 patients who completed the QWB questionnaire at their initial visit. Representative data from each of the cores were chosen for comparison. Specifically, we evaluated the validity of the QWB in correspondence with indicators of medical progress of the disease (CD4+ lymphocyte and β_2 -microglobulin levels), neuropsychological performance, clinical evaluation of neurological functioning, brain imaging, and psychiatric variables selected from the POMS.

RESULTS

Medical Core

Participants were categorized into four groups according to their CDC classification. CDC Group A

were infected but asymptomatic, Group B had minor conditions, and Group C had major conditions. The fourth group was uninfected controls with the similar occupations and sexual orientation. The CDC-A group scored 0.754 (standard error [SE] = 0.010) on the QWB scale; the combined CDC-B group scored 0.678 (SE = 0.006). The CDC-C group scored 0.614 (SE = 0.013). The uninfected male control group scored 0.801 (SE = 0.010; Figure 1, A). These differences were highly significant (F(3/507) = 53.82, p < .00001) with a very strong linear component (p < .00001). Pairwise comparisons using the Tukey-Kramer Honest Significant Difference (HSD) test showed each group to be significantly different from all other groups.

CD4+ lymphocytes were used as the primary indicator of severity of HIV-induced immunosuppression. Figure 1, *B* shows the relationship between CD4 lymphocytes and QWB scores. The CD4 counts were broken into three categories: 0 to 199, 200 to 499, and 500 or more cells/m³. Again, this linear relationship was statistically significant (F(1/487) =105.00, p < .0001). Curve-fitting procedures were used to evaluate the relationship. Further analysis used CD4 cells as a continuous variable. These



Fig. 1. Relationship between mean QWB and (A) CDC Group, (B) CD4+ lymphocyte count, (C) β_2 -microglobulin quartile, (D) neurologists' integrative clinical ratings for central dysfunction, (E) neuropsychological test summary scores of unimpaired, moderately impaired, and impaired; and (F) eventual death. Error bars show the 95% confidence intervals.

analyses demonstrated that the steepest portion of the relationship was for patients with a CD4+ cell count less than 500 cells/m³. This confirms clinical observations that most HIV-attributable morbidity occurs in patients with less than 500 CD4+ cells/m³.

In addition, we studied the relationship between QWB score and β_2 -microglobulin. This variable was broken into quartiles, and the relationship with QWB is given in Figure 1, *C*. There was a significant relationship between variables (F(3/498) = 6.05, p < .0005) with a strong linear component (F(1/498) = 17.62, p < .0001). Finally, there was a strong correlation between the QWB and the Karnofsky measure (r = .51, p < .001). The correlations with biological outcomes were consistently weaker for the Karnofsky than for the QWB measure.

Neurology Core

The HNRC has a wide variety of neurological indicators. For the purposes of this article, we selected a general indicator of neurological function characterized by neurologists' integrative clinical ratings for central and peripheral dysfunction. This rating is given on a 5-point scale with higher scores associated with lower neurological functioning. Figure 1, D shows a strong linear relationship between neurologist-determined clinical ratings and QWB scores (F(1/477) = 41.34, p < .0001). The wide confidence interval around Category 5 results because there are few cases.

Neuropsychology Core

Figure 1, *E* summarize some of the neuropsychological associations. On the basis of multiple neuropsychological tests, a summary score was created to indicate whether the patient was impaired (greater than 6 on a 9-point rating scale), moderately impaired (score of 5), or unimpaired (score of 4 or less). Two categories were selected for analysis: global rating and rating of motor function. In each case there was a strong linear relationship with QWB (global F(1/477) = 6.19, p < .01; motor F(1/477) = 13.92, p < .0002).

MRI Core

MRI images were used to determine total parenchymal abnormalities. Those with any evidence of parenchymal abnormalities had significantly lower QWB scores (mean = 0.663, SE = 0.11, N = 17) than

144

those with no evidence of abnormality (mean = 0.738, SE = 0.12, N = 446, F(1/462) = 5.97, p < .02).

Psychiatry Core

Relationships between the QWB and the variables from the POMS and with the BDI are shown in Table 4. In addition, the table shows correlations between the variables from the Psychiatry Core and the Karnofsky performance status measure. The QWB was significantly correlated with all POMS subscales and with the BDI. Parallel correlations with the Karnofsky measure were weaker in each comparison (see Table 4).

Work Status

A randomly selected subgroup of participants was classified according to their current work status. There was a significant relationship between QWB status for those currently working (0.754), those who had changed work because of their illness (0.720), and those who were now work disabled (0.692). The difference between those working and disabled was statistically significant (p < .05).

Prospective Prediction of Mortality

Figure 1, *F* shows the initial QWB values for those HIV-infected patients who died by the time of this report (mean follow-up time approximately 30 months, N = 46) and those who survived (N = 466). As the Figure suggests, the QWB is a significant prospective predictor of mortality ($t_{510} = 4.12$, p < .0001).

Multivariate Analysis

Validity studies attempt to characterize individual correlations between a criterion measure and various

Table 4.	Correlations Between QWB, Karnofsky, and
	Psychological Measures

Scale	QWB	Karnofsky
POMS anger/hostility	16**	08
POMS confusion/bewilderment	35**	17**
POMS depression/dejection	35**	03
POMS fatigue/inertia	34**	20**
POMS tension/anxiety	31**	11*
POMS vigor/activity	.32**	.14*
Beck Depression Inventory	49**	37**

* p < .05.

indicators believed to be associated with the outcome. However, clinical characteristics of a disease process are often highly intercorrelated. Thus, it is important to study the multivariate relationships between predictors and QWB as well as the univariate associations. Table 5 summarizes the multivariate relationships between QWB and measures of medical, neurological, neuropsychological, and psychiatric variables. One representative variable was selected from each of these domains. The variables were CD4+ cells (medical), BDI (psychiatric), Global Neuropsychological Score (neuropsychology), parenchymal abnormality (MRI), and neurologist rating of central dysfunction (neurology) Variables were entered into the equation in a single block. The multivariate analysis suggested that 28% of the variance in QWB scores could be accounted for by five predictor variable, with the most significant contributors to the multivariate equation being the BDI score (t = -8.85, p < .001), percent CD4+ cells (t =4.86, p < .001), and neurologist rating (t = -3.14, p < .002). Neuropsychological scores were also a significant contributor to the model (t = 2.32, p = .021). The MRI variable did not add significant information beyond the other measures (p = .33). The analysis was repeated controlling for antiretroviral therapy and socioeconomic status (SES). Although patients who were receiving therapy had lower QWB scores (0.02 units), controlling for therapy did not alter the significance of any of the other variables. Furthermore, controls for SES did not alter the pattern of results.

DISCUSSION

The term *validity* is defined as agreement between a test score and the quality it is believed to measure. In health measurement, obtaining evidence for validity is analogous to gathering support for a scientific theory. Some attributes, such as quality of life, have no clear definition. These hypothetical constructs must be validated through a series of activities. In this article, we present evidence for the construct validity of a general QWB. We suggest that the general QWB is associated with medical, neurological, neuropsychological, brain imaging, and psychiatric measures. Variables were selected for study because of their presumed relationship with quality of life. Continuing studies will attempt to validate these outcomes in relation to a wide variety of the scores and measures, including radiographic images of the brain and other more specific indicators of functioning from each domain.

The data also suggest that general quality of life is a significant prospective predictor of death. One of the major issues in the evaluation of health outcome measures is their sensitivity to clinical change. We previously reported that there are significant differences between HIV-infected patients randomly assigned to receive zidovudine treatment or placebo (23). These findings were largely influenced by the mortality component in the QWB scale. However, Wu et al. (24) demonstrated that the OWB can also detect the effect of zidovudine treatment even with the mortality factor removed. Ongoing clinical trials are evaluating the effect of zidovudine treatment for patients with less advanced HIV infection, and the results from these trials should be reported in the near future.

The results from the study have both general and HIV-specific implications for the psychometric qualities of the QWB, for clinical trial research, and for public policy analysis. Each of these is discussed briefly subsequently.

HIV infection affects virtually every organ system. Although there are multiple clinical markers of HIV disease, it is useful to characterize "how sick" HIVinfected patients are by placing them along a continuum from death to optimal function. The QWB has the advantage of stating approximately where patients fall along this continuum. This information may be useful to characterize populations of patients, identify the comparability of different groups of patients, and evaluate the general impact of infection on activities of daily living.

Table 5. Summary of Regression of Various Variables on QWB^a

Term	В	Standard Error of B	β Value	t Ratio	p Value
Intercept	.731	.025		29.18	.00001
CD4 cells	.002	.0004	.222	4.86	.00001
Beck Depression	007	.0008	381	-8.51	.00001
Neuropsychologist rating	.009	.0042	.104	2.32	.021
Neurologist rating	025	.008	143	-3.14	.002

^a Multiple R = .53, F(4,388) = 37.09, p < .0001.

Implication for QWB Validity and Reliability

Research studies require evidence for validity and reliability. Previous reports have described the reliability of the preference weights in the QWB system (15), the reliability of questionnaire classification (13), and the reliability of day-to-day assessment (14). The validity of the QWB has also been described for the general population (10) and for specific patient populations, such as patients with chronic obstructive pulmonary disease (25), cystic fibrosis (26), arthritis (27, 28), and a variety of other conditions (16). This article provides preliminary evidence for the validity of the QWB system for male patients with HIV infection. In addition, the data suggest that the QWB is associated with a wide variety of outcomes, including psychological depression. However, we must caution that the BDI is a relatively poor measure of depression in medically ill patients because of the confounding of vegetative signs associated with depression and the physical symptoms of progressive HIV infection. The finding that the QWB is associated with Beck depression scores does not necessarily provide supportive evidence that the QWB taps psychological distress. Continuing studies will provide further validation of the QWB in other populations.

It is also important to emphasize that the QWB may not assess the same range of outcomes as other quality-of-life measures. For example, the QWB does not directly measure psychological distress or sexual functioning. The HIV Overview of Problems-Evaluation System measure, which directly evaluates these dimensions, does not show the same linear relationship with CD4+ cells or clinical classification as the QWB (7). This may suggest that the QWB is more closely related to the physical dimension of life quality. On the other hand, the QWB does show higher correlations with psychological measures than does the Karnofsky performance status measure, which is a common indicator of physical disability.

Another concern about the QWB is the use of preference weights to value health states. These weights are derived from members of the general population rather than from patient groups. Studies have consistently shown that patient weights do not differ from those obtained from the general population (29), and HIV-infected patients have been shown to assign weights comparable to those from the general population (30). However, these weights are controversial because they are subjective. We encourage continuing study of this issue.

Implication for Clinical Trials

Quality-of-life measures are becoming increasingly common in clinical trials for several reasons. First, general measures of functioning can demonstrate both benefits and side effects of treatment. Second, many clinical trials use mortality rate as the end point. Quality-of-life measures provide more sensitive measures for use in outcome studies. A third advantage of the QWB for clinical trials is that it allows a common unit of effectiveness for comparisons across trials. For example, a generalized system allows the comparison of treatment benefit in HIV infection with that obtained in other disease states. An earlier comparison between those treated with zidovudine and a group treated with placebo showed the difference between the means of the two groups is very large. The zidovudine group was about 0.6; the control group was about 0.13, a difference of 0.47 units of well-being over this 1-year interval.

Implications for Health Policy

Different models for measuring health outcomes may have different applications. For example, a General Health Policy Model can produce data for cost-utility analysis. By placement of the outcomes in a common unit, it is possible to estimate the return on the dollar for investing in various components of health care. This is especially important in the selection of therapy for chronic, ultimately fatal conditions such as AIDS, which presently absorb growing amounts of health resources.

In summary, the QWB is significantly correlated with biological, neuropsychological, neurological, psychiatric, and mortality outcomes for male HIVinfected patients. These data suggest that the QWB may be an appropriate general outcome measure in studies of patients with HIV disease.

The principal support for the HNRC is provided by National Institute of Mental Health Center Grant 5 P50 MH45294 (HIV Neurobehavioral Research Center). Additional support is provided by 5 R01 MH43298 (Neuropsychiatric Sequelae of HTLV-III Infections), 5 R29 MH45688 (Object-Oriented Simulation of HIV and CNS/HIV Infection, 1 R01 MH46255 (Psychosocial Moderators of Disease Progression in AIDS), 1 R01 NS27810 (Role of Immune Responsiveness in HIV Encephalopathy), and the Henry M. Jackson Foundation (Psychiatric Natural History Study: Factors Related to Human Immunodeficiency Virus Transmission and Morbidity). The

views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of the Navy, Department of Defense, or the United States Government.

REFERENCES

- 1. World Health Organization: The HIV/AIDS Pandemic: 1993 Overview Global Program on AIDS. Geneva, World Health Organization, 1993, 1
- Choi S, Lagakos S, Schooley R, et al: CD4+ lymphocytes are incomplete surrogate marker for clinical progression in persons with asymptomatic HIV infection taking zidovudine. Ann Intern Med 118:674-680, 1993
- Hays RD, Shapiro MF: An overview of generic health-related quality-of-life measures for HIV research. Qual Life Res 1:91– 97, 1992
- 4. Lubeck DP, Fries JF: Changes in quality of life among persons with HIV infection. Qual Life Res 1:359-366, 1992
- 5. Ganz PA, Coscarelli Schag CA, Kahn B, et al: Describing the health-related quality of life impact of HIV infection: Findings from a study using the HIV Overview of Problems–Evaluation System (HOPES). Qual Life Res 2:109–119, 1993
- Wachtel T, Piette J, Mor V, et al: Quality of life in persons with human immunodeficiency virus infection: Measurement by the Medical Outcomes Study instrument. Ann Intern Med 116:129–137, 1992
- Schag CA, Ganz PA, Kahn B, et al: Assessing the needs and quality of life of patients with HIV infection: Development of the HIV Overview of Problems-Evaluation System (HOPES). Qual Life Res 1:397-413, 1992
- Kaplan RM, Bush JW: Health-related quality of life measurement for evaluation research and policy analysis. Health Psychol 1:61-80, 1982
- 9. Weinstein MC, Stason WB: Cost-Effectiveness of Coronary Artery Bypass Surgery. Boston, Harvard University Center for Analysis of Health Practice, 1983
- Kaplan RM, Bush JW, Berry CC: Health status: Types of validity for an index of well-being. Health Serv Res 11:478-507, 1976
- 11. Russell L. Is Prevention Better than Cure? Washington, DC, The Brookings Institution, 1986
- Kaplan RM, Anderson JP: The quality of well-being scale: Rationale for a single quality of life index. In Walker SR, Rosser R (eds), Quality of Life: Assessment and Application. London, MTP Press, 1988, 51-77
- Anderson JP, Kaplan RM, DeBon M: Comparison of responses to similar questions in health surveys. In Fowler F (ed), Health Survey Research Methods. Washington, DC, National Center For Health Statistics, 1989, 13–21

- 14. Anderson JP, Kaplan RM, Berry CC, et al: Interday reliability of function assessment for a health status measure: The quality of well-being scale. Med Care 27:1076-1084, 1989
- Kaplan RM, Bush JW, Berry CC: The Reliability, Stability, and Generalizability of a Health Status Index. Proceedings of the Social Status Section. Washington DC, American Statistical Association, 1978, 704–709
- Kaplan RM, Anderson JP: The general health policy model: An integrated approach. In Spilker B (ed), Quality of Life Assessments in Clinical Trials. New York, Raven, 1990, 131-149
- Heaton RK, Kirson D, Velin RA, et al: Utility of clinical ratings for detecting cognitive change in HIV infection. In Grant I, Martin A (eds), Neuropsychology of HIV Infection. New York, Oxford University Press, 1994, 188–206
- Gullivich SJ, Kalmijn JA, Thal LJ, et al: Sensory testing in HIV-I infected men. Arch Neurol 49:1281–1284, 1992
- Jernigan TL, Archibald SL, Hesselink JR, et al: Magnetic resonance imaging morphometric analysis of cerebral volume loss in human immunodeficiency virus infection. Arch Neurol 50:250-255, 1993
- NcNair DM, Lorr M, Droppleman LF: Profile of Mood States (POMS) Manual. San Diego, Educational and Industrial Testing Services, 1980
- 21. Beck AT: Cognitive Therapy and the Emotional Disorders. New York, Hoeber, 1976
- Dudley DL, Martin CJ, Holmes TH: Dyspnea: Psychologic and physiologic observations. J Psychosom Res 11:325–339, 1968
- Kaplan RM, Anderson JP, Wu AW, et al: The quality of well-being scale: Applications in AIDS, cystic fibrosis, and arthritis. Med Care 27(Suppl 3):S27-S43, 1989
- 24. Wu AW, Mathews WC, Brysk LT, et al: Quality of life in a placebo-controlled trial of zidovudine in patients with AIDS and AIDS-related complex. J Acquir Immune Defic Syndr 3:683-690, 1990
- 25. Kaplan RM, Atkins CJ, Timms R: Validity of a quality of well-being scale as an outcome measure in chronic obstructive pulmonary disease. J Chronic Dis 37:85–95, 1984
- Orenstein DM, Pattishall EN, Ross EA, et al: Quality of well-being before and after antibiotic treatment of pulmonary exacerbation in cystic fibrosis. Chest 98:1081–1084, 1990
- 27. Bombardier C, Ware J, Russell IJ, et al: Auranofin therapy, quality of life for patients with rheumatoid arthritis: Results of a multi-center trial. Am J Med 81:565–578, 1986
- Kaplan RM, Kozin F, Anderson JP: Measuring quality of life in arthritis patients (including discussion of a general healthdecision model). Qual Life Cardiovasc Care 4:131–139, 1984
- 29. Kaplan RM: Value judgment in the Oregon Medicaid experiment. Med Care 32:975–988, 1994
- Hughes TE, Coons SJ, Kaplan RM, et al: Reweighting the quality of well-being scale in HIV-infected patients. Qual Life Res 3:79-80, 1994