Movement Disorders
Vol. 22, No. 16, 2007, pp. 2332–2338

Measuring Health-Related Quality of Life in MSA: The MSA-QoL

Anette Schrag, MD, PhD,[1,2]* Caroline Selai, PhD,[2] Chris Mathias, MD, FMedSci,[2] Philip Low, MD,[3] Jeremy Hobart, MD, PhD,[2,4] Niall Brady, MSc,[2] and Niall Patrick Quinn, MD[2]

[1]*Department of Clinical Neurosciences, Royal Free and University College Medical School, University College London, London, United Kingdom*
[2]*Institute of Neurology, University College London, London, United Kingdom*
[3]*Department of Neurology, Mayo Clinic College of Medicine, Rochester, Minnesota, USA*
[4]*Department of Clinical Neuroscience, Peninsula Medical School, Plymouth, United Kingdom*

Abstract: The objective of this study was to develop a new patient-reported outcome measure for patients with multiple system atrophy (MSA) and to test its psychometric properties. There were three stages. First, a pool of potential scale items was generated from in-depth patient interviews. Second, these items were administered, in the form of a questionnaire, to a sample of people with MSA and traditional psychometric methods used to develop a rating scale satisfying standard criteria for reliability and validity. Third, the psychometric properties of the rating scale were examined in a second sample. In stage one, a pool of 105 items was generated from 20 patient interviews. In stage two, a scale with three subscales (motor, 14 items; nonmotor, 12 items; emotional/social functioning, 14 items), satisfying standard criteria for reliability and validity, was developed from the response data of 317 patients with MSA (response rate 71%). In stage three, the scale was examined in 286 people with MSA. Missing data were low, scores in both subscales were evenly distributed, and floor and ceiling effects were small. Reliability was high (Cronbach's alpha 0.83–0.93; test-retest ICC 0.88–0.92). Validity was supported by the interscale correlations (r = 0.47–0.59), known group differences, and the magnitude and pattern of correlations with four other rating scales, disease severity, and disease duration. In conclusion, the patient-rated MSA health-related Quality of life scale (MSA-QoL) may be a suitable patient-reported scale for use in clinical trials and studies in MSA. © 2007 Movement Disorder Society

Key words: multiple system atrophy; quality of life; scale; validation; development

Multiple system atrophy (MSA) is a neurodegenerative disease with a prevalence of about 2 to 5 per 100,000,[1–3] and leads to reduced life expectancy, increasing disability, and considerable impact on health-related quality of life (Hr-QoL).[4–6] Average age of onset is in the 6th decade[6,7] and mean survival from the onset of symptoms is around 5 to 9 years.[6–8] A scale for the objective assessment of this disease has been developed,[9] but no patient-reported outcome measure exists for use in clinical trials in this disorder. However, it is now widely acknowledged that patient-reported outcome measurement is an important adjunct to clinical data.[10–13] The aim of this study was to develop a Hr-QoL scale for patients with MSA based on clinician- and patient-involvement using standard scale development techniques.[14–17]

*Correspondence to: Dr. Anette Schrag, Department of Clinical Neurosciences, Royal Free and University College Medical School, Rowland Hill Street, London NW3 2PF, United Kingdom.
E-mail: a.schrag@medsch.ucl.ac.uk

Received 9 November 2006; Revised 6 February; 31 May 2007 accepted 31 May 2007

Published online 3 October 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/mds.21649

PATIENTS AND METHODS

Overview

The MSA-QoL was developed and tested in three stages. First, a pool of 105 items was generated from patient interviews. Second, these items, formatted as a questionnaire, were completed by a large sample of people with MSA, and standard item reduction methods were used to develop a clinically meaningful and psychometrically sound 40-item instrument. Finally, psy-

chometric properties of the MSA-QoL (data quality, scaling assumptions, acceptability, reliability, and validity) were evaluated in another sample of 286 people with MSA. The study was approved by the local research ethics committees.

Stage 1: Item Generation.

A pool of potential scale items was generated from semi-structured interviews of 20 people with MSA, review of the literature, and consultation with five neurologists with a special interest in MSA. Patients were selected from neurology clinics to represent the full spectrum of the disorder, and new interviews were conducted until no new themes emerged. Items generated by this process were given a standard[18] five response option format (0 = no problem to 4 = extreme problem). We also included a "not applicable " response option to account for people unable to answer individually the items, but this was not assigned a score. A time frame of 4 weeks was used as this was considered to be clinically sensible. The items were formatted as a questionnaire, reviewed by neurologists, and piloted in a small convenience sample of patients with MSA (n = 6) for readability, ambiguity, and clarity. Adjustments were made in response to the patient feedback.

Stage 2: Scale Development (First Field Test).

The preliminary questionnaire was distributed by the Sarah Matheson Trust for Multiple System Atrophy (SMT) in the UK to all eligible patient members (n = 444). No patients were excluded because they were considered unable to complete the questionnaire. The questionnaire included instructions requesting that the questions should be completed according to the patients' answers even if completed by a caregiver. It also included questions on demographic features, time of onset, a question on severity of MSA (range 0–4), and a visual analogue scale (VAS) of how satisfied the patient feels overall with his/her life. The questionnaire was also administered to 62 patients who fulfilled consensus criteria for probable or possible MSA[19] attending clinics at the Mayo Clinic, Rochester. Standard item reduction methods were used to develop a clinically meaningful and psychometrically sound instrument (see data analysis).

Stage 3: Scale Evaluation (Second Field Test).

The psychometric properties of the final questionnaire were examined in a second field test. The questionnaire was sent to 505 members of the SMT ~2.5 years after the first field test. Participants were asked to complete a booklet consisting of the 40-item MSA-QoL; self-rated MSA severity on a 5-point scale; 39-item Parkinson's Disease Questionnaire (PDQ-39)[20]; the EQ-5D (EQ-5D, a generic Hr-QoL measure[21]); the Hospital Anxiety and Depression Scale (HADS[22]) and a VAS of how satisfied the person felt with their life. In addition, the booklet contained questions on their demography and duration of MSA symptoms. A subset of 100 patients was sent a second identical questionnaire 2 weeks later to examine test–retest reliability.

Data Analysis

Data are presented as percentages, means and standard deviations, or medians, as appropriate.

Analysis of Item Response Data (First Field Test).

First, we examined the percent missing data for each item, and those with values exceeding 10% were removed.[23] Next, we performed an exploratory factor analysis (principal components analysis, varimax rotation) of the remaining items to determine the potential measurement dimensions within the item pool. A range of potential factor solutions was examined and compared to determine which was the most clinically sensible and statistically clear. For each factor solution we examined, the content of the factor, factor loadings, the degree of item cross-over (loadings of $\geq$0.40 on more than one factor), and equivalent loading (loadings on more than one factor within 0.10)[24] in order to obtain factors with meaningful content and little statistical overlap.

Further Item Reduction. To make the instrument short and user-friendly, we then examined the potential to reduce the number of items further. Item groups derived from factor analysis were examined to determine whether they could, or should, be refined in order to generate scales that satisfied criteria for scaling assumptions, acceptability, reliability, and validity. Specifically, we examined (1) each item's correlations with each of the potential scales to detect items that might be poor discriminators between constructs and therefore confound measurement[25] (items with item cross-over or equivalent loading were eliminated in order to obtain factors with meaningful content and little statistical overlap), (2) the item content of the groups to determine whether any items should be removed or retained on clinical validity grounds,[26] and (3) the correlations among the items in each group to determine whether any items might be considered redundant.[27] Items were eliminated if they had item–item correlations $\geq$0.70 and overlap in content with other items, low item-total correlations (i.e. the least reliable indicators of that construct) and/or the content did not fit within the context of the scale, and their elimination did not negatively affect reliability, content, and relative validity.

The final instrument had 40 items measuring the motor, nonmotor, and emotional/social impact of MSA. Scores for the three subscales were generated by summing items and, for ease of interpretation, transformed to a range of 0 to 100 (100 × [(observed score minus min possible score)/(max possible score minus min possible score)]).

Scale Evaluation (Second Field Test).

Parametric methods were used following examination of the distribution of the data. Correlations between the MSA-QoL subscales and other measures were assessed using Spearman's rank correlations due to the ordinal nature of the scales. Standard traditional methods were used to examine five psychometric properties.[14–17,28,29] This included examination of the data quality (percent missing data and percent computable scores), scaling assumptions (equivalent distributions of item response option frequency; similar mean item scores and standard deviations; substantial (>0.30) and similar item-total correlations with little overlap between the subscales), acceptability (even score distribution of the subscales total scores with means near midpoint, floor, and ceiling effects less than 20%, and low skewness (between −1 and +1)), reliability (Cronbach's alpha; test–retest reliability (one-way random effect model intraclass correlation coefficients)), and validity (interscale correlations, convergent and discriminant construct validity (correlations with other scales and variables), and group differences validity). In addition, the Standard Error of Measurement (SEM), a sample size-independent measure of scale precision, was calculated using the formula $SEM = SD^*\sqrt{1 - r_{xx}}$, where r_{xx} is the reliability of the instrument, in order to give an estimate of change scores that are likely to reflect a statistically significant change for individual patients.

RESULTS

Stage 1: Item Generation

The item generation interviews were conducted in 20 patients. Their median age was 62 (range 45–71) years and 50% were women. Ten had the parkinsonian subtype of MSA and 10 the cerebellar subtype.

Stage 2: Scale Development

The initial 105-item questionnaire was returned by 317 members of the SMT in the UK (71.4% response rate). Of these, 69 could not be included (37 had died; 8 had moved; 8 indicated their diagnosis was uncertain or changed; 16 returned blank). Sixty-two consecutive patients attending the neurology clinic at the Mayo clinic additionally agreed to complete the questionnaires, all of which could be included.

Patient characteristics of the analyzed 310 questionnaires are shown in Table 1. Forty-six percent of patients completed the questionnaires themselves, and 54% were completed by a caregiver. Those who completed the questionnaire themselves did not differ in age or gender from those whose caregiver had completed it, but they had a significantly more severe self-reported disease severity ($P < 0.001$) on the self-rated MSA staging system.

Analysis of Item Response Data

Seventeen items of the initial 105 were excluded because of >10% missing (n = 5) or not applicable (n = 12) data. Factor analysis of the remaining items did not produce a pure solution, and so 2, 3, 4, and 5-factor solutions were examined. The three-factor solution was the most statistically clear and clinically sensible. Factor 1 included items on mobility, coordination, self-care, and activities of daily living and was defined a motor scale. Factor 2 covered autonomic dysfunction, energy/sleep, pain, vision, and cognition and was defined a nonmotor scale. Factor 3 included items on social and emotional

TABLE 1. *Characteristics of patients participating in the first and second field tests*

	First field test			Second field test
	Overall sample	UK sample	US sample	Overall sample
N	310	248	62	286
Mean age (yr; SD)	66.2 (9.2)	65.5 (9.1)	68.9 (9.4)	65.4 (9.8)
Men (%)	55	57	47	55
Mean disease duration (yr; SD)	5.8 (3.7)	6.0 (3.7)	4.7 (3.2)	6.3 (3.9)
Mean disease severity[a] (range 0–4; SD)	2.7 (1.1)	2.7 (1.1)	2.8 (1.1)	2.9 (0.9)
Mean overall life satisfaction[b] (range 0–100; SD)	36.5 (24.8)	36.7 (23.8)	35.6 (28.6)	39.0 (23.2)

[a]As assessed on an MSA scale modified for self-completion.
[b]On a visual analogue scale from 0 (worst possible satisfaction with life) to 100 (best possible satisfaction with life).

aspects and was defined an emotional/social scale. All items could be referenced back to statements made by patients or caregivers. Further item reduction (see above) led to final subscales with 14 (motor), 12 (nonmotor), and 14 (emotional/social) items.

Stage 3: Scale Evaluation

The final scale was returned in the second field test by 346 patients (68.5% response rate). Sixty replies could not be included because of death (n = 32), change of diagnosis (n = 11), questionnaire blank (n = 14), or invalid (a different questionnaire or missing pages; n = 2), or as the patient had moved (n = 1). The response rate in the test–retest sample was 80%.

Data Quality

Item level missing data were low (range 0.3–2.8%), and subscale scores could be computed for 99% of the sample (Table 2).

Scaling Assumptions

In each scale, range of scores was distributed across the spectrum with small floor and ceiling effects, item response frequency distributions were relatively symmetrical and not unduly skewed, and item mean scores and standard deviations were similar. Factor analysis of the 40 items supported the grouping of items into the three scales.

Reliability

For both scales, corrected item-total correlations exceeded the recommended >0.30[25,30] and all reliability estimates exceeded 0.80.[14] Test–retest reliability was good with correlations around 0.9 for all three subscales. The SEM of both subscores was ~6, giving 95% confidence intervals for individual patients' scores of approximately ±12 scale points.

Validity

The content validity of the final subscales was assessed by the expert panel and considered to be good. Internal construct validity was supported by the moderate inter-scale correlation between the subscales (r = 0.47–0.59), implying that the two MSA-Qol subscales measure related but different health constructs. Convergent and discriminant construct validity was supported by finding correlations with other scales and variables as well as group differences that were consistent with a priori predictions (positive correlations of the motor subscale with motor severity, the mobility, ADL, and communications subscale of the PDQ-39 and EQ-5D; positive correlations of the emotional subscale with the depression and anxiety subscales of the HADS, the anxiety/depression subscales of the EQ-5D, and the emotional wellbeing subscale of the PDQ-39; positive correlation of the nonmotor subscale with the cognitive impairment subscale and bodily discomfort subscale of the PDQ-39; lack of correlation of the motor subscale with the cognitive impairment, emotional well-being and social support subscales of the PDQ-39, the anxiety/depression subscale of the EQ-5D, and the HADS anxiety scale; and negative and substantial correlations of all three scales with life satisfaction; Tables 3 and 4). There were no differences between scores for men and women

TABLE 2. *Score distributions and reliability of the MSA-QoL*

	Motor subscale	Nonmotor subscale	Emotional/social subscale
Data quality			
Missing data (%)	1.15	1.55	0.94
Computable scale scores (%)	99	99	99
Scaling assumptions			
Range of item mean scores	1.30–3.30	0.90–2.81	1.56–2.67
Range of item SD	1.02–1.38	1.15–1.41	1.21–1.48
Range of corrected item-total correlations	0.52–0.79	0.36–0.64	0.58–0.75
Acceptability			
Mean (SD)	61.3 (23.0)	51.3 (21.5)	54.1 (26.2)
Median	62.5	50.0	53.6
Range of scores (0–100)	3.6–100	10.4–100	1.8–100
Skewness/SE skewness	−0.18/0.15	0.26/0.15	0.02/0.15
Floor effects (%)	0	0	0
Ceiling effects (%)	4.2	0.7	2.1
Reliability			
Internal consistency (Cronbach's alpha)	0.93	0.84	0.94
Test–retest (intraclass correlation coefficient)	0.90	0.88	0.92
Standard error of measurement (SEM)	6.0	4.9	6.6
95% confidence intervals (1.96 × SEM)	11.8	9.7	13.0

TABLE 3. *Spearman's correlation coefficients between the MSA-QOL subscales and measures of health status and psychological well-being in the second survey*

	Motor subscale	Nonmotor subscale	Emotional/social subscale
Symptom severity (0–4)	0.72	0.26	0.26
Disease duration (yr)	0.20	0.17	0.15
Overall life satisfaction (0–100)	−0.40	−0.46	−0.57
Age (yr)	0.05	−0.08	−0.03
PDQ 39			
Mobility	0.76	0.41	0.49
Activities of daily living	0.82	0.43	0.43
Emotional wellbeing	0.25	0.47	0.75
Stigma	0.31	0.36	0.60
Social support	0.14	0.28	0.48
Cognitive impairment	0.12	0.52	0.40
Communication	0.55	0.31	0.46
Bodily discomfort	0.14	0.63	0.33
Index	0.58	0.65	0.75
EQ-5D			
Mobility	0.55	0.29	0.64
Activities of daily living	0.75	0.41	0.34
Self-care	0.63	0.28	0.37
Pain	0.06	0.48	0.3
Anxiety/depression	0.22	0.45	0.65
Index	−0.66	−0.59	−0.58
Visual analogue scale	0.76	0.41	0.49
Hospital Anxiety and Depression Scale			
Anxiety	0.11	0.55	0.75
Depression	0.4	0.58	0.64

Worse health is indicated by higher scores on the MSA-QoL subscales, the PDQ 39 subscales and index, the EQ-5D subscales and the Hospital and Anxiety subscales, and by lower scores on the overall life satisfaction and EQ-5D visual analogue scales and the EQ-5D index.

which, as MSA affects men and women similarly,[32] also supports the validity of the MSA-Qol.

DISCUSSION

To date, the assessment of Hr-QoL in patients with MSA has been difficult because of the lack of an MSA-specific measure. While generic instruments, such as the SF 36,[4,5,33] have been used in patients with MSA and have the advantage that they allow comparison between disease groups, these instruments have limited feasibility and acceptability in older patients[34] and patients with neurodegenerative disease.[35–38] In addition, both generic instruments and instruments developed for the related but different disorder of Parkinson's disease do not address, and are unlikely to be sensitive to, specific features important to patients with MSA, such as severe autonomic dysfunction, early sexual impairment, severe bulbar involvement, and incoordination, and may underestimate health problems in MSA. On the other hand, assessments of disease severity using clinical rating scales omit patient views about issues of importance to their health, particularly those of emotional and social functioning and the impact of dysfunction on activities of daily living. Thus, a subjective Hr-QoL scale derived from patient perspective in patients with MSA will provide an important complement to objective clinical rating scales.

We have developed a new instrument to assess patients with MSA, which is based on patient and clinician views and psychometric analysis of data from a large field test in patients with MSA. Items were selected to ensure a low number of missing values, applicability of the scale to patients at all stages of disease, no undue skewness, homogeneity of items with the scales allowing for replacement of missing values and robustness of the scale, high reliability, and clinical significance. We took care to include areas that have been identified to be important to patients with MSA from the literature and the initial interviews. As one of the main aims for a Hr-QoL scale is to capture change, we also selected items with good discriminative power between different patient groups of different severity of disease and health impairment, and for low ceiling or floor effects to ensure sensitivity at both ends of the spectrum of disease. The results of the psychometric testing of the instrument in a second field test of patients with MSA indicated that this instrument satisfies standard criteria as a summed rating scale and is both highly reliable and has good face and construct valid-

TABLE 4. *MSA-QoL known group differences*

Symptom severity	Motor subscale	Nonmotor subscale	Emotional/social subscale
EQ-5D mobility dimension			
No problems in walking about (n = 7)	21.8 (11.8)	38.9 (14.0)	29.6 (35.4)
Some problems in walking about (n = 209)	57.2 (20.2)	48.1 (18.7)	52.1 (24.0)
Confined to bed (n = 47)	86.8 (13.2)	67.2 (26.7)	66.2 (29.8)
F-value* (P-value)	60.8 (<0.0001)	17.5 (<0.0001)	9.2 (<0.0001)
Relative validity	1.0	0.84	0.15
EQ-5D self-care dimension			
No problems with self-care (n = 35)	34.9 (19.1)	35.1 (15.2)	37.6 (28.7)
Some problems washing or dressing myself (n = 124)	52.1 (15.6)	48.0 (16.8)	51.2 (22.5)
Unable to wash or dress myself (n = 105)	82.0 (13.2)	62.3 (23.8)	62.6 (26.5)
F-value* (P-value)	166.2 (<0.0001)	28.6 (<0.0001)	14.4 (<0.0001)
Relative validity	1.0	0.17	0.09
EQ-5D activities of daily living dimension			
No problems with performing my usual activities (n = 5)	11.1 (6.2)	26.7 (8.1)	20.7 (9.0)
Some problems with performing my usual activities (n = 94)	44.8 (17.1)	45.4 (17.9)	43.0 (24.0)
Unable to perform my usual activities (n = 164)	72.3 (18.1)	55.9 (22.6)	61.3 (24.9)
F-value* (P-value)	91.8 (<0.0001)	10.8 (<0.0001)	21.7 (<0.0001)
Relative validity	1.0	0.12	0.23
EQ-5D pain dimension			
No pain or discomfort (n = 57)	61.1 (23.8)	38.3 (24.8)	41.0 (27.2)
Moderate pain or discomfort (n = 160)	59.6 (23.2)	51.4 (17.9)	54.6 (23.6)
Extreme pain or discomfort (n = 46)	67.3 (21.1)	69.8 (17.0)	67.8 (26.8)
F-value* (P-value)	2.0 (ns)	28.8 (<0.0001)	14.8 (<0.0001)
Relative validity	0.07	1.0	0.51
EQ-5D anxiety/depression dimension			
Not anxious or depressed (n = 85)	54.7 (24.1)	41.5 (23.2)	30.8 (20.6)
Moderately anxious or depressed (n = 143)	62.3 (22.0)	53.1 (18.7)	59.5 (20.2)
Extremely anxious or depressed (n = 37)	71.5 (19.3)	69.1 (15.0)	82.7 (15.2)
F-value* (P-value)	7.5 (<0.01)	22.7 (<0.0001)	101.1 (<0.0001)
Relative validity	0.07	0.22	1.0

*Calculated as the ratio of paired-F-values using the largest as the denominator.[29,31]

ity. These data suggest that this scale could be a useful patient-reported outcome measure in clinical trials.

Our study has limitations. Although we were careful to include patients at all stages of MSA in the development of the questionnaire, it is likely that patients in more advanced stages who cannot respond accurately on their own, particularly those unable to communicate, were under-represented in the postal surveys of patient-generated answers. The applicability of the instrument to these patients is therefore unknown. In this context it is also important to note that only half of the patients could complete the questionnaire on their own, most requiring the help of their caregiver. The importance of completing the questionnaire according to the patient's answers should therefore always be emphasized when the questionnaire is used in these situations. A second limitation is that we have yet not tested the responsiveness of the MSA-QoL. However, we used item discrimination as a criterion for item selection at the second development stage, and have included an evaluation of scale discriminant validity. A third limitation is that both postal surveys were conducted primarily in patient members of the SMT, a patient organization, and we do not know how many of these patients had a diagnosis of probable or possible MSA. Finally, while the development of the scale was based on patient responses in both the UK and North America, the validation study was primarily performed in a UK-based sample. Thus, the validity of the scale has to be tested further in other populations.

With these caveats, the MSA-QoL is the first outcome measure for patients with MSA, which has been developed with input from patients, caregivers, clinicians, and literature review and using rigorous psychometric testing. It has been shown to have good psychometric properties and is reliable, acceptable, and valid. It may be used not only to assess Hr-QoL in individual patients but also in observational and epidemiological trials, to relate neuroimaging or pathophysiological findings to patients' Hr-QoL and, most importantly, in longitudinal studies and trials of therapeutic interventions alongside clinical rating scales. This is particularly important as a number of therapies are currently under consideration or in clinical trials for the treatment of patients with MSA. Whether these therapies can improve the patients' own appreciation of their health and health-related quality of life can be demonstrated using this new Hr-QoL measure.

Acknowledgments: This study was supported by the Sarah Matheson Trust for Multiple System Atrophy. The corresponding author had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

REFERENCES

1. Schrag A, Ben Shlomo Y, Quinn NP. Prevalence of progressive supranuclear palsy and multiple system atrophy: a cross-sectional study. Lancet 1999;354:1771–1775.
2. Chrysostome V, Tison F, Yekhlef F, Sourgen C, Baldi I, Dartigues JF. Epidemiology of multiple system atrophy: a prevalence and pilot risk factor study in Aquitaine, France. Neuroepidemiology 2004;23:201–208.
3. Vanacore N, Bonifati V, Fabbrini G, et al., for European Study Group on Atypical Parkinsonisms. Epidemiology of multiple system atrophy. ESGAP Consortium. Neurol Sci 2001;22:97–99.
4. Schrag A, Geser F, Stampfer-Kountchev M, et al. Health-related quality of life in multiple system atrophy. Mov Disord 2006;21: 809–815.
5. Benrud-Larson LM, Sandroni P, Schrag A, Low PA. Depressive symptoms and life satisfaction in patients with multiple system atrophy. Mov Disord 2005;20:951–957.
6. Wenning GK, Tison F, Ben Shlomo Y, Daniel SE, Quinn NP. Multiple system atrophy: a review of 203 pathologically proven cases. Mov Disord 1997;12:133–147.
7. Jellinger KA, Seppi K, Wenning GK. Grading of neuropathology in multiple system atrophy: proposal for a novel scale. Mov Disord 2005;20 (Suppl 12):S29–S36.
8. Ben Shlomo Y, Wenning GK, Tison F, Quinn NP. Survival of patients with pathologically proven multiple system atrophy: a meta-analysis. Neurology 1997;48:384–393.
9. Wenning GK, Tison F, Seppi K, et al. Development and validation of the unified multiple system atrophy rating scale (UMSARS). Mov Disord 2004;19:1391–1402.
10. Vickrey BG, Hays RD, Rausch R, et al. Outcomes in 248 patients who had diagnostic evaluations for epilepsy surgery. Lancet 1995; 346:1445–1449.
11. Vickrey BG, Hays RD, Engel J, Jr, et al. Outcome assessment for epilepsy surgery: the impact of measuring health-related quality of life. Ann Neurol 1995;37:158–166.
12. Merkies IS, Schmitz PI, van der Meche FG, Samijn JP, van Doorn PA. Quality of life complements traditional outcome measures in immune-mediated polyneuropathies. Neurology 2002;59:84–91.
13. Fletcher J. Patient page. Treatment and quality of life for people with ALS. Neurology 2004;62:E22–E23.
14. Nunnally JC, Bernstein IH. Psychometric theory, 3rd ed. New York: McGraw-Hill; 1994.
15. Juniper EF, Guyatt GH, Jaeschke R. How to develop and validate a new health-related quality of life instrument. In: Spilker B, editor. Quality of life and pharmacoeconomics in clinical trials. Philadelphia, NY: Lippincott-Raven; 1996.
16. Marx RG, Bombardier C, Hogg-Johnson S, Wright JG. Clinimetric and psychometric strategies for development of a health measurement scale. J Clin Epidemiol 1999;52:105–111.
17. Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use, 2nd ed. Oxford: Oxford University Press; 1995.
18. Nagata C, Ido M, Shimizu H, Misao A, Matsuura H. Choice of response scale for health measurement: comparison of 4, 5, and 7-point scales and visual analog scale. J Epidemiol 1996;6:192–197.
19. Gilman S, Low PA, Quinn N, et al. Consensus statement on the diagnosis of multiple system atrophy. J Neurol Sci 1999;163:94–98.
20. Peto V, Jenkinson C, Fitzpatrick R. PDQ-39: a review of the development, validation and application of a Parkinson's disease quality of life questionnaire and its associated measures. J Neurol 1998;245 (Suppl 1):S10–S14.
21. EuroQol—a new facility for the measurement of health-related quality of life. The EuroQol Group. Health Policy 1990;16:199–208.
22. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983;67:361–370.
23. World Health Organization. The World Health Organization Quality of Life Assessment (WHOQOL): development and general psychometric properties. Soc Sci Med 1998;46:1569–1585.
24. Ferguson E, Cox T. Exploratory factor analysis: a user's guide. Int J Sel Assess 1993;1:84–94.
25. Ware JE, Harris WJ, Gandek B, Rogers BW, Reese PR. MAP-R for windows: multitrait/multi-item analysis program-revised user's guide. Boston, MA: Health Assessment Laboratory; 1997.
26. Guyatt GH, Jaeschke R, Feeny DH, Patrick DL. Measurements in clinical trials: choosing the right approach. In: Spilker B, editor. Quality of life and pharmacoeconomics in clinical trials. Philadelphia, NY: Lippincott-Raven, 1996. p 41–48.
27. Juniper EF, Guyatt GH, Streiner DL, King DR. Clinical impact versus factor analysis for quality of life questionnaire construction. J Clin Epidemiol 1997;50:233–238.
28. Hobart J, Freeman J, Lamping D, Fitzpatrick R, Thompson A. The SF-36 in multiple sclerosis: why basic assumptions must be tested. J Neurol Neurosurg Psychiatry 2001;71:363–370.
29. Hobart J, Lamping D, Fitzpatrick R, Riazi A, Thompson A. The multiple sclerosis impact scale (MSIS-29): a new patient-based outcome measure. Brain 2001;124:962–973.
30. Likert RA. A technique for the development of attitudes. Arch Psychol 1932;140:5–55.
31. Hobart J, Kalkers N, Barkhof F, Uitdehaag B, Polman C, Thompson A. Outcome measures for multiple sclerosis clinical trials: relative measurement precision of the expanded disability status scale and multiple sclerosis functional composite. Mult Scler 2004; 10:41–46.
32. Geser F, Wenning GK, Seppi K, et al. Progression of multiple system atrophy (MSA): a prospective natural history study by the European MSA study group (EMSA SG). Mov Disord 2006;21: 179–186.
33. Brazier JE, Harper R, Jones NM, et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. BMJ 1992;305:160–164.
34. Brazier JE, Walters SJ, Nicholl JP, Kohler B. Using the SF-36 and Euroqol on an elderly population. Qual Life Res 1996;5:195–204.
35. Schrag A, Selai C, Jahanshahi M, Quinn NP. The EQ-5D—a generic quality of life measure-is a useful instrument to measure quality of life in patients with Parkinson's disease. J Neurol Neurosurg Psychiatry 2000;69:67–73.
36. Freeman JA, Hobart JC, Langdon DW, Thompson AJ. Clinical appropriateness: a key factor in outcome measure selection: the 36 item short form health survey in multiple sclerosis. J Neurol Neurosurg Psychiatry 2000;68:150–156.
37. Hobart JC, Williams LS, Moran K, Thompson AJ. Quality of life measurement after stroke: uses and abuses of the SF-36. Stroke 2002;33:1348–1356.
38. Neudert C, Wasner M, Borasio GD. Patients' assessment of quality of life instruments: a randomised study of SIP, SF-36 and SEIQoL-DW in patients with amyotrophic lateral sclerosis. J Neurol Sci 2001;191:103–109.