

Development and Testing of the Parkinson's Disease Quality of Life Scale

Mickie Welsh, RN, DNSc,¹ Michael P. McDermott, PhD,² Robert G. Holloway, MD, MPH,^{3*} Sandy Plumb, BS,³ Ronald Pfeiffer, MD,⁴ Jean Hubble, MD,⁵ and The Parkinson Study Group

¹*Department of Neurology, University of Southern California, Los Angeles, California, USA*

²*Departments of Biostatistics and Neurology, University of Rochester, Rochester, New York, USA*

³*Department of Neurology, University of Rochester, Rochester, New York, USA*

⁴*Department of Neurology, University of Tennessee, Tennessee USA*

⁵*Department of Neurology, Ohio State University, Ohio, USA*

Abstract: We report on the development and results of preliminary psychometric testing of a disease specific health-related quality of life (HRQoL) scale intended for use in individuals diagnosed with idiopathic Parkinson's disease (PD). Results from an initial qualitative study provided content for item development and scale construction of the Parkinson's disease quality of life scale (PDQUALIF). The 33-item instrument includes seven domains: social/role function, self-image/sexuality, sleep, outlook, physical function, independence, and urinary function, plus one item of Global HRQoL. Initial psy-

chometric testing of the instrument was conducted in 233 outpatient clinic attendees with physician-confirmed idiopathic PD. Factor structure, reliability and validity of the scale have been established in this cross-sectional study. Continuing development of the PDQUALIF will be directed at enhancing the psychometric properties, establishing responsiveness and determining appropriateness in culturally diverse samples. © 2003 Movement Disorder Society

Key words: health-related quality of life; Parkinson's disease

Parkinson's disease (PD), a chronic, degenerative movement disorder, is estimated to affect one million people in the United States.¹ Disability in PD results from progressive motor impairment, nonmotor complications, and treatment side effects.² Despite the availability of medications to treat symptoms and the emergence of surgical treatment options, a cure remains elusive and the protracted course of the disease continues to have major impact on the quality of life of patients and their families.³

A wide variety of outcome measures are used in PD therapeutics research. These range from disease staging scales (e.g., Hoehn and Yahr Scale), symptom-specific measures (e.g., dyskinesia scales), disability scales (e.g.,

Schwab and England scale), disability/impairment scales (e.g., Unified Parkinson's Disease Rating Scale), and health-related quality of life (HRQoL) instruments. To date, four PD-specific HRQoL instruments have been developed.⁴ Studies have shown that these PD specific HRQoL instruments correlate with disease severity, disability, somnolence, and depressive symptoms emphasizing their overall importance in measuring the impact that PD and its treatment has on patients.^{5–7} Limited data have yet emerged on the responsiveness of these scales to changes in the natural history of disease or to the use of treatment. In conjunction with the Parkinson Study Group, we sought to develop an instrument that could be ultimately used in multicenter clinical trials of various symptomatic, neuroprotective and surgical treatment options, and would compliment clinical measures already in use.

We present the development and initial psychometric testing results of a PD-specific HRQoL instrument, the Parkinson's Disease Quality of Life Scale: the PDQUALIF.

*Correspondence to: Robert G. Holloway, MD, MPH, University of Rochester, Department of Neurology, 1351 Mt. Hope Avenue, Ste 220, Rochester, NY 14620. E-mail: bholloway@mct.rochester.edu

Received 3 July 2002; Revised 28 September 2002; Accepted 13 November 2002

PATIENTS AND METHODS

Concept Clarification, Item Generation, Scale Construction

Fifty-two PD patients, 28 spouses or significant others, and 6 PD professionals answered an open-ended question: how does Parkinson's disease change or affect one's life quality? Using content analytic techniques, the qualitative data from this question were reduced to 73 indicators of quality of life.⁸ We categorized the indicators into one of the following domains: general health, personal functioning, psychological well being/distress and social and role functioning.⁹ Each indicator was then developed into a concept statement to be used for scale construction. Three movement disorder professionals rated the concept statements on a scale of 0 to 3 for relevance, clarity and conciseness.¹⁰ Concept statement total scores (relevance + clarity + conciseness) were averaged across judges and rank ordered. The 32 highest scoring indicators of quality of life were selected for inclusion in the instrument. An additional item rating global quality of life was added to complete the instrument entitled the PDQUALIF (the Parkinson's Disease Quality of Life Scale). Items were scaled using a 5-point Likert-style response format. The wording structure of 11 items was reversed.¹¹ Ten patients and 6 PD professionals reviewed the scale for face validity and editorial critique before psychometric testing. The instrument is shown in Figure 1.

Subject Selection

A convenience sample of 233 individuals was selected from 13 outpatient movement disorder clinics in the United States and Canada from July 1995 to June 1996. Subject informed consent was obtained. Subjects were limited to English speaking individuals without cognitive impairment whose diagnosis had been confirmed by the site movement disorder physician. The absence of cognitive impairment was left to the clinical judgment of the enrolling investigator. Fifty-eight of the subjects from 5 of 13 study sites provided additional consent to complete the PDQUALIF a second time, 2 weeks after the initial testing, to provide data for test-retest reliability.

Study Procedures

Each subject underwent an evaluation to determine disease stage using the Hoehn and Yahr (H&Y) Scale,¹² and impairments and disability using the Unified Parkinson Disease Rating Scale (UPDRS).¹³ The H&Y scale has been used in many observational studies and clinical trials and has excellent inter-observer agreement.¹⁴ The UPDRS demonstrates high internal consistency, shows

moderate construct validity, and has a stable factor structure.^{15,16}

Subjects then completed a self-administered instrument packet that included a demographic history, a current medication profile, the PDQUALIF, and two generic health profile questionnaires: the Medical Outcomes Study Short Form (SF-36)¹⁷ and the Sickness Impact Profile (SIP).¹⁸ The SF-36 is a widely used profile measure of generic HRQoL.¹⁷ The scale has shown good reliability and validity and has been used in hundreds of studies of chronic medical conditions. The SF-36 has been used in patients with PD and supplements information not found in the UPDRS.¹⁹ The SIP is a behaviorally based 136-item health status measure.¹⁸ It is a valid and reliable health status measure that has been used repeatedly in clinical trials and epidemiological studies.

Analysis

An exploratory principal component factor analysis, including varimax rotation to enhance interpretation of the resulting factors, was used to help assign the 32 PDQUALIF items to subscales. Multi-trait scaling analyses were used to examine the degree of convergence of items within a subscale and the discrimination of items between subscales.²⁰

Subscale scores were created by transforming raw scores to a 0 to 100 measurement scale by summing the items within a subscale, dividing by the maximum possible total score for that subscale, and multiplying by 100, with lower scores indicating a better quality of life. A total score for the scale was computed by taking an unweighted average of the individual subscale scores. Distributions of scale scores were examined using mean scores, standard deviations, ranges and percentages of respondents having minimum (floor) and maximum (ceiling) scores.

Internal consistency was estimated for the total scale and for each subscale using Cronbach's α coefficient.²¹ The test-retest reliability of subscale scores was estimated using the intraclass correlation coefficient (ICC) derived from a random effects one-way analysis of variance model.²² Test-retest reliability for the individual PDQUALIF items was estimated using the weighted κ statistic.²³

Convergent and discriminant validity of the PDQUALIF were assessed by examining the associations among the following variables: total and subscale scores of the PDQUALIF, SF-36, SIP and UPDRS, age, education, marital status, gender, employment status and years since the diagnosis of Parkinson's disease. Spearman rank correlation coefficients and *t*-tests were used for this purpose.

PDQUALIF[®]
M Welsh and the Parkinson Study Group[™], 1996.

PLEASE MAKE A CHECK MARK IN THE BOX BELOW THE ANSWER WHICH BEST DESCRIBES YOUR PERSONAL SITUATION.

EXAMPLE: I exercise regularly:

never *rarely* *sometimes* *occasionally* *always*

1. Changing position causes me to become lightheaded (lying to standing, or sitting to standing):

never *rarely* *sometimes* *occasionally* *always*

2. When walking, I have trouble keeping my balance:

never *rarely* *sometimes* *occasionally* *always*

3. When eating or drinking liquids, I have difficulty swallowing:

never *rarely* *sometimes* *frequently* *always*

4. My Parkinson's symptoms affect my ability to communicate with people:

never *rarely* *sometimes* *frequently* *always*

5. The need to go to the bathroom wakes me in the night:

never *rarely* *sometimes* *frequently* *always*

6. My Parkinson's symptoms affect my ability to show affection in intimate or sexual ways:

never *rarely* *sometimes* *frequently* *always*

7. I have aching/burning/coldness/numbness in my hand/feet:

never *rarely* *sometimes* *frequently* *always*

8. I have difficulty with bladder control (frequency, urgency, inability):

never *rarely* *sometimes* *frequently* *always*

9. Constipation is a problem:

never *rarely* *sometimes* *frequently* *always*

10. My Parkinson's symptoms cause me to have trouble falling asleep, or waking early:

never *rarely* *sometimes* *frequently* *always*

11. I have trouble staying asleep:

never *rarely* *sometimes* *frequently* *always*

12. My Parkinson's symptoms make it hard to maintain a positive outlook:

always *frequently* *sometimes* *rarely* *never*

13. My Parkinson's symptoms cause me to feel like a burden to other people:

never *rarely* *sometimes* *frequently* *always*

14. My Parkinson's symptoms have affected my social life:

never *rarely* *sometimes* *frequently* *always*

15. I worry about what the future has in store:

always *frequently* *sometimes* *rarely* *never*

16. Asking others for help is difficult for me:

always *frequently* *sometimes* *rarely* *never*

17. Maintaining my independence is important to me:

never *rarely* *sometimes* *frequently* *always*

18. It has been difficult to adjust to the changes which have taken place in my body:

very *moderately* *somewhat* *slightly* *not at all*

19. My Parkinson's symptoms have not affected my social life:

strongly agree *somewhat agree* *agree* *somewhat disagree* *strongly disagree*

20. Travel remains an important part of my leisure activities:

strongly agree *somewhat agree* *agree* *somewhat disagree* *strongly disagree*

21. My Parkinson's symptoms have affected my family role and relationship:

strongly agree *somewhat agree* *agree* *somewhat disagree* *strongly disagree*

22. My Parkinson's symptoms cause me to stay away from social gatherings:

strongly agree *somewhat agree* *agree* *somewhat disagree* *strongly disagree*

23. My spouse/children/friends' view of me has changed because of my illness:

strongly agree *somewhat agree* *agree* *somewhat disagree* *strongly disagree*

24. I feel I am less sexually desirable because of my illness:

strongly agree *somewhat agree* *agree* *somewhat disagree* *strongly disagree*

IN THE PAST 7 DAYS:

25. In my personal hygiene (bathing, hair care, make up, shaving, or toileting) I have been independent:

everyday *5-6 days* *3-4 days* *1-2 days* *never*

26. In food preparation or eating I am independent:

everyday *5-6 days* *3-4 days* *1-2 days* *never*

27. Written or spoken communication is a problem for me:

never *rarely* *sometimes* *frequently* *always*

28. Fatigue makes participation in activities, household chores, shopping or yard work a problem for me:

always *frequently* *sometimes* *rarely* *never*

29. My Parkinson's symptoms interfere with my ability to do my usual share in the home:

always *frequently* *sometimes* *rarely* *never*

30. My nighttime symptoms keep me from sleeping with my spouse/partner:

everyday *5-6 days* *3-4 days* *1-2 days* *never*

31. My Parkinson's symptoms have interfered with my driving ability:

doesn't apply *rarely* *sometimes* *frequently* *constantly*

32. My illness has caused a financial strain for me and my family:

doesn't apply *rarely a concern* *sometimes* *frequently* *constantly*

33. Compared to 6 months ago, my Parkinson's symptoms are:

much better *somewhat better* *about the same* *somewhat worse* *much worse*

FIG. 1. PDQUALIF. ©M. Welsh and the Parkinson Study Group[™], 1996.

Analysis of variance was also used to compare mean PDQUALIF, SF-36 and SIP scores across Hoehn and Yahr stages. The ratios of the resulting F-statistics for different variables provide a basis for evaluating the relative sensitivity (validity) of individual scales to these known disease stage differences.²⁴ The Tukey–Kramer method was used to adjust for multiple pairwise group comparisons.²⁵

All available data were used in all analyses and no imputation was carried out for missing data. Factor analyses, multitrait scaling analyses, and analyses of internal consistency involved the 222 subjects (95%) for whom no PDQUALIF items were missing. Analyses of convergent and discriminant validity involved up to 233 subjects, depending on the extent of missing data. All analyses were conducted using SAS software.

RESULTS

Subject Characteristics

Subject characteristics for the total sample and the subset of subjects participating in the test–retest reliability study are shown in Table 1. The mean age of the total sample was 65.4 (SD 10.3), 62% were male, 98% were Caucasian, and the average education level was 14.7 years (SD 3.3). Most were married (65%) and retired or unable to work (80%). The majority of subjects (82%)

were in mid-stage disease, either H&Y stages 2.0, 2.5, or 3.0, as reflected in the average years since PD diagnosis (7.2, SD 5.3). Only 14% of subjects were in H&Y stages 1.0 or 1.5 and even fewer were in stages 4.0 (42%) and 5.0 (0%). The characteristics of the subjects participating in the reliability study were similar to the total sample.

Feasibility of Administration

Completion of the entire questionnaire packet took approximately 35 to 40 minutes. Completion of the PDQUALIF took approximately 10 to 15 minutes.

Subscale Construction and Reliability

Factor analysis with varimax rotation supported a seven-factor solution (Table 2). Three of the original four domains used as the conceptual framework during the qualitative data collection and four additional domains emerged. The seven factors were identified as: 1) Social and Role Function (nine items); 2) Self Image and Sexuality (seven items); 3) Sleep (three items); 4) Outlook (four items); 5) Physical Functioning (five items); 6) Independence (two items); and 7) Urinary Function (two items). In the factor analysis, eight items had factor loadings >0.35 on more than one factor. None of the secondary factor loadings were greater than or equal to the primary factor loading. The Physical Functioning factor included one item (Item 31, driving ability) not expected to load with the other items in that subscale. Item 2, Imbalance, had been conceptualized as a part of the Physical Functioning factor; however, it loaded primarily on the Social and Role Function factor. Four items loaded on both the Social/Role Function and Self Image/Sexuality factors. The seven identified factors accounted for 55.6% of the total scale variance.

Item convergence and discrimination for the PDQUALIF were found generally satisfactory (Table 3). Only eight correlations between an item and its own subscale (corrected for item overlap) failed to reach the 0.40 threshold.²⁰ In five of these cases, the item–subscale correlation was above 0.30. The most problematic subscale in this regard was the Physical Function subscale. Item discrimination was well supported for all subscales.

Descriptive statistics for the PDQUALIF subscales are presented in Table 4. Subscales demonstrated reasonable score variability with the exception of the Independence subscale, which had a noticeable floor effect (78.9%), and the Sleep subscale, which demonstrated minor floor effects (8.7%). Table 4 also shows the internal consistency reliability estimates (Cronbach's α) for the total PDQUALIF scale and its seven subscales. The internal consistency estimate for the total scale was 0.89, and the range for the seven subscales was 0.55 to 0.85. The

TABLE 1. Subject characteristics

| | Total sample (n = 233) | Reliability subset* (n = 58) |
|-------------------------------|---------------------------|---------------------------------|
| Age (yr) | 65.4 (10.3) | 64.4 (8.8) |
| Male (%) | 62.2 | 65.5 |
| Caucasian (%) | 97.8 | 100.0 |
| Years since diagnosis | 7.2 (5.3) | 6.9 (5.4) |
| Education (yr) | 14.7 (3.3) | 15.1 (3.3) |
| Married (%) | 65.4 | 65.5 |
| Retired or unable to work (%) | 79.8 | 87.7 |
| H&Y Stage (%) | | |
| 1.0 | 5.6 | 5.3 |
| 1.5 | 8.6 | 14.0 |
| 2.0 | 44.8 | 50.9 |
| 2.5 | 16.8 | 15.8 |
| 3.0 | 20.3 | 10.5 |
| 4.0 | 3.9 | 3.5 |
| 5.0 | 0.0 | 0.0 |
| UPDRS | | |
| Total | 35.1 (17.8) | 30.3 (16.4) |
| Part I: mentation | 1.8 (1.7) | 1.5 (1.6) |
| Part II: ADL | 11.3 (6.5) | 10.7 (6.1) |
| Part III: motor | 22.2 (12.6) | 18.7 (12.0) |
| Sinemet use (%) | 65.5 | 69.1 |
| Motor fluctuations (%) | 58.8 | 57.1 |

Values are expressed as mean (SD) unless otherwise indicated.

*Subset of subjects participating in the test–retest reliability study.

H&Y, Hoehn and Yahr; UPDRS, Unified Parkinson's Disease Rating Scale; ADL, activities of daily living.

TABLE 2. Principal component factor analysis with seven factors

| Item name (item no.) | Subscale labels | | | | | | |
|-----------------------------------|----------------------|----------------------|-------|---------|-------------------|--------------|------------------|
| | Social/role function | Self image/sexuality | Sleep | Outlook | Physical function | Independence | Urinary function |
| Social/role function | | | | | | | |
| Usual share of work in home (29) | 0.66 | | | | | | |
| Social life (19) | 0.64 | | | | | | |
| Imbalance (2) | 0.64 | | | | (0.36) | | |
| Social life (14) | 0.62 | (0.40) | | | | | |
| Burden (13) | 0.58 | | | | | | |
| Fatigue (28) | 0.56 | | | | | | |
| Social isolation (22) | 0.56 | (0.43) | | | | | |
| Travel (20) | 0.54 | | | | | | |
| Adjust to change (18) | 0.53 | | | (0.50) | | | |
| Self image/sexuality | | | | | | | |
| Self concept/image (23) | | 0.79 | | | | | |
| Sexual ability (6) | | 0.68 | | | | | (0.37) |
| Family relationship (21) | | 0.62 | | | | | |
| Sexual desirability (24) | | 0.60 | | | | | |
| Communication (4) | (0.41) | 0.51 | | | | | |
| Communication (27) | (0.36) | 0.46 | | | | | |
| Financial strain (32) | | 0.44 | | | | | |
| Sleep | | | | | | | |
| Sleep initiation (10) | | | 0.85 | | | | |
| Sleep maintenance (11) | | | 0.81 | | | | |
| Sleep with partner (30) | | | 0.43 | | | | |
| Outlook | | | | | | | |
| Future (15) | | | | 0.64 | | | |
| Maintain independence (17) | | | | 0.64 | | | |
| Ask for help (16) | | | | 0.60 | | | |
| Outlook (12) | (0.41) | | | 0.51 | | | |
| Physical function | | | | | | | |
| Neuropathy (7) | | | | | 0.68 | | |
| Dizziness (1) | | | | | 0.68 | | |
| Swallow (3) | | | | | 0.48 | | |
| Driving (31) | | | | | 0.45 | | |
| Constipation (9) | | | | | 0.36 | | |
| Independence | | | | | | | |
| Independent hygiene (25) | | | | | | 0.82 | |
| Independent food preparation (26) | | | | | | 0.75 | |
| Urinary function | | | | | | | |
| Nocturia (5) | | | | | | | 0.77 |
| Uro-frequency (8) | | | | | | | 0.72 |

Factor loadings <0.35 omitted for clarity.

test-retest reliability estimates (intraclass correlation coefficients) for the total score and the seven subscales are also shown in Table 4 and were acceptable, ranging from 0.68 to 0.88. In addition, 28 of 32 individual scale items showed acceptable agreement (weighted $\kappa > 0.42$). The four items with weighted $\kappa < 0.40$ were Item 23: self concept (0.38), Item 17: remaining independent (0.34), Item 26: independent function in eating and food preparation (0.30), and Item 24: sexual desirability (0.20).

PDQUALIF and UPDRS, Generic HRQoL, and Other Measures

The correlations between the PDQUALIF total and subscale scores and UPDRS scores are provided in Table 5. The ADL scores from the UPDRS were associated

moderately with scores on the Social/Role Function, Self-Image/Sexuality, Physical Function and Independence subscales, as well as the total score ($r = 0.37-0.52$). The UPDRS Motor and Mental subscale scores were less strongly associated with the PDQUALIF scores. The associations between PDQUALIF scores and UPDRS scores were all highly statistically significant, with the exception of those involving the PDQUALIF Sleep and Outlook subscales. Mean PDQUALIF total scores also differed for individuals with motor fluctuations (45.2 ± 11.7) compared to those without fluctuations ($35.1 \pm 13.4, P < 0.0001$).

The associations between the PDQUALIF total and subscale scores and those of the SF-36 and SIP were all

TABLE 3. PDQUALIF item convergence and discrimination

| Subscale | Items (n) | Item convergent correlations ^a | Item convergence ^b | Item discriminant correlations ^c | Item discrimination ^d |
|--------------------------|-----------|---|-------------------------------|---|----------------------------------|
| Social/role/function | 9 | 0.36–0.74 | 8/9 | 0.05–0.62 | 53/54 |
| Self-image and sexuality | 7 | 0.40–0.58 | 7/7 | 0.02–0.54 | 42/42 |
| Sleep | 3 | 0.21–0.55 | 2/3 | 0.05–0.32 | 18/18 |
| Outlook | 4 | 0.15–0.48 | 2/4 | 0.13–0.58 | 23/24 |
| Physical function | 5 | 0.23–0.40 | 1/5 | 0.07–0.37 | 28/30 |
| Independence | 2 | 0.57 | 2/2 | 0.16–0.37 | 12/12 |
| Urinary function | 2 | 0.44 | 2/2 | 0.11–0.30 | 12/12 |

^aCorrelations between each item score and the score of the subscale to which that item belongs (corrected for overlap).

^bNumber of item-subscale correlations ≥ 0.40 /total number of correlations.

^cCorrelations between each item score and the score of the subscale to which the item does not belong.

^dNumber of cases in which an item correlated higher with its own subscale (corrected for overlap) than with another subscale/total number of correlations.

highly statistically significant; these associations tended to be stronger with the SIP than with the SF-36 (Table 5). Unlike the PDQUALIF and the SIP, however, higher SF-36 scores represent better HRQoL. Therefore, the correlations observed between the SF-36 and the PDQUALIF total and subscale scores were negative.

The PDQUALIF total score was associated with the number of years since PD diagnosis ($r = 0.38$, $P < 0.0001$), but not with age ($r = 0.02$, $P = 0.80$) or years of education ($r = -0.14$, $P = 0.04$); the same was true of most of the PDQUALIF subscales (Table 5). Mean PDQUALIF total scores did not differ significantly by gender or marital status; however, they did differ by whether patients were working (35.1 + 12.2), retired (42.0 + 13.7) or unable to work (48.8 + 11.1). All pairwise comparisons were statistically significant after adjustment for multiple comparisons.

PDQUALIF and Stage of Disease

Table 6 shows the mean scores on the PDQUALIF, the SF-36, and the SIP based on H&Y stage of disease. The mean scores for all subscales generally indicated worse quality of life with more advanced disease. The F -statistic was highest for the SIP Physical dimension score

(25.2), followed by the Social/Role Function subscale (11.6), the PDQUALIF total score (10.8), and the SF-36 Physical Component summary score (9.1).

DISCUSSION

The PDQUALIF provides an assessment of HRQoL as defined by the Parkinson's patients, caregivers and PD professionals who participated in this study. Questionnaire items reflect areas that were identified as important by patients suffering from the disease and, therefore, should be reflective of how the illness affects their lives.²⁶ The methods used for development of the PDQUALIF were based largely on methodology for the development of evaluative instruments in terms of item selection, questionnaire format, pretesting reproducibility, and validity.²⁷ The hypothesized conceptual framework for HRQoL in PD was supported partially by the results of the factor analysis, although additional factors emerged that were not apparent previously. Less emphasis on general health and a greater focus on self-concept and self-image were revealed.

Eleven of 12 areas of HRQoL identified as relevant to PD are included in the PDQUALIF.²⁸ The area not assessed directly in the PDQUALIF is PD-specific motor

TABLE 4. Descriptive statistics, internal consistency, and test-retest reliability for PDQUALIF

| Scale | n | Mean | SD | Scoring, % | | Cronbach's α | Intraclass correlation |
|----------------------|-----|------|------|------------|-----------|---------------------|------------------------|
| | | | | Min = 0 | Max = 100 | | |
| Social/role function | 231 | 47.7 | 20.4 | 0.9 | 0.4 | 0.85 | 0.86 |
| Self image/sexuality | 228 | 42.3 | 21.2 | 0.4 | 0 | 0.79 | 0.82 |
| Sleep | 229 | 41.2 | 26.2 | 8.7 | 3.1 | 0.59 | 0.70 |
| Outlook | 233 | 58.2 | 16.5 | 0.4 | 1.3 | 0.58 | 0.68 |
| Physical function | 232 | 36.5 | 17.9 | 3.0 | 0 | 0.55 | 0.86 |
| Independence | 232 | 9.4 | 22.3 | 78.9 | 3.0 | 0.72 | 0.69 |
| Urinary function | 233 | 52.8 | 23.7 | 3.0 | 2.6 | 0.62 | 0.85 |
| Overall scale score | 222 | 41.3 | 13.7 | 0 | 0 | 0.89 | 0.88 |

TABLE 5. Correlation of PDQUALIF with generic HRQoL, UPDRS, and other measures

| | Total score | Social/role function | Self-image/sexuality | Sleep | Outlook | Physical function | Independence | Urinary function |
|----------------------------|--------------------|----------------------|----------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| UPDRS | | | | | | | | |
| Total | 0.43 ^a | 0.39 ^a | 0.32 ^a | 0.14 ^b | 0.03 | 0.32 ^a | 0.46 ^a | 0.26 ^a |
| Mental | 0.39 ^a | 0.28 ^a | 0.37 ^a | 0.10 | 0.16 ^b | 0.30 ^a | 0.28 ^a | 0.22 ^a |
| Motor | 0.31 ^a | 0.27 ^a | 0.18 ^a | 0.06 | 0.02 | 0.23 ^a | 0.37 ^a | 0.16 ^b |
| ADL | 0.55 ^a | 0.50 ^a | 0.44 ^a | 0.26 ^a | 0.03 | 0.40 ^a | 0.47 ^a | 0.33 ^a |
| SF-36 health survey | | | | | | | | |
| Mental component summary | -0.48 ^a | -0.44 ^a | -0.41 ^a | -0.21 ^a | -0.49 ^a | -0.31 ^a | -0.19 ^a | -0.17 ^b |
| Physical component summary | -0.52 ^a | -0.60 ^a | -0.34 ^a | -0.24 ^a | -0.21 ^a | -0.37 ^a | -0.26 ^a | -0.29 ^a |
| Sickness impact profile | | | | | | | | |
| Total | 0.73 ^a | 0.68 ^a | 0.52 ^a | 0.29 ^a | 0.35 ^a | 0.57 ^a | 0.55 ^a | 0.29 ^a |
| Physical dimension | 0.61 ^a | 0.59 ^a | 0.36 ^a | 0.25 ^a | 0.22 ^a | 0.48 ^a | 0.55 ^a | 0.28 ^a |
| Psychosocial dimension | 0.70 ^a | 0.62 ^a | 0.61 ^a | 0.27 ^a | 0.43 ^a | 0.52 ^a | 0.42 ^a | 0.31 ^a |
| Age (yr) | 0.03 | 0.03 | -0.22 ^a | 0.02 | -0.21 ^a | 0.10 | 0.11 | 0.21 ^a |
| Education (yr) | -0.15 ^b | -0.09 | 0.03 | -0.12 | -0.12 | -0.09 | -0.17 ^a | -0.09 |
| Years since diagnosis | 0.32 ^a | 0.35 ^a | 0.32 ^a | 0.18 ^a | 0.08 | 0.20 ^a | 0.19 ^a | 0.07 |

^a*P* < 0.01; ^b*P* < 0.05.

UPDRS, Unified Parkinson's Disease Rating Scale.

symptoms. These were not endorsed strongly by participants in the qualitative study, and the instrument is intended as a compliment to the UPDRS, which provides a comprehensive assessment of motor symptomatology. On the contrary, the PDQUALIF emphasizes many non-motor symptoms including fatigue, sleep, autonomic dysfunction, and sexual function. These symptoms were endorsed in the qualitative study and found important to patients in other studies.²⁹⁻³²

The Cronbach's α for the total scale was 0.89. Four of the seven PDQUALIF subscales had Cronbach's α estimates of less than 0.70 (0.55-0.62), which may have resulted from the few numbers of items in the subscales.

Test-retest reliability was shown to be good for all of the subscales. Support for item convergence and discrimination was provided by the multitrait scaling analysis.

The low internal consistency estimate (0.55) for the Physical Function subscale may be due to two possibilities. First, our sample included a preponderance of patients in early and mid-stages of disease, and some of the Physical Function items (balance, swallowing, and numbness and tingling) address symptoms that may be more characteristic of individuals with later stage disease. Second, it may reflect the varied presentation and course of individual symptoms in any given patient. It is likely also that the large floor effect of the Independence

TABLE 6. Comparison of mean health-related quality of life scores across Hoehn and Yahr stages

| | Hoehn and Yahr Stage | | | | | | <i>F</i> * | Relative validity** |
|----------------------------|----------------------|-----------------|----------------|-----------------|---------------|--------------|----------------|---------------------|
| | 1 (n = 13) | 1.5 (n = 20) | 2 (n = 104) | 2.5 (n = 39) | 3 (n = 47) | 4 (n = 9) | | |
| PDQUALIF | | | | | | | | |
| Total | 27.4 | 37.7 | 38.8 | 40.9 | 48.7 | 58.5 | 10.8 | 1.00 |
| Social/role function | 31.0 | 41.1 | 44.5 | 45.0 | 57.5 | 79.6 | 11.6 | 1.07 |
| Self image/sexuality | 27.4 | 39.1 | 41.1 | 40.9 | 47.4 | 63.1 | 3.9 | 0.36 |
| Sleep | 23.1 | 40.8 | 37.9 | 43.6 | 48.0 | 58.3 | 3.2 | 0.29 |
| Outlook | 53.9 | 58.1 | 57.3 | 57.9 | 59.8 | 62.5 | 0.4 | 0.04 |
| Physical function | 18.1 | 35.3 | 34.0 | 36.2 | 46.6 | 44.4 | 7.4 | 0.68 |
| Independence | 00.0 | 0.7 | 5.3 | 9.0 | 18.1 | 44.4 | — ^a | — |
| Urinary function | 33.7 | 49.4 | 50.6 | 53.5 | 63.0 | 56.9 | 4.0 | 0.37 |
| SF-36 health survey | | | | | | | | |
| Mental component summary | 51.5 | 49.3 | 48.1 | 47.6 | 44.3 | 42.8 | 1.9 | 0.18 |
| Physical component summary | 47.7 | 40.7 | 40.5 | 37.3 | 33.7 | 27.8 | 9.1 | 0.84 |
| Sickness impact profile | | | | | | | | |
| Psychosocial dimension | 6.9 | 11.6 | 12.0 | 17.9 | 20.3 | 32.7 | 7.7 | 0.71 |
| Physical dimension | 3.3 | 7.1 | 6.3 | 11.0 | 20.8 | 31.6 | 25.2 | 2.32 |

**F*-statistic for testing equality of mean scores across Hoehn and Yahr stages.

**Ratio of the *F*-statistic for the given variable to that for the PDQUALIF total score.

^aThe scores on the physical function subscale were highly non-normally distributed; therefore, the *F* test was not performed.

subscale reflects the many individuals in this sample with early or milder PD, and relatively few with more advanced disease and impairments in maintaining personal hygiene and in preparing foods. Further analysis will be necessary to determine whether certain portions of the instrument are more useful for studies of early versus later-stage PD patients, or if certain items need to be eliminated or redefined when applied to particular subsets of patients (e.g., de novo patients)

Moderate to strong support for construct validity was evident using generic HRQoL scales (SIP and the SF-36), disease-specific instruments (UPDRS), stage of disease (H&Y stage), and other measures. The PDQUALIF total score was correlated strongly with the SIP total and subscale scores, the SF-36 physical component summary score, and the UPDRS ADL score. The PDQUALIF total score was correlated moderately with the total and mental UPDRS scores, the SF-36 mental component summary score, and years since diagnosis, but was correlated poorly with the motor UPDRS score, age, and education level. The relatively low correlation between PDQUALIF scores and UPDRS motor scores and the high correlation between PDQUALIF scores and SIP and SF-36 scores provide empirical evidence suggesting that the domain content of the PDQUALIF is different conceptually from that of the UPDRS, one of the most commonly used outcome measures in PD.

The PDQUALIF has similarities with, and differences from, the four existing PD-specific HRQoL instruments: the Parkinson's disease questionnaire-39 (PDQ-39),³³ the Parkinson's disease quality of life questionnaire (PDQL),³⁴ the Parkinson's impact scale (PIMS),³⁵ and the Parkinson LebensQualitat (PLQ).³⁶ Patients were involved in the original generation and evaluation of items for the PDQUALIF, the PDQ-39, the PDQL and the PLQ, whereas items were decided upon by a consensus of 10 nurses for the PIMS. All five instruments can be completed by the patient, but also can be administered easily by an interviewer. All five instruments use a 5-point ordinal scoring system. They can be completed in 10 to 20 minutes with a range from 10 (PIMS) to 44 (PLQ) questions. Compared with the seven subscales in the PDQUALIF, the PDQ-39 has eight subscales, the PDQL has four subscales, and the PLQ has nine.

The content of the existing scales differs. Compared with the other scales, the PDQUALIF has more emphasis on nonmotor impairments and disabilities and has more questions devoted to the "social" domain of HRQoL ($n = 12$). It is also the only instrument that has specific questions pertaining to fatigue and driving ability, both deemed important to patients with PD.^{31,32} At present, the PDQ-39 has been tested most thoroughly, has ade-

quate psychometric characteristics, and has been used in the largest number of studies. The PDQUALIF provides another option, however, particularly if the goals of the study are to monitor nonmotor impairments and social functioning in complement with the UPDRS. It is unlikely that any single HRQoL instrument will address all issues relevant to a particular therapy or subgroup of subjects; therefore, different measures may be applicable to different interventions or subgroups.²⁸

This study and its resulting instrument have several limitations. First, the instrument was tested in a sample of patients selected from Movement Disorder clinics across North America who had predominantly early to mid-stage disease, with relatively few patients in either very early or very late stages of disease. Second, the patients enrolled in this study, who were highly educated and nearly all Caucasian, may differ substantially in ways that impact self-reported HRQoL from patients in private practice settings, from other parts of the world, or with different socioeconomic backgrounds. Third, item reduction was based on the clinical judgment of health care professionals and not on the values and preferences of patients. Fourth, the most problematic subscales included the Independence subscale, which demonstrated significant floor effects, and the Physical Function subscale, which had low internal consistency.

Nonetheless, this study and instrument have several strengths. The item pool creation was carried out in interviews with many patients and spouses, many patients participated in the validation study across multiple sites, response rates were good, and the instrument was compared to two generic HRQoL scales, as well as relevant clinical data. Testing of the PDQUALIF has established preliminary reliability, validity, and factor structure. Reliability, validity, and responsiveness of the instrument will be assessed further in ongoing large randomized controlled trials in patients with both early and late PD.

Acknowledgments: This study was partially funded by the Pharmacia Corporation (Peacock, NJ). Special thanks to the Parkinson Study Group, Quality of Life Task Force; C. Brower, MPH, Department of Biostatistics, University of Rochester; G. Regalado, D.C. Robinson, and M. Lew, MD, Department of Neurology, University of Southern California; C. Waters, MD, Department of Neurology Columbia University; C. Casaceli and C. O'Connell, University of Rochester; and B. Vickrey, MD MPH, UCLA.

APPENDIX

Participating coordinators: E. Rost-Ruffner, RN BS, University of Virginia; D. McGuire, RN, University of Kansas; B. Fussell, RN, Yale University; L. Shulman, MD, University of Miami; P. Gray, RN, Ottawa Civic Hospital; D. Amyot, McGill Center for Studies of Aging;

C. Reider, Ohio State University; M. Brewer, RN, ANP, Barrow Neurologic Institute; D. Fontaine, University of San Diego; P. Lewis, RN, The Parkinson's Institute; M. Lannon, RN, Brown University; C. Wood, RN, Emory University; S. Rast, University of Tennessee; I. Gardner, RN, University of Rochester; J. Dobson, RN, University of Iowa.

REFERENCES

- Bennett DA, Beckett LA, Murray AM, Shannon KM, Goetz CG, Pilgrim DM, Evans DA. Prevalence of parkinsonian signs and associated mortality in a community population of older people. *N Engl J Med* 1996;334:71-76.
- Poewe WH, Wenning GK. The natural history of Parkinson's disease [review]. *Ann Neurol* 1998;44(3 Suppl. 1):S1-S9.
- Koller WC. Adverse effects of levodopa and other symptomatic therapies: impact on quality of life of the Parkinson's disease patient. Beyond the decade of the brain. Turnbridge Wells: Turnbridge Wells;1994. p 31-46.
- Marinus J, Ramaker C, van Hilten JJ, Stiggelbout AM. Health related quality of life in Parkinson's disease: a systematic review of disease specific instruments. *J Neurol Neurosurg Psychiatry* 2002;72:241-248.
- Karlsen KH, Tandberg E, Arsland D, Larsen JP. Health related quality of life in Parkinson's disease: a prospective longitudinal study. *J Neurol Neurosurg Psychiatry* 2000;69:584-589.
- Schrag A, Jahanshahi M, Quinn N. How does Parkinson's disease affect quality of life? A comparison with quality of life in the general population. *Mov Disord* 2000;15:1112-1118.
- Caap-Ahlgren M, Dehlin O. Insomnia and depressive symptoms in patients with Parkinson's disease. Relationship to health-related quality of life. An interview study of patients living at home. *Arch Gerontol Geriatr* 2001;32:23-33.
- Krippendorff K. Content analysis: an introduction to its methodology. Beverly Hills: Sage Publications; 1980.
- Ware JE. Conceptualizing disease impact and treatment outcomes. *Cancer* 1984;53:2316-2326.
- Waltz CF, Strickland OL, Lenz E. Measurement in nursing research. Philadelphia: FA Davis; 1984.
- Nunnally J. Psychometric theory. Second ed. New York: McGraw Hill; 1978.
- Hoehn MM, Yahr MD. Parkinsonism: onset, progression and mortality. *Neurology* 1967;17:427-442.
- Lang AE, Fahn S. Assessment of Parkinson's disease. In: Munsat TL, editor. Quantification of neurologic deficit. Boston, MA: Butterworth-Heinemann; 1989. p 285-309.
- Montgomery GK, Reynolds NCJ, Warren RM. Qualitative assessment of Parkinson's disease: study of reliability and data reduction with an abbreviated Columbia Scale. *Clin Neuropharmacol* 1985;8:83-92.
- Martinez-Martin P, Gil-Nigel A, Gracia LM, Gomez JB, Martinez-Sarries, Bermejo F. Unified Parkinson's Disease Rating Scale characteristics and structure. The Cooperative Multicentric Group. *Mov Disord* 1994;9:76-83.
- van Hilten JJ, van der Zwan AD, Zwinderman AH, Roos RAC. Rating impairment and disability in Parkinson's disease: evaluation of the Unified Parkinson's Disease Rating Scale. *Mov Disord* 1994;9:84-88.
- Ware JE. SF-36 health survey: manual and interpretation guide. Boston: The Health Institute; 1993.
- Bergner M, Bobbit RA, Carter WB, Gilson BS. The sickness impact profile: development and final revision of a health status instrument. *Med Care* 1981;19:787-805.
- Jenkinson C, Peto V, Fitzpatrick R, Greenhall R, Hyman N. Self-reported functioning and well-being in patients with Parkinson's disease: comparison of the Short-form Health Survey (SF-36) and the Parkinson's Disease Questionnaire (PDQ-39). *Age Ageing* 1995;24:505-509.
- Hays RD, Hayashi T. Beyond internal consistency reliability: rationale and user's guide for multitrait analysis program on the microcomputer. *Behav Res Methods Instrum Comput* 1990;22:167-175.
- Cronbach LJ. Coefficient α and the internal structure of test. *Psychometrika* 1951;53:63-70.
- Fleiss JL. Design and analysis of clinical experiments. New York: John Wiley and Sons; 1986.
- Fleiss JL. Statistical methods for rates and proportions. New York: John Wiley and Sons; 1981.
- Vickrey BG, Hays RD, Genovese BJ, Myers LW, Ellison GW. Comparison of a generic to disease-targeted health-related quality-of-life measures for multiple sclerosis. *J Clin Epidemiol* 1997;50:557-569.
- Miller R. Simultaneous statistical inference. Second ed. New York: Springer-Verlag; 2001.
- Guyatt GH, Jaeschke R, Feeny D, Patrick DL. Measurements in clinical trials: choosing the right approach. In: Spitzer B, editor. Quality of life and pharmacoeconomics in clinical trials. Philadelphia: Lippincott Williams and Wilkins; 1996. p 41-48.
- Guyatt GH, Bombardier C, Tugwell PX. Measuring disease specific quality of life in clinical trials. *CMAJ* 1986;134:889-895.
- Damiano AM, Snyder C, Strausser B, Willian MK. A review of health-related quality-of-life concepts and measures for Parkinson's disease. *Qual Life Res* 1999;8:235-243.
- Shulman LM, Taback RL, Bean J, Weiner WJ. Comorbidity of the nonmotor symptoms of Parkinson's disease. *Mov Disord* 2001;3:507-100.
- Friedman J, Friedman H. Fatigue in Parkinson's disease. *Neurology* 1993;43:2016-2019.
- Friedman JH, Friedman H. Fatigue in Parkinson's disease: a nine-year follow-up. *Mov Disord* 2001;16:1120-1122.
- Frucht S, Rogers JD, Greene PE, Gordon MF, Fahn S. Falling asleep at the wheel: motor mishaps in persons taking pramipexole and ropinirole. *Neurology* 1999;52:1908-1910.
- Peto V, Jenkinson C, Fitzpatrick R, Greenhall R. The development and validation of a short measure of functioning and well being for individuals with Parkinson's disease. *Qual Life Res* 1995;4:241-248.
- De Boer AGEM, Wijker W, Speelman JD, De Haes JCJM. Quality of life in patients with Parkinson's disease: development of a questionnaire. *J Neurol Neurosurg Psychiatry* 1996;61:70-74.
- Calne S, Schulzer M, Mak E. Validating a quality of life rating scale for idiopathic parkinsonism: Parkinson's impact scale. *Parkinsonism Relat Disord* 1996;2:55-61.
- Van den Berg M. Leben mit Parkinson: Entwicklung und psychometrische testung des fragenbogens PLQ. *J Neurol Rehabil* 1998;4:221-226.