*Movement Disorders*
Vol. 22, No. 13, 2007, pp. 1901–1911


# The Metric Properties of a Novel Non-Motor Symptoms Scale for Parkinson's Disease: Results from an International Pilot Study

Kallol Ray Chaudhuri,[1,2,3]* Pablo Martinez-Martin,[4] Richard G. Brown,[5] Kapil Sethi,[6] Fabrizio Stocchi,[7] Per Odin,[8] William Ondo,[9] Kazuo Abe,[10] Graeme MacPhee,[11] Doug MacMahon,[12] Paolo Barone,[13] Martin Rabey,[14] Alison Forbes,[1] Kieran Breen,[15] Susanne Tluk,[16] Yogini Naidu,[16] Warren Olanow,[17] Adrian J. Williams,[18] Sue Thomas,[19] David Rye,[20] Yoshio Tsuboi,[21] Annette Hand,[22] and Anthony H.V. Schapira[23,24]

[1]*National Parkinson Foundation Centre of Excellence, Kings College Hospital, London, United Kingdom*
[2]*Department of Neurology, University Hospital Lewisham, London, United Kingdom*
[3]*Kings College, Institute of Psychiatry, London, United Kingdom*
[4]*Unit of Neuroepidemiology, National Center for Epidemiology, Carlos III Institute of Health, Madrid, Spain*
[5]*Department of Psychology, Institute of Psychiatry, King's College London, London, United Kingdom*
[6]*Department of Neurology, Medical College of Georgia, USA*
[7]*Department of Neurology, IRCCS San Raffaele, Pisana, Roma, Italy*
[8]*Department of Neurology, Klinikum-Bremerhaven Reinkenheide, Bremerhaven, Germany*
[9]*Department of Neurology, Baylor College of Medicine, Houston, Texas, USA*
[10]*Department of Neurology, Osaka University Graduate School of Medicine, Osaka, Japan*
[11]*Medicine for the Elderly, Southern General Hospital, Glasgow, United Kingdom*
[12]*Medicine for Care of Elderly, Camborne-Redruth Hospital, Redruth, Cornwall, United Kingdom*
[13]*Department of Neurological Sciences, University of Napoli, Napoli, Italy*
[14]*Department of Neurology, Assaf Harofeh Medical Center, School of Medicine, Tel-Aviv University, Zerifin, Israel*
[15]*Research and Development Department, The UK Parkinson's Disease Society, London, United Kingdom*
[16]*Department of Neurology, University Hospital Lewisham, London, United Kingdom*
[17]*Department of Neurology, The Mount Sinai School of Medicine, New York, USA*
[18]*Lane Fox Sleep Centre, St. Thomas' Hospital, London, United Kingdom*
[19]*Policy Advice, Royal College of Nursing, London, United Kingdom*
[20]*Department of Neurology, Emory University, Atlanta, Georgia, USA*
[21]*Fifth Department of Internal Medicine, Fukuoka University, Fukuoka, Japan*
[22]*Department of Medicine, North Tyneside General Hospital, North Shields, United Kingdom*
[23]*Department of Clinical Neuroscience, Royal Free Hospital Medical School, London, United Kingdom*
[24]*National Hospital for Neurology and Neurosurgery, London, United Kingdom*

**Abstract:** Non-motor symptoms (NMS) in Parkinson's disease (PD) are common, significantly reduce quality of life and at present there is no validated clinical tool to assess the progress or potential response to treatment of NMS. A new 30-item scale for the assessment of NMS in PD (NMSS) was developed. NMSS contains nine dimensions: cardiovascular, sleep/fatigue, mood/cognition, perceptual problems, attention/memory, gastrointestinal, urinary, sexual function, and miscellany. The metric attributes of this instrument were analyzed. Data from 242 patients mean age 67.2 ± 11 years, duration of disease 6.4 ± 6 years, and 57.3% male across all stages of PD were collected from the centers in Europe, USA, and Japan. The mean NMSS score was 56.5 ± 40.7, (range: 0–243) and only one declared no NMS. The scale provided 99.2% complete data for the analysis with the total score being free of floor and ceiling effect. Satisfactory scaling assumptions (multitrait scaling success rate >95% for all domains except miscellany) and internal consistency were reported for most of the domains (mean α, 0.61). Factor analysis supported the *a prori* nine domain structure (63% of the variance) while a small test–retest study showed satisfactory reproducibility (ICC > 0.80) for all do-

*Correspondence to: K. Ray Chaudhuri, NPF Centre of Excellence, Movement Disorders Unit, 9th Floor Ruskin Wing, King's College Hospital, Denmark Hill, London SE5 9RS, United Kingdom
E-mail: ray.chaudhuri@uhl.nhs.uk
Received 6 March 2007; Revised 23 April 2007; Accepted 24 April 2007
**Published online 2 August 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/mds.21596**

mains except cardiovascular (ICC = 0.45). In terms of validity, the scale showed modest association with indicators of motor symptom severity and disease progression but a high correlation with other measures of NMS (NMSQuest) and health-related quality of life measure (PDQ-8) (both, rS = 0.70). In conclusion, NMSS can be used to assess the frequency and severity of NMS in PD patients across all stages in conjunction with the recently validated non-motor questionnaire. 

**Key words:** PD; non-motor; UPDRS; questionnaire; quality of life.

The cardinal clinical features of Parkinson's disease (PD) include asymmetric onset of bradykinesia, rigidity, and resting tremor. These are the consequences of the loss of dopaminergic neurons in the *substantia nigra pars compacta*. However, it is known that other neurotransmitter systems degenerate in PD with loss in the locus coeruleus, substantia innominata, dorsal motor nucleus, and other areas. Braak et al. introduced the concept of a six-stage pathological process, beginning at "induction sites" with the degeneration of the olfactory bulb, and the anterior olfactory nucleus at stage 1 while during stage 2 the pathological process progresses to lower brainstem involving the brainstem nuclei, which are thought to be key areas mediating non-motor symptoms (NMS) such as olfaction, sleep homeostasis, depression, and and central autonomic control. Studies have demonstrated that several symptoms such as olfactory loss, rapid eye movement sleep disorder, depression, and constipation may precede the motor features of PD.[1,2] Progression is associated with the development of more severe NMS that affects cognition, genitourinary, and gastrointestinal systems, sleep, and a variety of features that come to dominate PD symptomatology and significantly impair the quality of life.[2] Although some NMS such as constipation, sleep disorders, and depression are treatable, others maybe unresponsive to or even exacerbated by dopaminergic medications suggesting that their critical pathopyhsiological substrate to be outside the nigrostriatal system. In the Sydney long-term follow-up study of PD, 52 of the original cohort of 149 were still alive at 15 years and 40% were in institutional care and suffering from a high proportion of NMS.[3] Studies have also documented the adverse effect of NMS such as depression on quality of life, institutionalization rates, and health economics in PD.[2,4-6] However, although common, the NMS of PD are often not well recognized in clinical practice.[7,8] NMS are not identified by neurologists in over 50% of consultations and sleep disturbance, in particular, remain unrecognized in over 40% of PD patients.[9]

Several instruments are available for the quantification of motor features (e.g., Unified Parkinson's Disease Rating Scale, UPDRS) and health-related quality of life [e.g., Parkinson's Disease Questionnaires (PDQ) 8 and 39], but no scale has yet been developed or validated specifically for the assessment of the full range of NMS observed in PD.

We have recently developed and validated a self-reported questionnaire for PD patients for NMS (NMSQuest).[10] NMSQuest comprises 30 common symptoms scored "yes" or "no," and is designed to provide a rapid screen for problematic NMS as an aid to clinical management. It is not a rating scale, nor is it intended to evaluate the effect of treatment. To provide a method to quantify NMS, we have developed a scale covering the range of NMS each scored by the physician for symptom severity and frequency. We have now completed the first validation study of the NMSS in an international cohort of PD patients with a range of disease severity. The results confirm that the new scale provides a valid and reliable assessment of NMS in PD.

## METHODS

### Design

This was an international, cross-sectional, open, multicenter, one point-in-time evaluation with retest study.

### Patients

Consecutive PD patients, of all age groups and disease severity, satisfying the UK PD Brain Bank criteria for the diagnosis of idiopathic PD were included.[11] Patients with a diagnosis of parkinsonism because of alternative causes were excluded. Patients were recruited from movement disorder and care of the elderly clinics to include Hoehn and Yahr (HY) stages 1–5, and included a proportion of untreated cases. Only patients able to provide informed consent were selected for the study and demented patients were excluded. In practice, this meant that the clinician would exlcude patients with significant cognitive impairment that affected either (1) their ability to provide informed consent and/or (2) their ability to provide reliable self-report.

Several centers, who participated in the NMSQuest study, were included across Europe, USA, and Japan because we wanted to test the usefulness of the NMSS in different clinical settings and clinical population.

### Ethical Approval

Central ethical approval for the full study was initially obtained via the research ethics committee at University Hospital Lewisham, and subsequently all centers obtained site-specific ethical approval.

### Procedure

Scales were translated for non-English speaking patients. The scales were used by neurologists/care of the elderly specialists all of whom were involved in the validation process and development of NMSQuest and NMS scale. Translation of the scale was therefore deemed possible in a clinic setting by these investigators similar to the protocol followed in the NMSQuest study. The scales and questionnaires used in the study were applied following a standardized protocol, always in the same order and there was no reported patient fatigue while completing the scales.

### Assessments

The researcher then completed the following battery of standard assessment measures: a standard demography form, Hoehn and Yahr Scale (HY),[12] UPDRS (parts 3 and 4),[13] and various other measures validated for use in PD. These included the Cumulative Illness Rating Scale-Geriatrics (CIRS-G) to measure physical comorbidity,[14-16] the frontal assessment battery (FAB) to measure cognitive function,[17] and the NMSS scale. In addition, the patient (assisted by the research nurse, if necessary) completed the following self-assessments: PDQ-8 (a specific instrument for the assessment of health-related quality of life in PD),[18] a fatigue visual analogue scale (VAS-F, 0 (the worst imaginable fatigue state) to 100 (no fatigue at all),[19,20] Hospital Anxiety and Depression Scale HADS, a self-administered instrument developed for the detection of mood disorders in non-psychiatric outpatients attending hospital consulting rooms[21,22]), and the NMSQuest.[10] The total time required for a single assessment was ~65 min per patient. Details of clinimteric validation of these measures in PD are described elsewhere.[14,15,17-24]

### The NMSS

Originally 50 questions were devised after detailed literature review, expert experience (multidisciplinary members of the PDNMG, UK PD nurse specialist association), patient-response survey[23] (hospital based and the Parkinson's disease society of UK), and evaluation of pilot and extended NMSQuest study data, the latter involving 525 patients from an international setting.[10,24] Broadly, the NMSS reflected the questions flagged in the NMSQuest, and the scale was aimed to be a practical and quantitative measure that encompasses the NMS experienced by patients with PD. The 50 items were reduced to the final 30-item measure by the steering group after analyzing data from the pilot and the extended NMSQuest study[9,24] and item reduction based on factor analysis. Items were then grouped to nine relevant domains: cardiovascular (2 items); sleep/fatigue (4 items); mood/cognition (6 items); perceptual problems/hallucinations (3 items); attention/memory (3 items); gastrointestinal tract (3 items); urinary (3 items); sexual function (2 items); and miscellaneous (4 items) (annex 1).

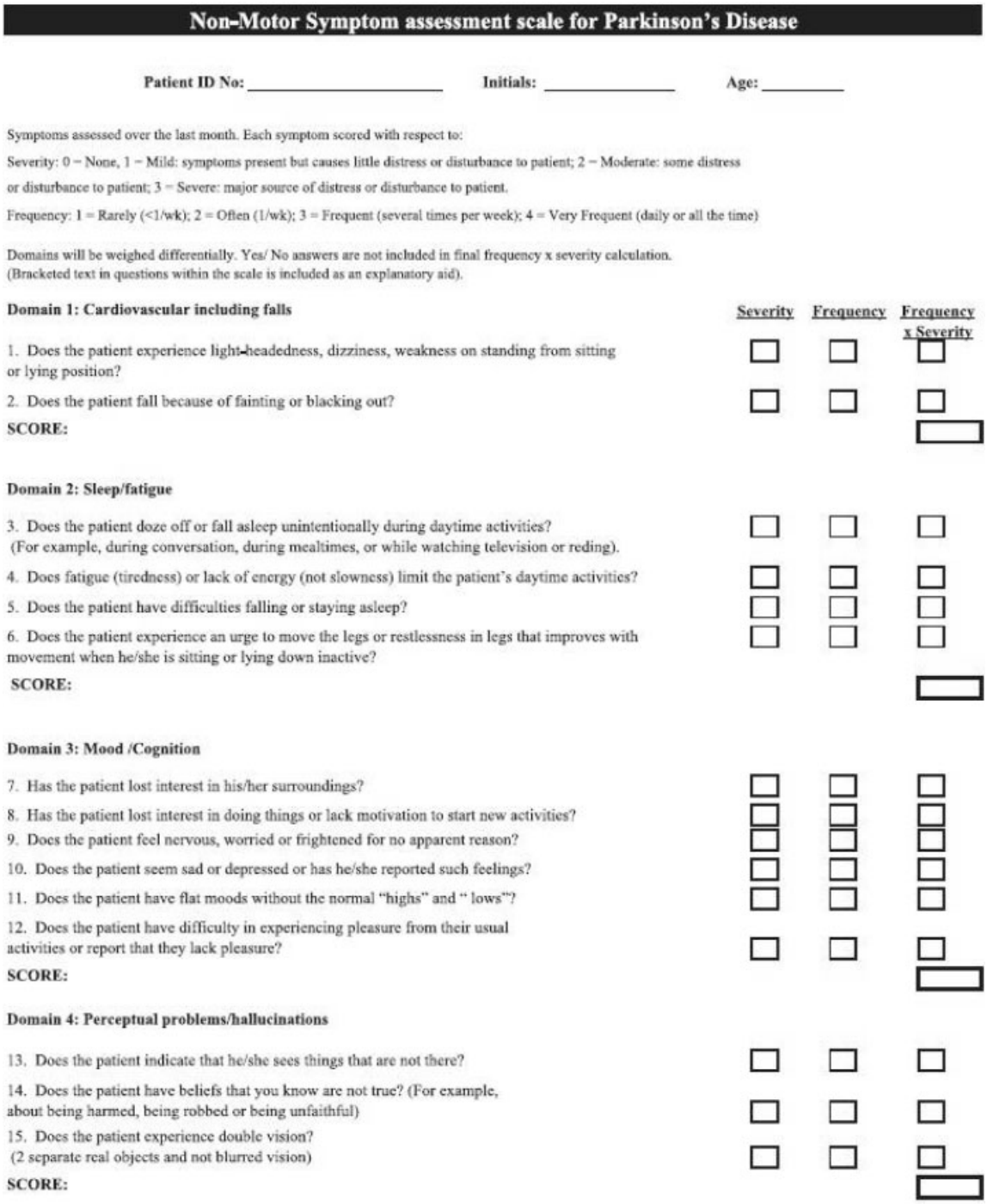
The NMSS is rated by the health professionals and obtained through interview. Score for each item is based on a multiple of severity (from 0 to 3) and frequency scores (from 1 to 4) (Fig. 1). The scale can therefore capture symptoms that are severe but relatively infrequent (e.g., hallucinations) and those that may be less severe but persistent (e.g., constipation, fatigue, or low mood). This method increases the weight of symptoms simultaneously that are persistent and severe.

### Data Analysis

At the conclusion of the study, coded data were transferred to the Neuroepidemiology Unit at Carlos III Institute of Health, Madrid (Spain) for clinimetric analysis.

Data did not fit normal distribution; therefore, nonparametric statistics were conveniently applied. The following attributes were tested:

1. Data quality (missing data, full computable scale scores). A 5% value for missing or noncomputable data was acceptable as the limit.[25]
2. The floor and ceiling effect, and skewness. Upper acceptable value for floor and ceiling effect was 15%[26] and the limits for skewness were $-1$ and $+1$.[27]
3. Scaling assumptions were explored by item-convergent/item-discriminant validity (multitrait analysis) and item-total correlation (corrected for overlap). Multitrait analysis was performed. We calculated correlation coefficients between a specific item and the sum of the items in the same domain. If this value was more than the correlation coefficient with the score of any other domain plus 2 $\times$ standard error of the correlation coefficient, then the results were considered as a success result, while a value higher than correlation coefficient with the score of any other domain but less than twice the standard errors was considered as a "probable success." This was based on the observations of Smith et al.[25] For the item-total correlation, Spearman rank correlation coefficient ($r_S$) values higher than 0.30 were deemed acceptable.[28,29]
4. Reliability was tested for both internal consistency



**Non-Motor Symptom assessment scale for Parkinson's Disease**

**Patient ID No:** _____ **Initials:** _____ **Age:** _____

Symptoms assessed over the last month. Each symptom scored with respect to:

Severity: 0 = None, 1 = Mild: symptoms present but causes little distress or disturbance to patient; 2 = Moderate: some distress or disturbance to patient; 3 = Severe: major source of distress or disturbance to patient.

Frequency: 1 = Rarely (<1/wk); 2 = Often (1/wk); 3 = Frequent (several times per week); 4 = Very Frequent (daily or all the time)

Domains will be weighed differentially. Yes/ No answers are not included in final frequency x severity calculation.
(Bracketed text in questions within the scale is included as an explanatory aid).

| **Domain 1: Cardiovascular including falls** | Severity | Frequency | Frequency x Severity |
|---|---|---|---|
| 1. Does the patient experience light-headedness, dizziness, weakness on standing from sitting or lying position? | ☐ | ☐ | ☐ |
| 2. Does the patient fall because of fainting or blacking out? | ☐ | ☐ | ☐ |
| **SCORE:** | | | ☐ |
| **Domain 2: Sleep/fatigue** | | | |
| 3. Does the patient doze off or fall asleep unintentionally during daytime activities? (For example, during conversation, during mealtimes, or while watching television or reding). | ☐ | ☐ | ☐ |
| 4. Does fatigue (tiredness) or lack of energy (not slowness) limit the patient's daytime activities? | ☐ | ☐ | ☐ |
| 5. Does the patient have difficulties falling or staying asleep? | ☐ | ☐ | ☐ |
| 6. Does the patient experience an urge to move the legs or restlessness in legs that improves with movement when he/she is sitting or lying down inactive? | ☐ | ☐ | ☐ |
| **SCORE:** | | | ☐ |
| **Domain 3: Mood /Cognition** | | | |
| 7. Has the patient lost interest in his/her surroundings? | ☐ | ☐ | ☐ |
| 8. Has the patient lost interest in doing things or lack motivation to start new activities? | ☐ | ☐ | ☐ |
| 9. Does the patient feel nervous, worried or frightened for no apparent reason? | ☐ | ☐ | ☐ |
| 10. Does the patient seem sad or depressed or has he/she reported such feelings? | ☐ | ☐ | ☐ |
| 11. Does the patient have flat moods without the normal "highs" and " lows"? | ☐ | ☐ | ☐ |
| 12. Does the patient have difficulty in experiencing pleasure from their usual activities or report that they lack pleasure? | ☐ | ☐ | ☐ |
| **SCORE:** | | | ☐ |
| **Domain 4: Perceptual problems/hallucinations** | | | |
| 13. Does the patient indicate that he/she sees things that are not there? | ☐ | ☐ | ☐ |
| 14. Does the patient have beliefs that you know are not true? (For example, about being harmed, being robbed or being unfaithful) | ☐ | ☐ | ☐ |
| 15. Does the patient experience double vision? (2 separate real objects and not blurred vision) | ☐ | ☐ | ☐ |
| **SCORE:** | | | ☐ |

**FIG. 1.** The Non-Motor Symptoms Scale for Parkinson's disease.

and stability of the measure. Internal consistency was analyzed by Cronbach's α-coefficient and item homogeneity. Criterion value for α was $\geq 0.70$[30] and $>0.30$ for item homogeneity.[31] An exploratory factor analysis by principal component factor method was carried out with orthogonal and oblique rotations. Test–retest reliability (individual, model 2, intraclass correlation coefficient, ICC) was assessed in a subset of patients ($n = 30$) who repeated the NMSS 2 weeks later. ICC values higher than 0.70 are considered acceptable.[29]

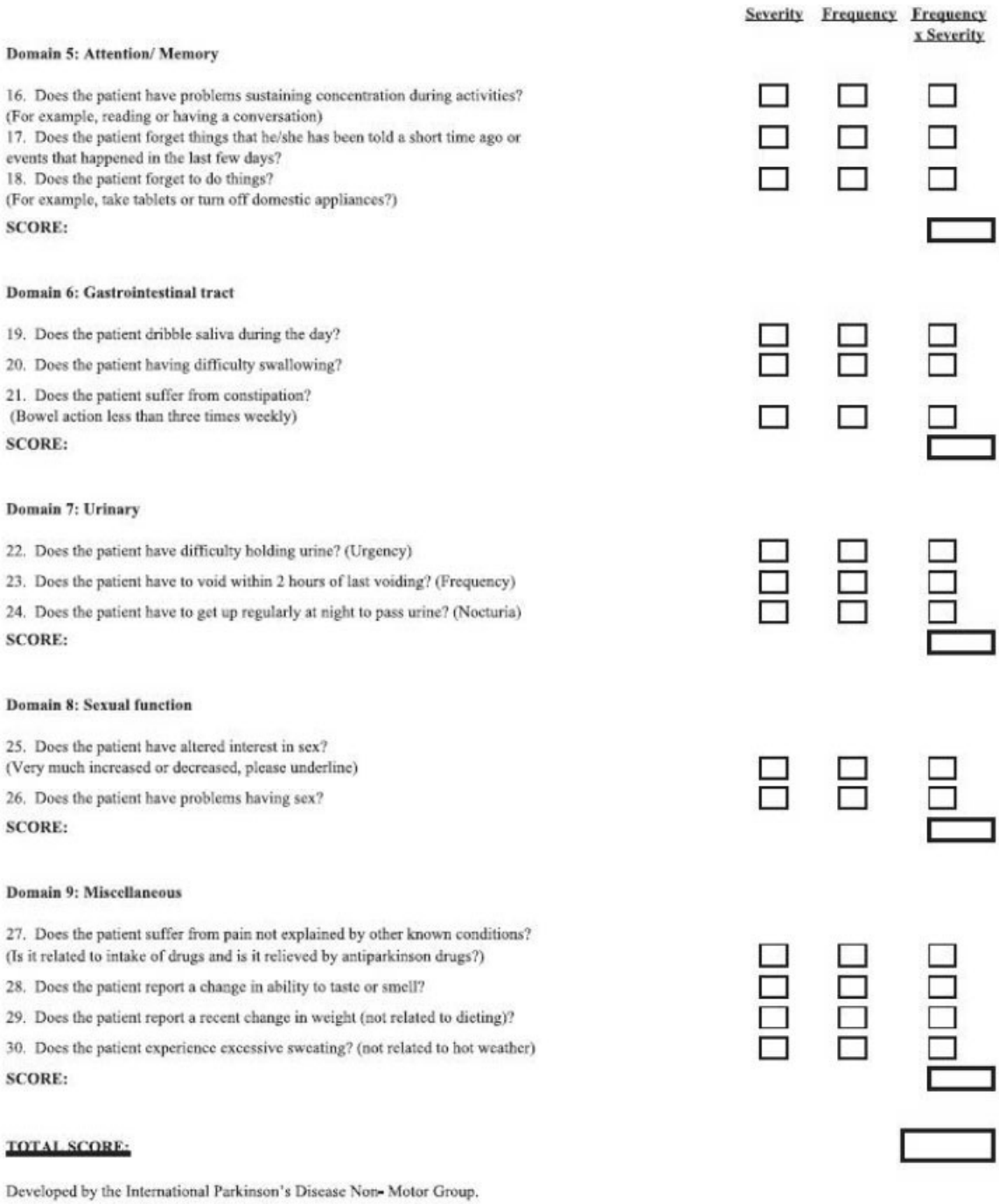



**FIGURE 1.** (Continued)

5. Construct validity. The convergent validity (correlation with other measures, which assess the same or similar constructs) and the discriminative validity (ability of the scale to differentiate known-groups) were the construct validity components tested in this study. It was hypothesized *a priori* that NMS have a progressive evolutionary course with a moderate association with duration of disease, HY, and UPDRS-Parts 3 and 4 ($r_S = 0.30-0.59$). Only some components of the NMSS are aimed at the evaluation of

**TABLE 1.** *Score distribution of the applied measures and non-motor symptoms scale*

| Measures | Mean | SD | Min. | Max. |
|---|---|---|---|---|
| UPDRS- section 3 | 17.43 | 8.16 | 2 | 47 |
| UPDRS- section 4 | 3.37 | 3.31 | 0 | 16 |
| CIRS-G index | 1.85 | 0.60 | 1 | 4 |
| FAB-total | 15.00 | 2.71 | 3 | 18 |
| PDQ-8 | 29.82 | 18.81 | 0 | 93.7 |
| Fatigue-VAS | 63.00 | 21.26 | 0 | 100 |
| HADS-anxiety | 8.70 | 5.00 | 0 | 21 |
| HADS-depression | 8.12 | 4.80 | 0 | 21 |
| NMS questionnaire | 10.47 | 5.01 | 0 | 27 |
| NMS scale | | | | |
| Cardiovascular | 2.04 | 3.60 | 0 | 21 |
| Sleep/fatigue | 12.60 | 10.35 | 0 | 48 |
| Mood/cognition | 9.94 | 13.13 | 0 | 72 |
| Perceptual problems | 1.71 | 4.03 | 0 | 36 |
| Attention/memory | 6.50 | 8.19 | 0 | 36 |
| Gastrointestinal | 5.44 | 5.85 | 0 | 26 |
| Urinary | 7.66 | 7.96 | 0 | 36 |
| Sexual function | 3.67 | 6.02 | 0 | 24 |
| Miscellaneous | 6.66 | 7.44 | 0 | 36 |
| Total NMSS | 56.46 | 40.66 | 0 | 243 |

SD, Standard deviation; Min., Minimum; Max, Maximum; UPDRS, Unified Parkinson's Disease Rating Scale; VAS, Visual analogue scale; CIRS-G, Cumulative Illness Rating Scale-Geriatrics; HADS, Hospital Anxiety and Depression Scale; FAB, Frontal assessment battery; PDQ-8, Parkinson's Disease Questionnaire- 8 items; NMS, Non-Motor Symptoms; NMSS, Non-Motor Symptoms Scale.

fatigue, mood state, and attention/memory; therefore, the overall correlation with the applied measures for these constructs, we felt, are likely to be weak ($r_S < 0.30$). Finally, it was expected a high correlation of the NMSS score with the number of positive NMSQuest responses and PDQ-8 score ($r_S \geq 0.60$) while the association of the NMS with age, age at diagnosis, and comorbidity could be weak to moderate ($r_S \leq 0.40$).[10,32] High correlations (coefficients 0.40 or higher) were expected between the corresponding domains of the NMSS (score) and NMSQuest (number of positive responses).[29]

The discriminative validity was explored by the comparison of the NMSS scores broken down by severity level based on HY staging (mild = stages 1–2.5; moderate = stage 3; severe = stages 4 and 5 (Kruskal–Wallis test).

6. Precision of the scale domains (ability of the instrument to detect small differences) was determined by means of the standard error of measurement (SEM).[30]

## RESULTS

The total number of patients recruited were 242 across centers in the UK, Italy, Germany, USA, and Japan. Five cases were excluded because of missing data on the NMSS.

The age ranged between 35 and 88 years (67.2 ± 11.1 years) while the duration of PD ranged from "just diagnosed," 0 to 45 years (6.4 ± 6.1years). Fifty-seven percentage of patients were males, and most were clustered around HY stages 2–3 (stage 2 = 19%; 2.5 = 17%; 3 = 32.5%) although patients in stage 1 (9.3%) and stage 5 (2.1%) were also included. Twenty-four patients had untreated PD at the time of assessment (age, 63.4 ± 13 years; duration of PD, 3.0 ± 4.5 years). The descriptive statistics of the applied measures are shown in Table 1.

The mean NMSS score was 56.5 ± 40.7 (median: 46; interquartile rank: 25–81.5) with a range from 0 to 243 of a possible maximum of 360. Maximum score (standardized to percentage on the maximum possible score) was recorded for six domains: sleep (3 patients); mood (2); perceptual disorders (1); attention/memory (4); urinary (4); and sexual function (5). Only one patient declared no NMS while another one scored a maximum of 243.

The overall floor and ceiling effect of the total NMSS score were 0.42% under the upper standard limit of 15%. Skewness was 1.17 (exceeding the limit +1 by a borderline 0.17 factor only). Both the exception of sleep significant proportions (17–66%) reported no problems in the other domains.

The multitrait scaling reached a success, and probable success rate was higher than 95% for all domains (Table 2), except the miscellaneous domain (47% success rate), which contained wide ranging, unrelated questions from diplopia to weight change. Twenty three of the 26 items (88.5%) contained in eight domains (not including miscellaneous) of the NMSS gained item-total correlation coefficients higher than the criterion 0.30 (Table 3). The dimension Gastrointestinal showed the weakest scaling assumptions by both methods.

Four domains (mood/cognition, attention/memory, urinary, and sexual function) showed Cronbach's α-coefficient higher than the criterion 0.70 (0.71, urinary, to

**TABLE 2.** *Multitrait scaling results*

| Domain | Number of correlations* | Success | Probable success | Failure | Succes rate (%) |
|---|---|---|---|---|---|
| Cardiovascular | 16 | 16 | 0 | 0 | 100 |
| Sleep/fatigue | 32 | 22 | 9 | 1 | 97 |
| Mood | 48 | 48 | 0 | 0 | 100 |
| Perceptual problems | 24 | 23 | 1 | 0 | 100 |
| Attention/memory | 24 | 24 | 0 | 0 | 100 |
| Gastrointestinal | 24 | 8 | 15 | 1 | 96 |
| Urinary | 24 | 23 | 1 | 0 | 100 |
| Sexual Function | 16 | 16 | 0 | 0 | 100 |
| Miscellany | 32 | 4 | 11 | 17 | 47 |

*Discounting item-own domain correlations.

**TABLE 3.** *Item-total correlation and test–retest of the Non-Motor Symptoms Scale*

| Non-Motor Symptoms Scale | Item-total correlation* | Test–retest (ICC) |
|---|---|---|
| Cardiovascular | | 0.45 |
| Light-headedness | 0.46 | 0.53 |
| Fainting | | 0.49 |
| Sleep/fatigue | | 0.92 |
| Daytime sleep | 0.32 | 0.85 |
| Fatigue | 0.47 | 0.95 |
| Difficulty falling asleep | 0.36 | 0.87 |
| Restless legs | 0.47 | 0.74 |
| Mood/cognition | | 0.96 |
| Lost interest in surroundings | 0.65 | 1.00 |
| Lack motivation | 0.68 | 0.50 |
| Feel nervous | 0.61 | 0.88 |
| Seem sad | 0.62 | 0.90 |
| Flat mood | 0.62 | 0.62 |
| Difficulty experiencing pleasure | 0.65 | 0.82 |
| Perceptual problems | | 0.83 |
| Hallucinations | 0.37 | 0.77 |
| Delusions | 0.44 | 0.86 |
| Double vision | 0.42 | 0.31 |
| Attention/memory | | 0.94 |
| Concentration | 0.54 | 0.94 |
| Forget things or events | 0.73 | 0.93 |
| Forget to do things | 0.66 | 0.87 |
| Gastrointestinal | | 0.84 |
| Saliva | 0.23 | 0.96 |
| Swallowing | 0.24 | 0.60 |
| Constipation | 0.21 | 0.99 |
| Urinary | | 0.83 |
| Urgency | 0.51 | 0.68 |
| Frequency | 0.58 | 0.64 |
| Nocturia | 0.51 | 0.94 |
| Sexual dysfunction | | 0.94 |
| Interest in sex | 0.65 | 0.78 |
| Problems having sex | | 1.00 |
| Miscellaneous | | 0.65 |
| Pains | 0.17 | 0.63 |
| Taste or smell | 0.30 | 0.95 |
| Weight change | 0.10 | 0.78 |
| Excessive sweating | 0.15 | 0.83 |

*In domains with two items, interitem correlation.
ICC, Intraclass correlation coefficient.

0.85, mood). Two other, sleep and perceptual problems reached values around 0.60 and the rest were under the limit. Mean Cronbach's α was 0.61. Item homogeneity analysis showed that all domains exceeded the accepted standard 0.30 ($I_H$ = 0.35–0.57), except sleep/fatigue ($I_H$ = 0.29), gastrointestinal ($I_H$ = 0.18), and miscellany ($I_H$ = 0.11).

Factor analysis identified nine factors that explained 63% of the variance and corresponded approximately to the *a priori* defined domains. Factor 1 was composed of six items included in mood/cognition domain; factor 2, of three items of perceptual problems dimension; factor 3, of three urinary domain items; and factor 4, of two items of cardiovascular. Factor 5 included three items (restless legs, difficulty falling asleep, and fatigue) of the four-item domain sleep/fatigue. Factor 6 was formed by the sexual function items plus weight change (miscellaneous). Factor 7 included constipation and excessive sweating. Swallowing was the only item loading on factor 8. Items comprising factor 9 were coincident with the items of attention/memory.

Twenty-three of the 30 NMSS items (76.7%) showed ICC higher than the standard 0.70 (Table 3). All dimensions gained ICC values higher than 0.80, except cardiovascular (ICC = 0.45).

NMSS score was not significantly associated with age and only a weak association was found with age at diagnosis ($r_S = -0.16$, $P < 0.05$) and duration of disease ($r_S = 0.23$, $P < 0.001$). Correlation of the NMSS score with measures related to PD (HY, UPDRS-Parts 3 and 4) resulted low to moderate ($r_S$ = 0.33–0.49, $P < 0.0001$). As a whole, they were also low with VAS of fatigue, comorbidity index, HADS, and FAB ($r_S$ = 0.23–0.31; $P < 0.001$–$P < 0.0001$) and high with NMS Quest and PDQ-8 (both, $r_S = 0.70$, $P < 0.0001$) (Table 4).

The correlation of NMSS domains with independent measures of related constructs resulted as follows: sleep/fatigue with UPDRS-Part 4, $r_S = 0.51$; sleep/fatigue with fatigue VAS, $r_S = -0.26$; mood/cognition with HADS-anxiety and depression, $r_S = 0.31$ and 0.35, respectively; mood/cognition with FAB, $r_S = -0.28$; and attention/memory with FAB, $r_S = -0.24$.

As foreseen, convergent validity between corresponding domains contained in the NMSS and NMSQuest reached values of 0.44 (NMSS-mood/cognition with

**TABLE 4.** *Convergent validity of the Non-Motor Symptoms Scale*

| Variable | Spearman r* |
|---|---|
| Hoehn and Yahr staging | 0.33 |
| UPDRS- section 3 | 0.35 |
| UPDRS- section 4 | 0.49 |
| UPDRS4-dysk & flct. | 0.43 |
| UPDRS4-other compl. | 0.45 |
| Fatigue-VAS | −0.30 |
| CIRS-G index | 0.27 |
| HADS-anx | 0.23[a] |
| HADS-dep | 0.28 |
| FAB-total | 0.31 |
| NMSQ-total | 0.70 |
| PDQ-8 | 0.70 |

*All, $P < 0.0001$, except [a]$P < 0.001$.
UPDRS, Unified Parkinson's Disease Rating Scale; Dysk & Flct, Dyskinesias and fluctuations; Other compl, Other complications; VAS, Visual analog scale; CIRS-G, Cumulative Illness Rating Scale-Geriatrics; HADS, Hospital Anxiety and Depression Scale; FAB, Frontal assessment battery; NMSQ, Non-Motor Symptoms Questionnaire; PDQ-8, Parkinson's Disease Questionnaire- 8 items.

**TABLE 5.** *Precision (sensitivity) of the Non-Motor Symptoms Scale*

| Domain | SEM | SD/2 |
|---|---|---|
| Cardiovascular | 2.71 | 1.80 |
| Sleep/fatigue | 2.93 | 5.17 |
| Mood/cognition | 2.62 | 6.56 |
| Perceptual problems | 1.66 | 2.01 |
| Attention/memory | 2.01 | 4.10 |
| Gastrointestinal | 2.34 | 2.92 |
| Urinary | 3.28 | 3.98 |
| Sexual function | 1.47 | 3.01 |
| Miscellany | 2.90 | 3.72 |
| Non-Motor Symptoms Scale | 7.04 | 20.33 |

SEM, Standard error of the measure; SD/2. Half a standard deviation.

NMSQuest-depression/anxiety) to 0.74 (both, NMSS and NMS Quest, and sexual dysfunction).

NMSS scores, stratified for severity levels of disease, resulted 46.1 ± 34.5 for mild disease; 61.0 ± 37.5 for moderate, and 87.5 ± 53.0 for severe disease patients (Kruskal–Wallis test, $P < 0.0001$).

SEM for NMSS domains ranged from 1.47 (sexual dysfunction) to 3.28 (urinary). For the total NMSS, SEM was 7.04 (Table 5).

NMSS scores were not significantly different by gender or ethnic group. Akinetic subtype of PD had higher NMS scores (Kruskal–Wallis test, $P = 0.05$).

## DISCUSSION

We describe the development and evaluation of a new scale (Fig. 1) for the comprehensive assessment of NMS in PD patients. The scale is relatively easy to administer, requiring the investigator to administer a total of 30 questions to the patient to be scored for both severity and frequency. The NMSS alone takes ~10–15 min to administer. Data from nine domains provide information on NMS ranging from gastrointestinal, sexual, and sleep dysfunction to cognitive problems. In spite of the complexity of construct, the scale appears robust, reproducible, and has acceptable clinimetrics. There is a significant relationship between NMSS score and severity of disease based on HY, NMSQuest, and health-related quality of life assessments.

This is the first study providing information on the metric attributes of the NMSS with data captured on a PD population from countries in three continents. As usual in clinical samples, intermediate stages of disease were overrepresented, but patients in all stages of disease were included. Furthermore, we did not use cut-off scores on cognitive tests to define dementia. A patient may therefore have cognitive impairment and still be included and indeed data from FAB and corresponding domains of NMSQuest and NMSS suggest this. However, this highlights that we have included a "real life" patient population as indeed, we would expect a significant proportion to show cognitive impairment. The results should therefore be generalizable to the broad range of patients except those with the most severe cognitive impairment.

Quality of data was very satisfactory, with over 99% of cases fully computable. Score distribution covered the complete range of scoring for six domains and for the total NMSS the difference between mean and median was low (4.3% of the maximum observed score). There were no floor or ceiling effects for the NMSS total score and the skewness was only slightly higher than the standard limit. NMSS is a complex scale assessing quite different and often unlinked symptoms not necessarily related to each other constructs. Therefore, we expect that a proportion of patients will experience symptoms in one or more domains, but rarely in all the included domains. This explains why the less prevalent symptoms produce the highest floor effects for their domains. However, the key aim of the scale is to explore the wide range of NMS that may occur in a patient with PD and provide a holistic assessment. The diversity of symptoms explored and their relative prevalence (only one patient had no NMS) makes it possible to obtain a unified score free of floor effect.

Scaling assumptions, that test whether the scale items are correctly grouped and if their scores can be directly summed to produce a total score, were satisfactory (success rate, 96–100%), except for the miscellany domain (47%) (Table 2). This latter result was expected since the individual components of this domain (pains, taste/smell, weight change, and sweating) are, in principle, unrelated to each other. However, we feel these items need to be included as the NMSQuest and other studies have indicated the importance of these symptoms to the patients.[10,31,32] In the NMSQuest study, diplopia and weight change was cited as important by 21.9 and 22% of an unselected cohort of PD patients across all stages.[10]

As a whole, internal consistency was acceptable for a scale of this type, with multiple domains and a short number of items per domain (mean N = 3). The most widely used coefficient, Cronbach's α, reflects intercorrelations between items but it is also dependent on the number of items and even on the type of assessed construct.[28,29,33] To ensure that the NMSS structure was robust, factor analysis was performed and it supported, globally, the current structure of the NMSS, although the items of gastrointestinal domain loosely loaded on several factors. Gastrointestinal dimension showed weakness in consistency, reaching values under the threshold

value also for the item-total correlation. However, the domain is maintained because it contains relevant symptoms in the digestive area: saliva dribbling, dysphagia, and constipation, cited in the NMSQuest study by patients to be important and relevant.[10] In addition to the statistical analysis, other criteria (for instance, expert criterion or relevance for the target population) are also regarded as a guide to the construction of a health measure. Thus, the importance of symptoms to patients in real life has been taken into account independent of the statistical results.[34,35] Nonetheless, we envisage that similar to many other scales, future revisions of the measure may improve the design of the NMSS components.

A test–retest assessment suggested that the NMSS reliability was satisfactory except for cardiovascular items and domain, where a single patient response ranging from 0 (1st assessment) to 12 (2nd assessment) induced a dramatic fall of the ICC. The cause of this large change in a single question remains unclear and as the number of patients participating in the test-retest evaluation was relatively small ((n = 30), we cannot draw any conclusions on this finding.[35]

We hypothesized that the NMSS will show a high association with the number of positive responses both on the NMSQuest and with the PDQ-8 (Table 4). A similar structure in contents explains the high correlation between NMSS and NMSQuest. PDQ-8 is a validated measure of health-related quality of life in PD and it is recognized that some NMS such as mood disorders, sleep disturbances, pain, and fatigue are determinants or strongly associated with health-related quality of life of patients with PD.[2–4,7,8,36-38] Our results therefore, support these observations highlighting the importance of NMS for self-reported health status as measured by PDQ-8 in PD.

The correlation with duration of disease was low, as also reported in the NMSQuest study.[10] It may be anticipated that the diversity of domains included in the NMSS will preclude an homogeneous linear progression over time. As in the NMSQuest, NMSS score correlated moderately with HY stages, as well as motor examination and motor complications (Table 4),[10] Total NMSS score increased at each level of disease severity based on HY stages similar again to the NMSQuest study.[10] Therefore, we conclude that progression in motor impairment and disability is accompanied by increasing burden of NMS as reflected by the NMSQuest and NMSS studies.

SEM provides an indicator of scale precision.[30,39] Both 1 SEM and ½ standard deviation of data can be regarded as the "minimal important difference" representative of a real change.[40,41] We have thus cited both values for the NMSS domains and total to gain some information on potential responsiveness of the scale (Table 5). In the domains where the reliability index (ICC) was high, the SEM was lower than half a standard deviation and vice versa.

There are several potential limitations of this study. These are discussed as follows:

1. First, the lack of NMSS testing with other measures to cover all dimensions represented in the NMSS (for instance, with scales for measurement of sleep, sexual function, and gastrointestinal disorders) could be considered a weakness. Methodologically, this would be difficult to perform in one study because of likely patient fatigue trying to complete a large number of scales within a limited time frame in clinic. In this explorative study, therefore, we have concentrated on a few specific well-established scales such as the UPDRS and some others, which would be practicable to use without causing undue strain on the patient and excessive respondent and administrative burden.
2. Another potential criticism of this study was the lack of normal controls to determine the difference between this population and PD patients. However, we felt that the ability of the NMSS to discriminate PD from non-PD or controls is not crucial during the testing of NMSS at a pilot stage. This scale is not intended to be a diagnostic scale but is an evaluation scale. Therefore, we are not attempting to distinguish the symptoms from normal subjects but intending to measure the NMS, which have been indicated by patients completing the NMSQuest as important to them. As an example, in the NMSQuest study, pain was not significantly different between PD and controls but it will still be important to be able to measure pain in PD, an important NMS.[7] We envisage a smaller second-stage study during which health controls could be included to characterize what is "pathological" and what is a function of the normal ageing process.
3. Finally, the small size of the sample for test-retest could also be considered a limitation.

The issue of the setting where NMSS can be used merits discussion. It has become clear that a comprehensive assessment of PD is an unmet need, much of it related to lack of assessment of NMS along with the motor state in a single practical scale, which provides a reasonable overview of the burden of NMS in PD. We feel that the benefits of a specific scale for NMS would include the ability to identify accurately NMS in patients, follow their evolution during disease progression and possibly, determine their response to treatment. The

strong correlation of NMSS with advancing disease stages of PD supports the above proposal. It can be argued that following identification of NMS using NMSQuest, in depth specific validated scales such as depression or sleep scales can be applied instead of NMSS. However, such a process is likely to be very time consuming and impractical in the clinic. The NMSQuest study identified an average of 10 NMS per patient across all disease stages and as such several disease specific scales may thus need to be applied if these symptoms are to be explored in detail. This can be avoided by the use of NMSS, which captures a range of relevant NMS within one scale.

Evidence-based treatment of NMS in PD is urgently needed, but in part its development will depend not just on the identification of NMS but also on the quantification of the effects of treatment on patients' baseline disability. This requires the use of validated assessment. Existing PD-specific rating scales largely concentrate on motor symptoms and a revised movement disorders society sponsored version of the UPDRS (MDS-UPDRS) will include some screening questions on NMS.[42,43] Individual aspects of non-motor disability in PD are also an integral part of Scales for Outcomes in Parkinson's disease (SCOPA) project.[44-46] The NMSS, however, enables a comprehensive assessment of the range of NMS that occur in PD, both for the identification of problems and measurement of intervention outcome. No such instrument exists to date and the NMSS provides the first step for the health care professional with an objective and reliable tool to grade and rate NMS. Although future studies on the NMSS may lead to revisions to improve its metric properties, until then the NMSS may be considered a comprehensive, useful measure for evaluation of NMS in PD patients.

**Acknowledgments:** We thank all members of the PDNMG, a nonprofit multidisciplinary academic group dedicated to raising awareness of non-motor symptoms of PD, our coordinators in the UK, USA, Italy, and Israel and the UK Parkinson's Disease Society for support. The additional member of the PDNMG apart from the authors cited is Professor K Bhattacharya. The study was also supported by unrestricted educational grants from Pfizer, GSK, Boehringer-Ingelheim, and Britannia Pharamaceuticals.

## REFERENCES

1. Braak H, Del Tredici K, Rüb U, de Vos RAI, Jansen Steur ENH, Braak E. Staging of brain pathology related to sporadic Parkinson's disease. Neurobiol Aging 2003;24:197-210.
2. Chaudhuri KR, Healy D, Schapira AHV. The non-motor symptoms of Parkinson's disease. Diagnosis and management. Lancet Neurol 2006;5:235-245.
3. Hely MA, Morris JGL, Reid WGJ, Trafficante R. Sydney multicenter study of Parkinson's disease: non-L-dopa-responsive problems dominate at 15 years. Mov Disord 2005;20:190-199.
4. Schrag A, Jahanshahi M, Quinn N. What contributes to quality of life in patients with Parkinson's disease? J Neurol Neurosurg Psychiatry 2000;69:308-312.
5. Findley L, Aujla M, Bain PG, et al. Direct economic impact of Parkinson's disease: a research survey in the United Kingdom. Mov Disord 2003;18:1139-1145.
6. Hagell P, Nordling S, Reimer J, Grabowski M, Persson U. Resource use and costs in a Swedish cohort of patients with Parkinson's disease. Mov Disord 2002;17:1213-1220.
7. Global Parkinson's Disease Survey Steering Committee. Factors impacting on quality of life in Parkinson's disease: results from an international survey. Mov Disord 2002;17:60-67.
8. Clarke CE, Zobkiw RM, Gullaksen E. Quality of life and care in Parkinson's disease. Br J Clin Pract 1995;49:288-293.
9. Shulman LM, Taback RL, Rabinstein AA, Weiner WJ. Non-recognition of depression and other non-motor symptoms in Parkinson's disease. Parkinsonism Relat Disord 2002;8:193-197.
10. Chaudhuri KR, Martinez-Martin P, Schapira AHV, et al. An international multicenter pilot study of the first comprehensive self-completed non motor symptoms questionnaire for Parkinson's disease: the NMSQuest study. Mov Disord 2006;21:916-923.
11. Hughes AJ, Daniel SE, Kilford L, Lees AJ. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. J Neurol Neurosurg Psychiatry 1992;55:181-184.
12. Hoehn M, Yahr MD. Parkinsonism: onset, progression and mortality. Neurology 1967;17:427-442.
13. Fahn S, Elton RL, Members of the UPDRS Development Committee. Unified Parkinson's disease rating scale. In: Fahn S, Marsden CD, Calne DB, Goldstein M, editors. Recent developments in Parkinson's disease, Vol. 2. Florham Park, NJ: Macmillan Health Care Information; 1987. p 153-164.
14. Miller MD, Paradis CF, Houck PR, et al. Rating chronic medical illness burden in geropsychiatric practice and research: application of the cumulative illness rating scale. Psychiatry Res 1992;41:237-248.
15. Visser M, Marinus J, van Hilten JJ, Schipper RGB, Stiggelbout AM. Assessing comorbidity in patients with Parkinson's disease. Mov Disord 2004;19:824-828.
16. Linn BS, Linn MW, Gurel L. Cumulative illness rating scale. J Am Geriatr Soc 1968;16:622-626.
17. Dubois B, Slachevsky A, Litvan I, Pillon B. The FAB. A frontal assessment battery at bedside. Neurology 2000;55:1621-1626.
18. Jenkinson C, Fitzpatrick R, Peto V, Greenhall R, Hyman N. The PDQ-8: development and validation of a short-form Parkinson's disease questionnaire. Psychol Health 1997;12:805-814.
19. Dittner AJ, Wessely SC, Brown RG. The assessment of fatigue: a practical guide for clinicians and researchers. J Psychosom Res 2004;56:157-170.
20. Martinez-Martin P, Catalan MJ, Benito-Leon J, et al. Impact of fatigue in Parkinson's disease: the fatigue impact scale for daily use (D-FIS). Qual Life Res 2005;15:597-606.
21. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983;67:361-370.
22. Marinus J, Leentjens AF, Visser M, Stiggelbout AM, van Hilten JJ. Evaluation of the hospital anxiety and depression scale in patients with Parkinson's disease. Clin Neuropharmacol 2002;25:318-324.
23. Gulati A, Forbes A, Stegie F, Kelly L, Clough C, Chaudhuri KR. A clinical observational study of the pattern and occurrence of non-motor symptoms in Parkinson's disease ranging from early to advanced disease. Mov Disord 2004;19 (Suppl 9):S406 (Abstract).
24. Clayton L, Martinez-Martin P, Odin P, et al. Prevalence of non motor symptoms in Parkinson's disease: an international survey using NMSQuest in 525 patients. Park Relat Disord 2006;12 (Suppl 1):I.P2 (Abstract).
25. Smith SC, Lamping DL, Banarjee S, et al. Measurement of health-related quality of life for people with dementia: development of a new instrument (DEMQOL) and an evaluation of current methodology. Health Technol Assess 2005;9:16-19.

26. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? Qual Life Res 1995;4:293-307.
27. Hays RD, Anderson R, Revicki D. Psychometric considerations in evaluating health-related quality of life measures. Qual Life Res 1993;2:441-449.
28. Nunnally JC, Bernstein IH. Psychometric theory, 3rd ed. New York: McGraw-Hill; 1994. p 304-308
29. Fayers PM, Machin D. Quality of life. Assessment, analysis and interpretation. Chichester: Wiley; 2000. p 45-90
30. Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality-of-life instruments: attributes and review criteria. Qual Life Res 2002;11:193-205
31. Eisen M, Ware JE, Jr, Donald CA, Brook RH. Measuring components of children's health status. Med Care 1979;17:902-921.
32. Fisk JD, Brown MG, Sketris IS, Metz LM, Murray TJ, Stadnyk KJ. A comparison of health utility measures for the evaluation of multiple sclerosis treatments. J Neurol Neurosurg Psychiatry 2005; 76:58-63.
33. McDowell I. Measuring health. A guide to rating scales and questionnaires, 3rd ed. Oxford: Oxford University Press; 2006. p 10-54.
34. 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.
35. Terwee CB, Bot SDM, de Boer MR, et al. J Clin Epidemiol 2007;60:34-42.
36. 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.
37. Larsen JP, Karlsen K, Tandberg E. Clinical problems in non-fluctuating patients with Parkinson's disease: a community-based study. Mov Disord 2000;15:826-829.
38. Kuopio AM, Marttila RJ, Helenius H, Toivonen M, Rinne UK. The quality of life in Parkinson's disease. Mov Disord 2000;15:216-223.
39. Streiner DL, Norman GR. Health measurement scales. A practical guide to their development and use, 3rd ed. Oxford: Oxford University Press; 2003. p 142-143.
40. Wyrwich KW, Bullinger M, Aaronson N, Hays RD, Patrick DL, Symonds T, and The Clinical Significance Consensus Meeting Group. Estimating clinically significant differences in quality of life outcomes. Qual Life Res 2005;14:285-295.
41. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life. The remarkable universality of half a standard deviation. Med Care 2003;41:582-592.
42. Goetz CG, Fahn S, Martinez-Martin P, et al. Movement disorders society sponsored revision of the unified Parkinson's disease rating scale (MDS-UPDRS): process, Format and clinimetric testing plan. Mov Disord 2007;22:41-47.
43. Dodel RC, Dubois B, Fahn S, et al. Addressing non-motor impairments in Parkinson's disease: the new version of the UPDRS. Mov Disord 2005;20 (Suppl 10):S83 (p 277) (Abstract).
44. Marinus J, Visser M, Verwey NA, et al. Assessment of cognition in Parkinson's disease. Neurology 2003;61:1222-1228.
45. Marinus J, Visser M, van Hilten JJ, Lammers GJ, Stiggelbout AM. Assessment of sleep and sleepiness in Parkinson disease. Sleep 2003;26:1049-1054.
46. Visser M, Marinus J, Stiggelbout AM, Van Hilten JJ. Assessment of autonomic dysfunction in Parkinson's disease: the SCOPA-AUT. Mov Disord 2004;19:1306-1312.