Measuring Health-Related Quality of Life Among Patients Infected with Human Immunodeficiency Virus

Diane Franchi and Richard P. Wenzel

From the Division of Quality Health Care, Medical College of Virginia, Virginia Commonwealth University, Richmond, Virginia

Optimizing health-related quality of life (HRQL) has become an important treatment focus for patients infected with human immunodeficiency virus (HIV). Consideration of HRQL is especially relevant with the development of new antiretroviral agents that have significant side effects. The measurement of HRQL is still in evolution, and several methods have been used to quantify HRQL in the HIV-infected population. A review of existing studies shows that HRQL scores do not always correlate with disease stage or health indices and that symptoms have a significant impact on HRQL. Studies have also revealed that certain therapies for HIV and opportunistic infections exact a significant cost in terms of HRQL. HRQL outcomes will play a major role in treatment decisions for HIV-infected patients and in the development and marketing of new pharmaceutical agents in the near future.

Quality of life has become an important outcome of interest in clinical medicine and has “grown from a cottage industry to a large academic enterprise” [1]. Nevertheless, controversy and criticism have surrounded its definition, scope, and applicability [1,2]. The purpose of this article is to review the types of quality-of-life instruments developed to date, the methods used to evaluate them, and their applications to HIV disease through July 1996.

In the most general terms, quality of life extends beyond health. It may include satisfaction with one’s geographic location, access to education, or level of income. In a clinical setting, the definition is limited to those aspects of life directly affected by the health state or health care, extends beyond conventional assessments of health, and is often referred to as health-related quality of life (HRQL) [3]. HRQL is the focus of this review, and this term will be used throughout the article.

The increasing number of HRQL instruments being developed reflects a lack of consensus for a single instrument that adequately measures HRQL with enough sensitivity to detect significant changes for all diseases. HRQL instruments have been used in several different ways: to compare treatment options in clinical trials, to assess outcome measures in health services or evaluation research, to assist in cost-utility analyses, to screen and monitor individual patient care, to survey populations for perceived health problems, and to conduct medical audits [4]. Although some of these uses are viewed skeptically, their utility in clinical trials to identify optimal treatment regimens with quantitative certainty is difficult to deny when the appropriate instrument is used.

Methods

Information concerning the categorization and review of HRQL instruments was obtained from a literature search using the National Library of Medicine Bibliographic Retrieval Services, Inc., or MEDLINE. Only English-language review articles found with the key words “quality of life” or “health-related quality of life” were considered.

Journal articles concerning HRQL for HIV-infected patients were obtained through a MEDLINE search using combinations of the key words “quality of life,” “health-related quality of life,” “health status indicators,” “disability,” and “human immunodeficiency virus,” “acquired immunodeficiency syndrome,” or “communicable diseases.” Only English-language articles published before July 1996 were considered for review. The bibliographies of the articles discovered in this way were also used as a source of articles for review.

Categorization of HRQL Instruments

The most distilled categorization of HRQL instruments puts them into two realms, generic and specific [3]. Generic instruments are designed to incorporate multiple dimensions of HRQL, such as the physical, social, and emotional dimensions. Because of their breadth, generic instruments can be used to compare HRQL among different diseases or populations. Generic instruments can further be broken down into “health profiles” and “utility-based” instruments [3].

Health profiles are batteries of questions directed to the patient, covering multiple aspects of HRQL separately. They vary with respect to their emphasis, the number of questions, and the amount of time it takes to complete them. The scales employed may have questions for individual categories, yielding separate scores, or may generate one total score. An instrument that yields one score is called an index [3]. Health profiles that have been well studied for validity and reliability for various diseases include the Sickness Impact Profile, the Medical Out-
comes Study General Health Survey, the Nottingham Health Profile, and the Health Assessment Questionnaire. Although these instruments may be very useful in comparing HRQL between some diseases, their utility for other diseases may vary. For example, the VF-14, a questionnaire designed to measure functional impairment specifically caused by cataracts, was able to detect impairment of HRQL that the Sickness Impact Profile was not, when the scales were compared directly with each other [5].

Utility-based instruments evolved from economic theory of choice and decision analysis and are used when a decision must be made between alternatives with limited information available [6]. The most widely used scales are those that measure HRQL by a single numerical value on a continuum from death (0.0) to complete health (1.0). This number indiscriminately includes all aspects of a patient’s HRQL. Examples of instruments used for obtaining this rating include the Standard Gamble and Time Trade-Off analyses.

The Standard Gamble asks patients to “gamble” (for example, with therapeutic options) between their own present health state and complete health or immediate death. The Time Trade-Off approach involves how much time in full health subjects would trade for their own present health state [7]. The Quality-Adjusted Time Without Symptoms of Disease and Toxicity of Treatment (Q-TWiST) instrument partitions HRQL choices and examines the partitions over time. Specifically, it measures quality-weighted survival time by summing weighted scores through well-defined health states [8]. The Q-TWiST and other utility-based instruments are most useful when there is a cost-utility comparison of treatments required. These instruments generally place less emphasis on the individual patient than on populations of patients, and they do not distinguish the effects of a disease or treatment on particular aspects of HRQL [3].

Specific instruments measure HRQL for a specific disease (Living with Asthma Questionnaire), a particular aspect of patient care (Hospice Quality-of-Life Index), a particular dimension of HRQL (Profile of Mood States), or a particular population such as geriatric or pediatric patients. Specific measures, therefore, address health-related issues of special interest to the appropriate patient. Because of their targeted inquiry, specific instruments are not used to compare HRQL among diseases.

Currentl

Ideal HRQL Instruments

Instruments used to evaluate HRQL should possess the attributes of reproducibility, responsiveness, validity, and interpretability [3, 4]. Reproducibility or reliability means that the instrument elicits similar scores on repeated testing at times when HRQL is not expected to change. If several interviewers are administering the questionnaire, interrater reliability should also be established [9].

The four measures of validity are sensitivity, specificity, and positive and negative predictive values. If a “gold standard” exists, an instrument can be evaluated against that standard. Currently, a universal gold standard does not exist, but if a longer version of a questionnaire has already been established as valid for a disease, a shorter version can be validated against the longer one [3]. When no gold standard exists for explicit comparison, one could establish validity implicitly by asking patients, doctors, nurses, social scientists, or allied health care workers if the method of measurement seems to cover the intended aspects of HRQL (face validity) comprehensively (content validity) [4]. Construct validity is based on the comparison of an instrument’s performance to a theoretical construct assumed by the investigator to be true [3]. Alternatively, validity can be established as in the case of the Health Assessment Questionnaire, which was originally used to measure HRQL for patients with rheumatoid arthritis (nurse assessors were sent to patients’ homes to confirm their responses to functional disability questions) [10].

Measuring responsiveness, or the ability of an instrument to detect significant changes in HRQL quantitatively by a significant change in score, also becomes difficult when there is no established standard. It may be that the HRQL instrument must be compared to other health status measures or to the patient’s or the doctor’s perception of a change in health status [4].

Interpretability is the ability to translate a significant change or difference in HRQL score to a change in health status. For example, does a change in score mean that the patient is much better or just moderately better, and how seriously does that impact on the patient’s daily life?

HRQL Studies for Patients with HIV Disease

HRQL instruments are most applicable to chronic diseases. At the start of the HRQL movement, the majority of studies focused on pulmonary, cardiac, and rheumatological diseases. In the last 7–8 years, there has been an explosion of HRQL work done with HIV disease. Although it is not the only infectious disease studied, HIV infection is the most extensively studied. The HIV-infected population presents a special challenge to those measuring HRQL because it is generally younger than other populations and often belongs to “marginalized” communities (homosexual males, intravenous drug abusers, prisoners) [11]. The utility and value of the studies published are often difficult to interpret: one must look at the measure being used, whether it represents an observed or self-reported HRQL measurement, whether the measure is utility based or is a psychometric health scale, and whether the measure has been validated.

Initial studies incorporated mostly functional scales in clinical trials but did not include formal validation of these scales. The Karnofsky Performance Status (KPS) score, a physician-rated functional scale, was used in a double-blind, placebo-controlled trial to determine the efficacy of administering 250 mg of zidovudine (AZT) every 4 hours to patients with AIDS and AIDS-related complex (ARC) over a 24-week period (table 1) [12].
### Table 1. Studies of health-related quality of life (HRQL) for HIV-infected patients.

<table>
<thead>
<tr>
<th>Patient population or disease (treatment) studied</th>
<th>HRQL instrument(s), reference</th>
<th>Major outcome/point</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIDS and ARC (zidovudine vs. placebo)</td>
<td>KPS [12]</td>
<td>Zidovudine recipients’ functional scores improved transiently at 12 w vs. placebo recipients’, but this difference was no longer evident at 16 weeks.</td>
</tr>
<tr>
<td>AIDS and ARC (zidovudine vs. placebo)</td>
<td>QWB, KPS [13]</td>
<td>QWB scores were better for zidovudine group at 24 w and 1 y (reflects differences in mortality). When death was not considered as an outcome, functional scores were transiently better for zidovudine group at 24 w. No difference was evident at 1 y.</td>
</tr>
<tr>
<td>HIV encephalitis</td>
<td>KPS [14]</td>
<td>Advanced disease by Walter Reed Classification predicted a rapid progression to functional impairment from encephalitis</td>
</tr>
<tr>
<td>HIV− vs. asymptomatic HIV+ patients</td>
<td>QLI [15]</td>
<td>QLI is valid for measuring HRQL in HIV+ patients, but no difference was found between HRQL for asymptomatic HIV+ and HIV− patients.</td>
</tr>
<tr>
<td>Asymptomatic HIV+, ARC, and AIDS patients</td>
<td>SIP, SDS [16]</td>
<td>ARC and AIDS patients scored the worst in the psychosocial dimensions of the SIP and for the SDS; ARC patients had the worst scores overall for both scales. Neither scale was able to distinguish between HRQL for AIDS and ARC patients.</td>
</tr>
<tr>
<td>Asymptomatic HIV+ and early symptomatic HIV+ patients</td>
<td>MOS HIV Survey [17]</td>
<td>MOS HIV Survey is able to distinguish between asymptomatic and early symptomatic HIV+ patients. Asymptomatic HIV+ patients scored better than patients with hypertension, diabetes mellitus, myocardial infarction, or depression.</td>
</tr>
<tr>
<td>AIDS</td>
<td>MOS SF-20 [18]</td>
<td>MOS SF-20 is a reliable and sensitive measure of HRQL for AIDS patients but does not contain assessment of energy, cognitive function, or distress secondary to health problems. Presence of symptoms was the strongest indicator of low score.</td>
</tr>
<tr>
<td>Early symptomatic HIV disease (high-dose zidovudine vs. placebo)</td>
<td>MOS HIV Survey [19]</td>
<td>Zidovudine provided 9 mo of disease-free survival at a moderate cost to HRQL at 1 y. HRQL for both groups became similar after that, presumably secondary to disease progression in the placebo group.</td>
</tr>
<tr>
<td>Early symptomatic HIV disease (high-dose zidovudine vs. placebo)</td>
<td>Q-TWiST [20]</td>
<td>Treatment provided more Q-TWiST in an 18-mo period if HRQL after disease progression is considered 10%–20% worse than HRQL after a severe symptomatic event.</td>
</tr>
<tr>
<td>Early symptomatic HIV disease (low-dose zidovudine vs. placebo)</td>
<td>Q-TWiST [21]</td>
<td>Reduction in HRQL secondary to severe side effects of zidovudine nearly equaled the increase in HRQL associated with delay in disease progression.</td>
</tr>
<tr>
<td>AIDS patients receiving prophylaxis with rifabutin against <em>Mycobacterium avium</em> complex</td>
<td>Q-TWiST, KPS, HUI [22]</td>
<td>HRQL scores were much better for those patients receiving prophylaxis against <em>M. avium</em> complex.</td>
</tr>
<tr>
<td>Asymptomatic HIV+, symptomatic HIV+, and AIDS patients</td>
<td>MOS HIV Survey, SIP–home management scale, Standard Gamble, other health-status measures [23]</td>
<td>Standard Gamble is not able to discriminate between HIV health states when compared directly with the MOS HIV Survey.</td>
</tr>
<tr>
<td>AIDS</td>
<td>Combination of several existing scales compiled by Cleary et al. [24]</td>
<td>Instrument is valid and reliable. Aggregates of symptoms and fatigue scores were the best predictors of functioning. Physical and mental health are predictors of life satisfaction but cannot explain all the variance.</td>
</tr>
<tr>
<td>Asymptomatic HIV+, symptomatic HIV+, and AIDS patients</td>
<td>MOS HIV Survey, Cleary’s HRQL Instrument, Hospital Anxiety and Depression Scale [25]</td>
<td>MOS HIV Survey and Cleary’s HRQL Instrument are reliable and valid. Psychological and cognitive measures did not correlate to health indices or CDC stage.</td>
</tr>
<tr>
<td>Asymptomatic HIV+, symptomatic HIV+, and AIDS patients</td>
<td>AIDS-HAQ [26]</td>
<td>Both symptomatic patients without a diagnosis of AIDS and AIDS patients showed significant drops in HRQL over a 12-mo period, especially in role functioning and symptoms. Symptomatic HIV+ and AIDS patients also showed decline in ability to work as many hours as before.</td>
</tr>
<tr>
<td>Asymptomatic HIV+, symptomatic HIV+, and AIDS patients (some with cancer)</td>
<td>HOPES [27]</td>
<td>Medical and demographic variables could not account for all of the variability of HRQL ratings.</td>
</tr>
<tr>
<td>HIV+ patients with chronic diarrhea vs. HIV+ patients without diarrhea</td>
<td>AIDS-HAQ [28]</td>
<td>Patients with diarrhea had worse declines in HRQL, especially in the areas of role functioning and general health. They also accrued more medical care costs.</td>
</tr>
<tr>
<td>Recurrent cytomegalovirus retinitis (ganciclovir vs. foscarnet vs. combination therapy)</td>
<td>MOS HIV Survey, cytomegalovirus retinitis–specific HRQL instrument [29]</td>
<td>Combination therapy proved to be more effective for relapsed cytomegalovirus retinitis but at a higher cost in terms of HRQL.</td>
</tr>
<tr>
<td>Anemia associated with zidovudine</td>
<td>Visual analogue scales for energy, activities of daily living, and overall HRQL [30]</td>
<td>All three aspects of HRQL improved for erythropoietin recipients whose hematocrit reached ≥38% without transfusion or discontinuation of zidovudine and who started with endogenous erythropoietin levels of ≤500 IU/L.</td>
</tr>
<tr>
<td>Disseminated <em>M. avium</em> complex infection</td>
<td>Eastern Cooperative Oncology Group Functional Scale [31]</td>
<td>Patients whose blood culture for <em>M. avium</em> complex converted from positive to negative showed improved performance overall.</td>
</tr>
</tbody>
</table>

**NOTE.** AIDS-HAQ = AIDS–Health Assessment Questionnaire; ARC = AIDS-related complex; CDC = Centers for Disease Control and Prevention; HOPES = HIV Overview of Problems Evaluation System; HUI = Health Utility Index; KPS = Karnofsky Performance Status; MOS = Medical Outcome Survey; QLI = Quality of Life Index; Q-TWiST = Quality-Adjusted Time Without Symptoms of Disease and Toxicity of Treatment; QWB = Quality of Well-Being Scale; SDS = Symptom Distress Scale; SIP = Sickness Impact Profile; + = seropositive; − = seronegative.
In addition to transient improvement in KPS scores, AZT recipients developed fewer opportunistic infections and had transient increases in CD4 cell counts, informally establishing validity of KPS scores. This was most apparent in those patients whose CD4 cell count initially was <100/mm³. However, there was no difference in the therapeutic effects of AZT on HRQL for patients stratified by an initial diagnosis of AIDS vs. ARC. The effects of AZT toxicity on HRQL were not established. No information on interrater reliability of KPS scores was given.

The Quality of Well-Being (QWB) scale combines preference-weighted measures of symptoms and functioning that are expressed on a continuum from death (0.0) to asymptomatic, optimum functioning (1.0). It includes elements of mobility, physical activity, and social activity and is therefore more comprehensive than the KPS scale. The QWB and KPS scales were used in a small study in the San Diego arm of the above-mentioned trial comparing AZT and placebo in AIDS and ARC patients (table 1) [13]. The mortality in the AZT group was less than that in the placebo group in this study.

Because the QWB scale factors death into its score, the QWB score was skewed in favor of the AZT group. However, the QWB scores showed trends similar to those in the KPS scores when data concerning those who died were removed, adding validity to the use of the QWB scale for HIV-infected patients. It should be noted that the QWB scale was administered by two trained interviewers who were blinded to treatment groups. The KPS score was determined by one care provider, and no interrater reliability was established.

In 1989 Moller et al. used the Karnofsky Index to follow 19 patients with HIV encephalitis and determined the predictors of their clinical course over time in terms of function (table 1) [14]. This was successful in linking a more conventional measure of HIV health status, the Walter Reed Classification, to a change in functional status score. While it was a small study and no mention was made of who determined KPS scores, it indirectly validated the KPS scale for use with HIV-infected patients.

As time went on, more formal attempts at validating HRQL scales for HIV disease were being made. Spitzer’s Quality of Life Index (QLI) was examined for validity in measuring HRQL for HIV-infected patients by comparing it to several well-established measures of psychopathology and psychosocial and physical functioning, including the Social Supports Scale and the Global Assessment of Functioning Scale (table 1) [15]. There was a highly significant correlation found between the support domain of the QLI and the Social Supports Scale. A significant correlation was found between the outlook domain and most measures of psychopathology, and a modest correlation between the activity domain of the QLI and a medical staging scale and the Global Assessment of Functioning Scale. It was concluded that the five-item physician-rated scale was valid for use in measuring HRQL, but its ability to distinguish between levels of severity of disease and progression of disease was not established.

The Sickness Impact Profile (SIP) and the Symptom Distress Scale (SDS) have been used together in a study to validate their applicability in HIV disease (table 1) [16]. Both instruments showed similarly poor scores in psychosocial domains for AIDS and ARC patients. It was not in the power of either scale to distinguish between AIDS and ARC patients, and this study was not designed to assess the scales’ ability to detect changes in clinical status over time. What is noteworthy, however, is that these scales were patient-rated, not physician-rated.

Various versions of the Medical Outcomes Study (MOS) General Health Survey, originally developed by the Rand Corporation (Santa Monica, CA), have been applied to HIV-infected patients. In 1991 Wu and colleagues formulated a short “HIV-relevant,” patient-rated, 30-item questionnaire from the functional and overall well-being subscales of the MOS General Health Survey, called the MOS HIV Survey. They showed that the MOS HIV Survey reliably distinguished between asymptomatic HIV-infected patients and those with early symptomatic disease not categorized as having AIDS, in the areas of overall health, pain, physical function, role function, cognitive function, and overall quality of life (table 1) [17]. This was consistent with the findings of the SIP and SDS and had the advantage of being less time-consuming for the patients to complete. Again, this study looked at patients at only one point in time. Because the MOS HIV Survey was adapted from a generic, psychometric scale, the scores for HIV-infected patients could be compared with those for patients with other chronic diseases (table 1).

The original short form of the MOS General Health Survey, containing 20 questions (MOS SF-20), was used in a large multicenter trial to help establish the MOS SF-20 as a reliable and sensitive measure for AIDS patients (table 1) [18]. This was achieved by comparing the results from the MOS SF-20 to patients’ reports of the presence of symptoms, their age and sex, non-white race, and intravenous drug use. Unlike the MOS HIV Survey, the MOS-SF 20 does not include the assessment of energy, cognitive function, or distress secondary to health problems.

In 1993 the MOS HIV Survey was used to compare functional status and well-being in a sample of patients from a placebo-controlled trial of high-dose AZT for early symptomatic HIV disease (AIDS Clinical Trials Group [ACTG] Study 016) (table 1) [19]. It is of interest that at 12 weeks there were no differences in any subscale tested between the AZT-treated group and the placebo group. At 24 weeks, there was a significantly better HRQL measured in all areas of well-being for the placebo group, including overall health, energy, mental health, health distress, and pain, but no differences were found in the functional subscales. At 52 weeks the AZT-treated and placebo groups had similar scores overall. The magnitude of decline in HRQL scores due to treatment with AZT was similar to that of going from being asymptomatic to symptomatic in the previously mentioned study using the MOS HIV Survey.

The Q-TWIST method was applied retrospectively to the same study patients in ACTG Study 016; however, data from...
all patients were included in this study (table 1) [20]. Time-related health states were defined and categorized as (1) the time from study entry until the occurrence of an adverse symptomatic event or disease progression, (2) the period of time between the occurrence of the first severe, adverse symptomatic event and disease progression, and (3) the remaining survival time after disease progression. The final results were very similar to those of the study using the MOS HIV Survey, showing a cost in terms of HRQL for treatment with AZT and thus validating those results. The Q-TWiST, however, is unable to identify which aspects of HRQL are more significantly affected and is therefore less interpretable in terms of individual patients but is better suited for financial analyses.

In 1994 the Q-TWiST method was utilized again to compare low-dose AZT (500 mg/d) and placebo for asymptomatic HIV-infected patients (table 1) [21]. The original study (ACTG Study 019) compared AZT at two different doses (500 mg/d and 1,500 mg/d) to placebo in these patients. Progression-free survival at 18 months in the original study was 94% for low-dose AZT recipients and 89% for placebo recipients. Health states were defined similarly to those in the aforementioned study.

The two groups had approximately equal amounts of time before disease progression or an adverse symptomatic event, and AZT recipients had more time after an adverse event than did the placebo group. Theoretical threshold utility analyses were again constructed to define the limits of patient-ascribed utility weights that would produce optimal treatment planning. Even at lower doses, AZT was found to impinge on HRQL.

A modified Q-TWiST method was used by Revicki et al. in 1995 to assess the impact on HRQL associated with rifabutin prophylaxis for Mycobacterium avium complex (MAC) in patients with AIDS (table 1) [22]. This was a secondary, retrospective analysis of two previously reported multicenter, randomized, placebo-controlled clinical trials. Eighteen health states were defined with use of a combination of the Karnofsky Index, symptoms of fever and night sweats, anemia, hospitalizations, and positive or negative blood cultures for MAC. Physicians were then asked to rate the health states by developing a utility score using the Health Utility Index (HUI), which covers dimensions of sensory ability, mobility, emotional function, self-care, pain, and discomfort. Q-TWiST scores with these measures and health states were significantly higher for patients who received rifabutin prophylaxis. It should be noted, however, that there was little accounting for adverse effects of rifabutin therapy in this study and that patient input was not utilized for assigning utility weights.

The Standard Gamble instrument has been shown not to have the ability to discriminate between HIV disease states and shows no correlation with the MOS HIV Survey and other health status measures (table 1) [23]. Therefore, the value of the utility-based measure for HIV disease, other than the Q-TWiST method, is still elusive.

Cleary et al. developed an instrument to determine HRQL for AIDS patients with use of questions and scales from several different sources (table 1) [24]. It covers the dimensions of life satisfaction, general health perception, physical functioning, emotional well-being, fatigue, disability, pain, memory problems, and other symptoms. When examining the extent to which symptoms predict reported levels of functioning and how functioning and symptoms impact on overall health assessment and life satisfaction, Cleary et al. [24] found that each symptom from the symptom scale correlated with level of functioning, but the best predictors of functioning were the summary physical-symptom score and the fatigue score. The instrument was administered face-to-face in this study.

Later, a self-administered version of the instrument and the MOS HIV Survey were compared to the Hospital Anxiety and Depression Scale score, the Centers for Disease Control and Prevention (CDC) stage, and CD4 and CD8 cell counts of HIV-infected homosexual men, including asymptomatic patients, symptomatic patients, and patients with the diagnosis of AIDS (table 1) [25]. This study confirmed the reliability and validity of the two scales in measuring HRQL for HIV-infected patients by comparing them to the Hospital Anxiety and Depression Scale. However, psychological and cognitive measures, including global ratings of HRQL, did not correlate with health indices or CDC stage.

Changes in HRQL in HIV-infected patients have also been assessed over a 12-month period with use of a hybrid questionnaire, called the AIDS-HAQ, consisting predominantly of the Health Assessment Questionnaire with subscales from the MOS General Health Survey concerning mental health, cognition, energy, social functioning, and symptom scale (table 1) [26]. There was a striking loss of HRQL in symptomatic patients, even if they did not have an AIDS-defining illness. Patients with AIDS also had a significant decline in HRQL, even more than that for symptomatic patients without AIDS.

Both groups had significant declines in all aspects of role functioning and had significantly high levels of disease symptoms, but they showed no significant decline in cognition or mental health. Patients with AIDS and symptomatic HIV-infected patients also reported fewer hours at work and more disability days than asymptomatic patients. These findings were consistent with those of other studies except that the AIDS-HAQ demonstrated worse HRQL for HIV-infected patients with AIDS than for those who were only symptomatic.

More recently, the HIV Overview of Problems Evaluation System (HOPES) was developed [27]. In a cross-sectional study examining the relationship between clinical and biological factors and HRQL measured by the HOPES instrument, it was noted that medical and demographic variables explained only a portion (35%) of the variability of HRQL ratings for these patients (table 1).

HRQL instruments have been utilized for specific problems associated with HIV, such as the MAC bacteremia study previously cited. The AIDS-HAQ was used to compare HRQL for HIV-infected patients with diarrhea and low CD4 cell counts vs. that for those with low CD4 cell counts alone (table
The results revealed that patients with chronic diarrhea had more severe declines in HRQL over 1 year than their counterparts in role functioning (social activity, daily living, energy, and cognition) and in the general health area. They also accrued higher costs for medical care and home health care services.

In a comparison of foscarin, ganciclovir, and combination therapy for relapsed cytomegalovirus (CMV) retinitis in a randomized, controlled clinical trial, combination therapy proved to be more effective in preventing the progression of CMV retinitis, yet the cost in terms of HRQL to the patient was significant (table 1) [29]. Both a visual-specific HRQL scale and the MOS HIV Survey were used in this study. In addition, more patients were withdrawn from the combined-therapy arm of the study than from the monotherapy arms for reasons of toxic effects revealed by laboratory results.

Visual analogue scales for energy, activities of daily living, and overall quality of life were used in a randomized, double-blind, placebo-controlled trial of treatment with recombinant human erythropoietin (r-HuEPO) for anemia in patients with AIDS receiving AZT therapy [30]. The trial showed that all three aspects of HRQL improved significantly for r-HuEPO recipients starting with endogenous erythropoietin levels of ≤500 IU/L whose hematocrit reached 38% or more without transfusion or discontinuation of AZT therapy. However, changes in HRQL were not analyzed for patients whose initial endogenous erythropoietin levels were >500 IU/L and who did not have a significant change in hematocrit. Therefore, the correlation between improvement in hematocrit and improvement in HRQL could not be made definitively.

A retrospective study in 1994 examined the effects of MAC bacteremia culture conversion on functioning, survival, and morbidity among patients with AIDS (table 1) [31]. The patients had previously enrolled in an open, multicenter HIV treatment study. The Eastern Cooperative Oncology Group (ECOG) functional scale was used, and information was extracted from the patient’s case record for scoring by a physician. The result was improved overall performance in those patients whose blood culture for MAC converted from positive to negative. The study was limited by the fact that HRQL scores were physician-assigned, the reliability of physician raters was not established, and no formal attempts at validation were made.

The HIV Patient-Reported Status and Experience Scale (HIV-PARSE instrument) can also be found in the HIV literature [32]. It has already been used in several ACTGs.

**Discussion**

Most of the initial studies concerning HRQL for HIV-infected patients focused on establishing the responsiveness, reliability, and validity of existing and newly developed HRQL instruments for this population. The measurement of HRQL for patients with HIV infection is still in evolution, as it is for patients with other diseases. The reader should be aware that requirements for formal validation of HRQL instruments are becoming more stringent in order to make the process more scientific and accurate. Those instruments that have been formally validated—the MOS HIV Survey, the AIDS-HAQ, and the QLI—show that for patients with (early) symptomatic HIV infection there is a significant impact on HRQL that is similar to that for those with AIDS (late-stage HIV infection). Specifically, there is little correlation between psychological scores and disease stage or health indices.

New and expensive combination therapies for HIV infection that lengthen life but manifest serious side effects and impinge on daily living have made utilization of the measurement of HRQL imperative. Both the financial burden associated with these regimens and compliance issues also warrant investigation of HRQL associated with them. Convincing data from the studies involving AZT and combination therapy for CMV retinitis show that the benefits of effective treatments do not come without a price in terms of HRQL. This is valuable information for the physician and patient for predicting the impact of new treatments on HRQL and determining optimal treatment regimens.

With new drug combinations for HIV and opportunistic infections, there will likely always be an emphasis on utility analysis for financial and marketing purposes. However, this is not to say that patient-rated psychometric elements cannot be incorporated into these analyses. For example, use of the modified Q-TWiST method by Revicki et al. [22] to determine the effects of prophylaxis for MAC on HRQL was a more compelling study than that using the Q-TWiST for AZT treatment of early HIV disease because health states in the former study were more specific and gauged by the HUI, which determines utility weights with psychometric properties. The challenge in the future will be to allow patients to determine their own utility weights, such as HUI weights, in a meaningful, interpretable way for health states.

Some psychometric models, such as the AIDS-HAQ, include dimensions that focus on cost and drug toxicity. Whereas these psychometric measures are not always optimal in clinical trials involving new medications, they are excellent for comparing the specific HRQL manifestations of complications of HIV infection and opportunistic infections, such as chronic diarrhea, while also accounting for the financial burden and iatrogenic therapeutic toxicities of these phenomena.

Finally, HRQL scales developed from generic psychometric instruments such as the MOS HIV Survey enable one to compare HRQL with HIV infection to that with other chronic diseases. This is an advantage that HIV-specific scales do not have. If these generic health profiles are used for specific opportunistic infections, it is probably best that they are used in conjunction with a disease-specific or organ-specific instrument to ensure their sensitivity.

Future refinements of HRQL instruments should include improvement of discrimination across disease stages that allows for improved interpretability of scores. Whatever the future
refinements will be, however, there is a necessity for infectious disease clinicians to understand HRQL scales now in order to evaluate clinical trials critically.

References