*Quality of Life Research*, 5, pp. 339–347

# A study of the reliability, validity and responsiveness of the HIV Overview of Problems Evaluation System (HOPES) in assessing the quality of life of patients with AIDS and symptomatic HIV infection

**J. B. De Boer, M. A. G. Sprangers, N. K. Aaronson, J. M. A. Lange and F. S. A. M. van Dam***

Division of Psychosocial Research and Epidemiology, The Netherlands Cancer Institute, Amsterdam, The Netherlands (J. B. De Boer, M. A. G. Sprangers, N. K. Aaronson, F. S. A. M. van Dam); World Health Organization Collaborating Center on Quality of Life, Amsterdam, The Netherlands (J. B. De Boer, N. K. Aaronson, F. S. A. M. van Dam); Department of Medical Psychology and Psychotherapy, Erasmus University Rotterdam, Rotterdam, The Netherlands (J. B. De Boer); National AIDS Treatment Evaluation Center (NATEC), Academic Medical Center, Amsterdam, The Netherlands (J. M. A. Lange); Department of Clinical Psychology, University of Amsterdam, The Netherlands (F. S. A. M. van Dam)

**The objective of this study was to evaluate the feasibility, reliability, validity and responsiveness of the HIV Overview of Problems Evaluation System (HOPES) in a Dutch sample. The HOPES was administered three times in a one-year period to a sample of 106 outpatients with a symptomatic HIV-infection ($n = 23$) or AIDS ($n = 83$). The HOPES is a self-report HIV-specific quality of life (QOL) questionnaire including five scales: physical, psychosocial and sexual functioning, medical interaction and partner relationship. QOL was also assessed with the EORTC Quality of Life Questionnaire (EORTC QLQ-C30), a 30-item self-report instrument. Clinical data included Centers for Disease Control and Prevention (CDC) stage, date of diagnosis and CD4 cell count. Patients needed approximately 20–30 minutes to complete the questionnaire. The five scales had good internal consistency reliability. Multitrait scaling analysis provided moderate support for item discriminant and convergent validity. The HOPES exhibited adequate levels of construct validity: (1) the inter-scale correlations and correlations with the EORTC QLQ-C30 were in the predicted direction; (2) it discriminated clearly between patients with AIDS and ARC and (3) it was able to document changes in QOL over time. Moreover, the HOPES was responsive to changes in clinical status over time as indicated by CD4 counts. This study provides further evidence of the reliability and validity of the HOPES and shows that this instrument is responsive to changes in CD4 cell counts.**

*Key words:* AIDS; HIV infections; HOPES; quality of life; reliability; validity.

## Introduction

AIDS or a symptomatic HIV infection can have devastating effects on patients' quality of life (QOL). In the short-run, patients are confronted with the physical and (possibly) neurological deficits associated with the disease, but they also have to accept the resulting discontinuation of work, limitations in social activities and the increased dependence on significant others (e.g., partners, relatives or friends). To provide appropriate medical and psychosocial care to HIV-infected patients, the immediate and long-term effects of the disease and treatment on patients' QOL need to be thoroughly understood. Toward this end, reliable and valid instruments are required that are also responsive to clinical changes in health status over time.

In 1989, when the current study was initiated, psychometrically robust measures that were specifically designed to assess the QOL and needs of HIV-infected patients were scarce. To our knowledge, the only

This research was made possible by a grant from the Dutch Ministry for Welfare, Public Health and Culture (grant no. 88-52).

* To whom correspondence should be addressed at Division of Psychosocial Research and Epidemiology, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands.

instrument then available was the HIV Overview of Problems Evaluation System (HOPES). The HOPES is derived from the Cancer Rehabilitation System (CARES), a 136-item self-administered, cancer-specific instrument that has been studied extensively in a range of cancer patient populations. The HOPES covers a wide range of daily problems HIV-infected patients may experience, including sexual problems, and problems with a partner and the medical team. In a recent study among 316 HIV-infected patients, the HOPES demonstrated adequate internal consistency reliability.[1,2] Further, concurrent validity was demonstrated with reference to the following criteria: the Medical Outcomes Study questionnaire adapted for HIV (MOS-HIV),[3,4] the Profile of Mood States (POMS),[5] the Perceived Adjustment to Chronic Illness Scale (PACIS)[6] and the Physical Activity Scale (PAS).[7]

The current study was undertaken to investigate the feasibility of employing the HOPES among patients treated in the Netherlands, to provide further evidence of the instrument's validity and reliability and to examine its responsiveness to changes in clinical status over time.

## Methods

### Subjects

One hundred and six patients were recruited from two patient groups: 45 patients were enrolled in a randomized clinical trial comparing the efficacy and tolerance of zidovudine monotherapy vs. a combination of zidovudine and interferon-α; the remaining 61 subjects were outpatients (not treated in a trial context) attending one of the five hospitals in Amsterdam providing AIDS care.

Inclusion criteria for the QOL study were: (1) a diagnosis of a symptomatic HIV infection or AIDS, (2) age 18 years and older and (3) ability to speak and read Dutch or English. An additional requirement for the non-trial patients was that they had to have been diagnosed within the previous six months. The trial patients also had to meet the specific inclusion criteria of the trial (e.g., CD4 count >150 cells/mm$^3$, Karnofsky Performance Status (KPS)[9] score >60).[8]

All patients were asked by their physician or a nurse to participate in the QOL study. Of the 49 trial patients approached, 45 (92%) agreed to participate. Among the four non-respondents, three patients preferred not to talk about their disease and one patient was already participating in another psychosocial study. The exact number of eligible non-trial patients could not be determined due to privacy regulations of the participating hospitals.

## Design

The first interviews with the non-trial patients were conducted within six months after their diagnosis of AIDS. The trial patients were interviewed after randomization, but prior to the start of the treatment. The follow-up assessments for both groups took place 6 and 12 months after the first interview. Most patients were interviewed in their homes. In case of hospitalization, the questionnaires were administered on the hospital ward.

## Study instruments

### Demographic and clinical data

A brief questionnaire was administered at the first interview to obtain information on age, living situation, level of education and risk group. From the medical records, data were abstracted on the date of diagnosis of symptomatic HIV infection or AIDS, CDC classification, CD4 counts at the three assessment points and date of death. At the time of this study, CD4 counts were not performed routinely on non-trial patients, resulting in missing data for 22 of these patients. For the trial patients, the clinical data also included the KPS score[9] as rated by the trial physician.

### The HOPES

In the current study, we employed the revised version of the HOPES described by Ganz and associates in a recent paper.[1] This version consists of 142 problem statements. There are 33 subscales that can be aggregated into five higher order summary scales and a miscellaneous scale: (1) Physical—eight subscales pertaining to physical problems and problems in daily functioning; (2) Psychosocial—including nine subscales relating to emotional functioning and problems in communicating and interacting with relatives and friends; (3) Medical Interaction—three subscales pertaining to problems with interaction and communication with doctors and nurses; (4) Sexuality—containing two subscales relating to sexual interest, activities and functioning; and (5) Partner—five subscales relating to problems in communication and

interaction with a partner(s). The remaining six subscales were grouped under a miscellaneous scale.

Patients were asked to rate the extent to which a problem statement applied to him/her during the past month on a 5-point scale ranging from 0 'not at all' to 4 'very much'. Eighty-nine items were relevant to all patients. The remaining sections of the questionnaire, each of which is preceded by a screening question, were rated only by patients to whom they applied. For example, patients in a steady relationship were asked questions pertaining to their relationship with their partner, whereas patients without such a relationship were asked questions relating to dating and initiating new relationships.

The HOPES can be scored on three levels: a global score (calculated by dividing the sum of all the 1–4 rated items by the total number of potential problems); the summary scales (the sum of the 1–4 rated items belonging to the summary scale divided by the number of potential problems in that scale); and the 33 more specific subscales. For all scales and the global score, higher scores indicate more problems.

The current paper is restricted to the results pertaining to the five summary scale scores (Physical, Psychosocial, Medical Interaction, Sexual and Partner) and the global score. The content of the subscales of the Miscellaneous scale is extremely diverse, and should therefore only be scored and interpreted on a subscale level.

The HOPES was translated into Dutch following a standard forward–backward procedure. The content of the translated version was reviewed by a range of internists and AIDS consultants who provide medical and psychosocial care to HIV-infected individuals. Additionally, the questionnaire was pretested on a number of patients. Doctors, nurses and patients reported that the problem statements reflected the experiences of patients with the disease and its treatment, thus lending support to the face and content validity of the questionnaire.

## The EORTC QLQ-C30

Quality of life was also assessed with the EORTC QLQ-C30, a 30-item self-report questionnaire originally developed for use with cancer patients.[10] The QLQ-C30 includes five functioning scales (physical, role, emotional, cognitive and social functioning), three symptom scales (pain, fatigue and nausea and vomiting), and an overall quality of life scale. The remaining six single items refer to additional physical symptoms (e.g., shortness of breath, diarrhea), and to financial difficulties patients may experience as a result of their disease or treatment. The time-frame of the questions encompasses the past week, and the majority of the items can be answered on a 4-point Likert scale. Scores of all of the QLQ-C30 scales and individual items were transformed linearly to a 0–100 point scale. High values on the functioning and overall QOL scales indicate high levels of functioning and QOL. Conversely, higher scores on the symptom scales indicate higher levels of symptomatology.

In a study of 156 patients with a symptomatic HIV infection or AIDS (including the 106 patients of the present study), the QLQ-C30 exhibited acceptable levels of reliability and validity.[11] In the current analysis, the QLQ-C30 is used as a concurrent measure to examine the construct validity of the HOPES.

## Statistical analysis

*Scale structure.* The multitrait analysis program (MAP-R)[12] was used to evaluate the hypothesized scale structure of the HOPES. Item–scale correlation matrices were employed to evaluate the convergent and discriminant validity of the items of the five summary scales of the HOPES. Evidence of item convergent validity is provided when an item correlates substantially (≥0.40) with its own scale, corrected for overlap. Item divergent validity is evidenced when an item correlates significantly (≥2 standard errors) higher with its own hypothesized scale than with other scales.[12] Given that the standard error ($SE = 1/\sqrt{N}$) used as the criterion for judging scaling successes and errors is highly sensitive to sample size and that, in the current case, the sample size was relatively small ($n = 106$ for the Physical, Psychosocial and Medical Interaction scales; $n = 55$ and $n = 54$ for the Sexual and Partner scales, respectively) a scaling success was counted when an item correlated higher (≥1 SE) with its own scale than with another scale. A scaling error was counted when an item correlated lower with its own scale than with another scale.

Due to the presence of sections within the HOPES that apply only to specific subgroups of patients, some summary scales were not completed by all patients. Since the MAP-R program does not accommodate missing data, three multitrait analyses were performed; one for the items of the three summary scales Physical, Medical Interaction and Psychosocial; one for the items of the Sexual scale; and one for the items of the Partner scale.

*Reliability.* The internal consistency reliability of the five summary scales was estimated by Cronbach's

alpha coefficient. Reliability was considered acceptable when Cronbach's α were ≥0.70.[13]

*Validity.* Four sets of analyses were carried out to evaluate the construct validity of the HOPES. First, we examined the correlations among the five summary scales, and between the summary scores and the global score. Moderate correlations (in the range 0.30–0.50) were expected among the summary scales. More substantial correlations (>0.50) were expected between the five summary scales and the global score, as they were all considered aspects of global QOL.

Second, the correlations between the HOPES summary and global scores, and the QLQ-C30 subscales were examined. High correlations (>0.60) were expected among scales pertaining to the same construct (e.g., between the physical scale of the HOPES and the subscales physical and role functioning of the QLQ-C30). Weaker associations were expected between subscales of the HOPES and the QLQ-C30 that were conceptually distinct (e.g., sexuality and cognitive functioning). Correlations were calculated by means of Pearson product-moment correlations.

The third set of analyses evaluated the extent to which the summary scales of the HOPES could discriminate between subgroups of patients known to differ in clinical status. At all three assessment points, scores of patients with AIDS vs. those with a symptomatic HIV-infection [previously termed AIDS-Related Complex (ARC)] were compared. Previous studies have indicated that patients with AIDS report more problems in physical and role functioning, but less psychological problems than patients with ARC.[14,15,16,17] Thus, it was hypothesized that patients with AIDS would report poorer physical functioning, but better psychosocial functioning. Student's *t*-tests were used to compare the mean scale scores for the subgroups at baseline and follow-up.

Fourth, changes over time in scale scores for the 82 surviving patients were examined by means of repeated measures ANOVA. In a longitudinal study design with symptomatic and AIDS patients, Lubeck and colleagues[15,16] found an increase in disease symptoms, and significant declines in all aspects of role functioning, but no significant declines in cognitive functioning or mental health. Accordingly, we predicted that, over time, patients would generally report an increase in physical problems, but a decrease in psychosocial problems.

Finally, a valid measure of QOL should be responsive to clinically important changes in health status over time. The observed changes in patients' CD4 lymphocytes are commonly used as a prognostic marker of clinical progression. When subjects' CD4 counts decrease, their risk of clinical progression increases.[18] To evaluate the responsiveness of the HOPES to changes in CD4 counts, the changes over time in mean HOPES scores of stable or improved patients were compared with those of deteriorated patients by means of repeated measures ANOVA. Deteriorated patients were defined as those whose CD4 count had decreased significantly (>10%) over a one year period. Stable/improved patients were characterized by stable (±10%) or increased (>10%) CD4 counts at 12 months follow-up. We hypothesized that stable/improved patients would have approximately the same or better HOPES scores at 12 months follow-up as at baseline, whereas patients with decreased CD4 counts were expected to score worse on the HOPES, particularly the Physical scale.

All data analyses were performed using SPSS 4.0. For *t*-tests and repeated measures ANOVA, *t*- and *F*-values associated with chance probabilities of less than 0.05 were considered statistically significant.

## Results

### Description of the sample at baseline and follow-up

One hundred and six patients were enrolled in the study and evaluated at baseline. At 6 and 12 months follow-up, 93 and 82 patients completed the questionnaires, respectively. During the one year follow-up, 18 patients had died, four patients declined further participation, and data for two patients were missing at either the 6 or 12 month assessments due to severe illness.

As is shown in Table 1, the sample consisted primarily of well-educated, male homo-/bisexuals with AIDS. On average, patients had been diagnosed with AIDS or symptomatic HIV infection within the past four months. The mean KPS score (available for the trial patients only) was high, indicating nearly normal physical functioning. Approximately half of the sample had CD4 counts below 200 cells/mm$^3$. At 12 months follow-up, the mean KPS score was still high (Mean = 90.5; SD = 12.6; $n$ = 39). The mean CD4 count had decreased to 176 cells/mm$^3$ (SD = 178; $n$ = 61) with 70% of the patients having CD4 counts below 200 cells/mm$^3$.

## Feasibility

On average, patients needed approximately a half hour ($n$ = 103, range 9–70 minutes) to complete the

**Table 1.** Baseline characteristics of patients, $n = 106$

| Characteristic | n | % |
|---|---|---|
| Sex | | |
| Male | 99 | 93.4% |
| Female | 7 | 6.6% |
| Mean (SD) age in years | 38.0 | (7.8) |
| Range | 22–65 | |
| Risk group | | |
| Bi-/homosexual | 98 | 92.5% |
| Heterosexual | 5 | 4.7% |
| Heterosexual/ex-IVDU* | 3 | 2.8% |
| Education† | | |
| Low | 32 | 30.5% |
| Middle | 34 | 32.4% |
| High | 39 | 37.1% |
| Living arrangement | | |
| Alone | 53 | 50.0% |
| With partner | 38 | 35.8% |
| Other (family, friends, nursing home) | 15 | 14.2% |
| Diagnosis§ | | |
| AIDS | 69 | 65.1% |
| Symptomatic HIV infection (ARC) | 23 | 21.7% |
| AIDS/Kaposi Sarcoma | 14 | 13.2% |
| Mean time (SD) since diagnosis | 127 days | (114) |
| Range | 1–631 days | |
| CD4+ count | ($n = 89$) | |
| Mean (SD) | 212 | (162) |
| Range | 10–950 | |
| KPS score | ($n = 48$) | |
| Mean (SD) | 92.7 | (7.6) |
| Range | 80–100 | |

KPS=Karnofsky performance status;
IVDU=intravenous drug user
* Revised version of the 1987 Centers for Disease Control surveillance case definition of AIDS.
† Low education corresponds to 0–11 yrs, middle to 12–15 yrs, and high to ≥16 yrs of education.

HOPES at baseline, and 22 minutes at month 6 and 12 (Month 6: $n = 91$, range 10–45 min; Month 12: $n = 80$, range 12–45 min). The wide range of time required to complete the HOPES was due to the fact that some patients took the opportunity, while filling out the questionnaire, to talk about their illness with the research assistant.

At baseline, and at 6 and 12 months follow-up, three, three and five patients, respectively, were unable to complete the questionnaire on their own because of difficulty writing (due to polymyelopathy), or because they were too ill. On these occasions, the questionnaire was administered in the form of an interview.

The system of the screening questions incorporated in the HOPES was generally clear to patients, and they reported few problems in understanding the questions. However, some double-negative question-response combinations caused problems (e.g. item: 'I have pain not controlled by medication'; response: 'not at all'). Further, some items were not relevant to all of the patients (e.g. 'Do you have difficulty driving a car', for patients who either did not have a driver's license or had not driven a car during the past month).

## Descriptive statistics

The majority of the patients scored toward the healthy end of the five summary scales and the global scale (Table 2). The highest baseline score (most problems) was obtained on the sexuality scale. At follow-up, the highest scores were observed on the Physical and Sexual scales.

## Scale structure and reliability

The results of the three multi-trait scaling analyses of the baseline HOPES indicated that 29 of the 106 items (27%) did not meet the standard criterion for item convergent validity (item–scale correlation >0.40); 13 of these items belonged to the Physical scale; six to the Psychosocial scale; seven to the Partner scale; and three to the Sexual scale. In total, there were 240 tests of item discriminant validity. Scaling successes were noted in 82.3% of the tests of item-discriminant validity at baseline. Thirty-four scaling errors (17.7%) were noted, five for items of the Physical scale, eight for items of the Psychosocial scale, 17 for items of the Partner scale, and four for items of the Sexual scale.

The reliability coefficients of the five summary scales were all high, with a mean $\alpha$ coefficient of 0.87 at baseline, and 0.88 at 6 and 12 months (see Table 2).

## Validity

Correlations among the summary scales. The inter-scale correlations at baseline and follow-up were examined (data not shown). In accordance with *a priori* expectations the correlations among the scales were of a moderate magnitude, with the exception

**Table 2.** The HOPES: descriptive statistics and reliability for the summary scales and global score at baseline and follow-up

| No. of Summary scales* | Items | Baseline: ($n = 106$) | | | 6 months: ($n = 93$) | | | 12 months: ($n = 82$) | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Mean | (SD) | α | Mean | (SD) | α | Mean | (SD) | α |
| Physical | 30 | 0.69 | (0.47) | 0.89 | 0.74 | (0.58) | 0.92 | 0.92 | (0.61) | 0.92 |
| Psychosocial | 37† | 0.76 | (0.52) | 0.93 | 0.68 | (0.47) | 0.92 | 0.70 | (0.45) | 0.92 |
| Medical interaction | 11 | 0.36 | (0.55) | 0.89 | 0.28 | (0.39) | 0.84 | 0.25 | (0.40) | 0.86 |
| Sexual§ | 10 | 1.13 | (0.80) | 0.80 | 1.21 | (0.81) | 0.83 | 1.25 | (0.85) | 0.84 |
| Partner§§ | 18 | 0.39 | (0.42) | 0.86 | 0.39 | (0.47) | 0.90 | 0.42 | (0.47) | 0.88 |
| Global‡ | | 0.67 | (0.41) | | 0.65 | (0.39) | | 0.71 | (0.38) | |

SD = standard deviation; α = Cronbach's alpha coefficient
* Higher values indicate more problems
† The subscales 'children' and 'at work concerns' (8 items) are excluded because these questions only applied to a small number of patients.
‡ Dependent on applicability of the screening questions
§ Sample size at baseline: $n = 54$; 6 months: $n = 88$; 12 months: $n = 77$
§§ Sample size at baseline: $n = 54$; 6 months: $n = 50$; 12 months: $n = 45$

**Table 3.** Mean HOPES scores at baseline and 12 months follow-up for patients with AIDS vs. patients with a symptomatic HIV infection (ARC)

| | Baseline | | | | | 12 month follow-up | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | ARC ($n = 23$) | | AIDS ($n = 83$) | | | ARC ($n = 22$) | | AIDS ($n = 60$) | | |
| Summary scale* | Mean | (SD) | Mean | (SD) | $p$ | Mean | (SD) | Mean | (SD) | $p$ |
| Physical | 0.66 | (0.44) | 0.70 | (0.48) | 0.77 | 0.64 | (0.44) | 1.02 | (0.64) | 0.01 |
| Psychosocial | 1.01 | (0.59) | 0.70 | (0.49) | 0.01 | 0.95 | (0.58) | 0.61 | (0.36) | 0.02 |
| Medical interaction | 0.40 | (0.61) | 0.35 | (0.53) | 0.65 | 0.27 | (0.40) | 0.24 | (0.40) | 0.78 |
| Sexual | 1.26[a] | (0.82) | 1.09[b] | (0.79) | 0.42 | 1.18[c] | (0.86) | 1.28[d] | (0.86) | 0.67 |
| Partner | 0.41[e] | (0.35) | 0.38[f] | (0.44) | 0.82 | 0.54[g] | (0.53) | 0.37[h] | (0.45) | 0.30 |
| Global | 0.78 | (0.45) | 0.63 | (0.40) | 0.13 | 0.73 | (0.43) | 0.70 | (0.36) | 0.70 |

* Higher values indicate more problems
[a] $n = 20$; [b] $n = 76$; [c] $n = 20$; [d] $n = 57$; [e] $n = 11$; [f] $n = 47$; [g] $n = 12$ ; [h] $n = 33$

of the sexual and medical interaction scales that were correlated in the range of 0.08–0.23. All summary scales correlated substantially with the global score (range 0.47–0.92). Consistently, the strongest correlate of the global score was the Psychosocial scale (0.79–0.92), followed by the Physical scale (0.75–0.79). Over time, the magnitude of the correlations declined. The mean inter-scale correlations declined from 0.48 at baseline to 0.39 and 0.28 at 6 and 12 months follow-up, respectively.

*Correlations among the HOPES and QLQ-C30.* Consistent with expectations, the correlations between the summary scales of the HOPES and the QLQ-C30 subscales indicated that the conceptually related scales were highly intercorrelated (data not presented in tabular form). For example, the correlation between the subscales assessing physical functioning ranged from 0.49–0.77; those among the scales assessing psychosocial functioning ranged from 0.41–0.68. The subscales pertaining to conceptually distinct domains (e.g., sexuality and cognitive functioning) exhibited weak correlations (<0.10).

*Comparison of AIDS and ARC patients.* The mean scores at baseline and 12 months follow-up of patients with AIDS vs. ARC are shown in Table 3. As expected, patients with AIDS reported significantly fewer psychosocial problems at baseline and 12 months follow-up than patients with ARC. At 6 months follow-up, no significant differences could be detected between the two groups. Contrary to our expectations, the patients with AIDS reported significantly more physical problems at 12 months follow-up only.

*Changes over time.* For the 82 surviving patients, main effects for time were found for the Physical and Psychosocial scales (see Table 4). As expected, patients reported significantly more physical problems at month 12 than at baseline and month 6. Psychosocial functioning had improved significantly at month 6, and remained stable thereafter. Additionally, a trend toward reporting fewer problems in medical interactions was found ($p = 0.06$). No significant change was observed for the global score, which may be explained by the fact that the deterioration in physical functioning had cancelled out the improvement in psychosocial functioning.

*Responsiveness to clinical changes.* Prior to performing the repeated measures ANOVA, statistical tests were performed to confirm whether the increase or decrease in CD4 counts at follow-up was statistically significant. The 19 stable/improved patients evidenced significantly higher CD4 counts at 12 months follow-up [Baseline: Mean = 269 (SD = 176); 12 months: Mean = 339 (SD = 198); $p < 0.001$]. The CD4 count of the 46 deteriorated patients had declined significantly [Baseline: Mean = 230 (SD = 170); 12 months: Mean = 98 (SD = 104); $p < 0.001$].

The repeated measures ANOVA revealed significant differences over time between the patients with

**Table 4.** The HOPES summary and global scores: changes over time for 82 survivors

| | Baseline | | 6 months | | 12 months | | | |
|---|---|---|---|---|---|---|---|---|
| **Summary scales*** | **Mean** | **(SD)** | **Mean** | **(SD)** | **Mean** | **(SD)** | $F^{\dagger}$ | $p$ |
| Physical | 0.68 | (0.48) | 0.69 | (0.54) | 0.92 | (0.61) | 11.99 | 0.00 |
| Psychosocial | 0.83 | (0.55) | 0.70 | (0.48) | 0.70 | (0.45) | 5.80 | 0.00 |
| Medical interaction | 0.37 | (0.56) | 0.29 | (0.41) | 0.25 | (0.40) | 2.86 | 0.06 |
| Sexual ($n = 72$) | 1.08 | (0.72) | 1.14 | (0.81) | 1.24 | (0.81) | 1.95 | 0.15 |
| Partner ($n = 35$) | 0.34 | (0.35) | 0.31 | (0.41) | 0.37 | (0.46) | 0.62 | 0.54 |
| Global | 0.70 | (0.44) | 0.64 | (0.40) | 0.71 | (0.38) | 2.14 | 0.12 |

* Higher values indicate more problems
† F-value and $p$-value reported for main effect for time

**Table 5.** Changes over time in HOPES scores for patients with stable/improved CD4 counts vs. deteriorated CD4 counts

| | | Baseline | | 6 months | | 12 months | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Summary scales*** | $n$ | **Mean** | **(SD)** | **Mean** | **(SD)** | **Mean** | **(SD)** | $F^{\dagger}$ | $p$ |
| Physical | 19†† | 0.71 | (0.57) | 0.59 | (0.46) | 0.63 | (0.50) | 4.62 | 0.01 |
| | 46‡ | 0.68 | (0.44) | 0.71 | (0.57) | 1.01 | (0.65) | | |
| Psychosocial | 19 | 1.01 | (0.58) | 0.81 | (0.52) | 0.80 | (0.50) | 0.90 | 0.41 |
| | 46 | 0.78 | (0.52) | 0.69 | (0.50) | 0.71 | (0.45) | | |
| Medical interaction | 19 | 0.54 | (0.76) | 0.35 | (0.52) | 0.23 | (0.35) | 2.66 | 0.07 |
| | 46 | 0.31 | (0.48) | 0.26 | (0.36) | 0.30 | (0.47) | | |
| Sexual ($n = 72$) | 18 | 1.18 | (0.72) | 1.05 | (0.89) | 0.95 | (0.78) | 3.74 | 0.03 |
| | 41 | 1.11 | (0.71) | 1.16 | (0.77) | 1.38 | (0.81) | | |
| Partner ($n = 35$) | 7 | 0.29 | (0.19) | 0.25 | (0.31) | 0.24 | (0.38) | 0.75 | 0.48 |
| | 27 | 0.34 | (0.38) | 0.24 | (0.23) | 0.38 | (0.47) | | |
| Global | 19 | 0.83 | (0.52) | 0.66 | (0.43) | 0.64 | (0.37) | 5.22 | 0.01 |
| | 46 | 0.67 | (0.40) | 0.63 | (0.41) | 0.76 | (0.40) | | |

* Higher values indicate more problems
† F-value and $p$-value reported for interaction effect (group x time)
†† $n = 19$: stable/improved CD4 counts
‡ $n = 46$: diminished CD4 counts

sustained or improved CD4 counts vs. those with decreased CD4 counts for the Physical and Sexual scales, and the global scores (Table 5). Whereas the scores of the stable/improved patients improved or remained at approximately the same level, patients with decreased CD4 counts reported a deterioration of their physical and sexual functioning, and global QOL over time. No significant differences between the two groups over time were found with respect to psychosocial functioning, problems in their partner relationship or in their interaction with the medical team.

## Discussion

The primary aim of this study was to examine the feasibility and psychometric properties of the HOPES in longitudinal research with HIV-infected and AIDS patients. In a group of predominantly male, bi-/homosexuals with a symptomatic HIV infection or AIDS, the HOPES proved to be a feasible instrument for assessing health-related quality of life. Patients reported few problems in understanding the questions and seldom made mistakes with the screening questions. However, the level of education of our patients was very high and the HOPES was administered under optimal conditions whereby an interviewer was always present to clarify misunderstandings and to guide the respondents through the branching system of the questionnaire.

The HOPES proved to be a highly reliable instrument, with the internal consistency coefficients for the five summary scales being ≥0.80 at all assessments. The multitrait scaling analysis provided moderate support for the hypothesized scale structure of the HOPES. Twenty-seven and 18 per cent of the items did not meet the criteria for item-convergent and item-divergent validity, respectively. In part, this could be explained by the effect of the relatively small sample size, particularly for the Sexual and Partner scale.

The results generally were supportive of the construct validity of the HOPES. The majority of the inter-scale correlations were moderate and in the predicted direction. In addition, the correlations between the summary scales of the HOPES and relevant subscales of the QLQ-C30 were all substantial and in the predicted direction. The HOPES was responsive to changes in CD4 counts over time. A deterioration in immune function was reflected in poorer scores on the HOPES, whereas a stable or improved immune function was accompanied by equal or even better scores on the HOPES.

The HOPES discriminated between subsamples of AIDS and ARC patients, in accordance with *a priori* hypotheses, particularly for the Psychosocial scale. Contrary to our expectations, patients with AIDS only reported significantly poorer physical functioning at 12 months follow-up. This might indicate that patients with AIDS experience a more rapid deterioration in physical functioning than patients with ARC.

Further, the HOPES was able to detect significant changes over time in accordance with hypotheses derived from previous studies. Over time, the surviving patients reported more physical problems and fewer psychosocial problems. When the changes in CD4 counts were also taken into account, deterioration in physical functioning was found to be more prominent among patients with decreased CD4 counts. Improvement in psychosocial functioning did not appear to be related to changes in CD4 counts. In addition, among the surviving patients, no significant deterioration in sexual functioning was found. Upon closer inspection, however, it appeared that patients with decreased CD4 counts reported more sexual problems over time while patients with stable or improved CD4 counts reported fewer sexual problems.

The HOPES provides both a global score and summary scores for specific QOL dimensions. However, prudence is called for in interpreting the global score. We found no change in the global score over time while we did observe simultaneous improvement and deterioration on the summary scales. Moreover, the absolute value of the HOPES global score is dependent on the number of sections that apply to a given patient. Patients for whom all sections apply may obtain a higher global score than those for whom only a subset of sections apply. This scoring artifact could lead one to erroneously conclude that the former patients have a poorer overall QOL.

In conclusion, the results of the study support the feasibility, reliability and validity of the HOPES in assessing the quality of life of patients with AIDS or a symptomatic HIV-infection. The strength of the HOPES is that it allows for a detailed assessment of the daily problems of HIV-infected patients, including topics such as sexuality and the interaction with health care providers that are usually not covered by other QOL instruments. However, the price for this detailed assessment is the 20–30 minutes needed for completion. For use in clinical trials, where time for QOL assessments is usually limited, a short-form (SF) version analogue to the CARES-SF[19] may be required. In future studies, the feasibility of the HOPES in populations with lower educational levels and under less optimal administration conditions needs to be addressed. In addition, the scale structure of the

HOPES requires closer scrutiny in larger samples. To further examine the construct validity of the HOPES, studies in healthier, asymptomatic as well as in late stage disease populations are needed.

## Acknowledgements

The authors wish to thank all patients for their co-operation and Elly Visser-Pol, Wytske Kylstra and Wilmy Cleijne for their help in interviewing patients. Further, we extend special thanks to the following physicians and nurses for recruiting patients onto the study: Sven Danner, Jan-Karel Eeftinck Schattenkerke, Jos Frissen, Peter Reiss and Theo de Graaff from the Department of Internal Medicine/AIDS Unit, Academic Medical Center, Amsterdam; Hugo Weigel and Edwin de Vries from the Onze Lieve Vrouwe Gasthuis, Amsterdam; Pieter Meenhorst, Jeanine te Nijenhuis and Liliane van Belle from the Slotervaart Hospital, Amsterdam; Cissy van den Berg from the Prinsengracht Hospital, Amsterdam; Reinier ten Kate from the Vrije Universiteit Academic Hospital, Amsterdam; Gerrit Schreij from the Academic Hospital Maastricht; Ineke van den Ende from Dijkzigt Academic Hospital Rotterdam; and Robert Kauffmann from the Leyenburg Hospital, The Hague.

### References

1. Ganz PA, Coscarelli Schag CA, Kahn B, Petersen L, Hirji K. 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* 1993; **2**: 109–119.
2. Schag CAC, Ganz PA, Kahn B, Petersen L. 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* 1992; **1**: 397–413.
3. Stewart AL, Hays RD, Ware JE. The MOS short-form general health survey: reliability and validity in a patient population. *Med Care* 1988; **26**: 724–735.
4. Wu AW, Rubin HR, Mathews WC, *et al.* A Health Status Questionnaire using 30 items from the Medical Outcomes Study; a preliminary validation in persons with early HIV Infection. *Med Care* 1991; **29**: 786–798.
5. McNair DM, Lorr M, Droppelman LF. *Manual for the Profile of Mood States.* San Diego: Educational and Industrial Testing Service, 1981.
6. Rahe RH. Epidemiological studies of life change and illness. *Int J Psychiatry Med* 1975; **6**: 133–146.
7. Bloom JR, Gorsky RD, Fobair P, *et al.* Physical performance at work and at leisure: validation of a measure of biological energy in survivors of Hodgkin's disease. *J Psychosoc Oncol* 1990; **8**: 49–63.
8. Frissen PHJ, van der Ende ME, Ten Napel CH, *et al.* Zidovudine and Interferon-alpha combination therapy vs. zidovudine monotherapy in subjects with symptomatic human immunodeficiency virus type I infection. *J Infect Dis* 1994; **169**: 1351–1355.
9. Karnofsky DA, Abelman WH, Craver LF, Burcheneal JH. The use of nitrogen mustards in the palliative treatment of carcinoma. *Cancer* 1948; **1**: 634–656.
10. Aaronson NK, Ahmedzai S, Bergman B, *et al.* The EORTC QLQ-C30: a quality of life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993; **85**: 365–376.
11. De Boer JB, Sprangers MAG, Aaronson NK, Lange JMA, van Dam FSAM. The feasibility, reliability and validity of the EORTC QLQ-C30 in assessing the quality of life of patients with a symptomatic HIV infection or AIDS (CDC IV). *Psych Health* 1994; **9**: 65–77.
12. Hays RD, Hayashi T, Carson S, Ware JE. *User's guide for the Multitrait Analysis Program (MAP).* Santa Monica, CA: Rand Corporation, 1988.
13. Nunnally JC. *Psychometric theory, 2nd ed.* New York: McGraw–Hill, 1978.
14. King MB. Psychosocial status of 192 outpatients with HIV infection and AIDS. *Br J Psychiatry* 1989; **154**: 237–242.
15. Lubeck DP, Fries JF. Changes in quality of life among persons with HIV infection. *Qual Life Res* 1992; **1**: 359–366.
16. Lubeck DP, Fries JF. Health status among persons infected with human immunodeficiency virus. *Med Care* 1993; **31**: 269–276.
17. Chuang HT, Devins GM, Hunsley J, Gill MJ. Psychosocial distress and well-being among gay and bisexual men with Human Immunodeficiency Virus Infection. *Am J Psychiatry* 1989; **146**: 876–880.
18. De Gruttola V, Wulfsohn W, Fischl MA, *et al.* Modeling the relationship between survival and CD4 lymphocytes in patients with AIDS and AIDS-related complex. *J Acquir Immune Defic Syndr* 1993; **6**: 359–36.
19. Coscarelli Schag CA, Ganz PA, Heinrich RL. Cancer Rehabilitation Evaluation System-Short Form (CARES-SF): a cancer-specific rehabilitation and quality of life instrument. *Cancer* 1991; **68**: 1406–1413.

*(Received 11 August 1995; accepted 9 September 1995)*