




Non-Motor Fluctuations in Parkinson's Disease: Validation of the Non-Motor Fluctuation Assessment Questionnaire

Galit Kleiner, MD,^{1,2*} Hubert H. Fernandez, MD,³ Kelvin L. Chou, MD,⁴  Alfonso Fasano, MD, PhD,^{5,6} Kevin R. Duque, MD,⁷  Diana Hengartner, MD,³ Albie Law, BA, CCRA,^{1,2} Adam Margolius, MD,³ Yu-Yan Poon, RN,⁵ Michel Sáenz Farret, MD,⁵ Philip Saleh, MD, MSc,^{1,2} Joaquin A. Vizcarra, MD,⁷ Glenn T. Stebbins, PhD,⁸ Alberto J. Espay, MD, MSc,⁷  and PSG NoMoFA Study Group

¹Jeff and Diane Ross Movement Disorders Clinic/ATC, Baycrest Center for Geriatric Health Toronto, Toronto, Ontario, Canada

²Division of Neurology, University of Toronto, Toronto, Ontario, Canada

³Department of Neurology and Center for Neurological Restoration, Neurological Institute, Cleveland Clinic Lerner College of Medicine, Cleveland, Ohio, USA

⁴Department of Neurology and Neurosurgery, University of Michigan Medical School, Ann Arbor, Michigan, USA

⁵Edmond J. Safra Program in Parkinson's Disease, Morton and Gloria Shulman Movement Disorders Clinic, Toronto Western Hospital, UHN, Toronto, Ontario, Canada

⁶Krembil Brain Institute, Toronto, Ontario, Canada

⁷James J. and Joan A. Gardner Center for Parkinson's Disease and Movement Disorders Department of Neurology, University of Cincinnati, Cincinnati, Ohio, USA

⁸Department of Neurological Sciences, Rush University Medical Center, Chicago, Illinois, USA

ABSTRACT: Background: In patients with Parkinson's disease (PD), sleep, mood, cognitive, autonomic, and other non-motor symptoms may fluctuate in a manner similar to motor symptoms.

Objectives: To validate a final version of a patient-rated questionnaire that captures the presence and severity of non-motor fluctuations in levodopa-treated PD patients (NoMoFA).

Methods: We recruited PD subjects from five movement disorders centers across the US and Canada. We assessed the internal consistency, floor and ceiling effects, test-retest reliability, and concurrent validity of NoMoFA. Classical test theory and item response theory methods informed item reduction and Delphi process yielded a final questionnaire.

Results: Two hundred subjects and their care-partners participated in the study (age: 66.4 ± 9.6 years; disease duration: 9 ± 5.5 years; median Hoehn and Yahr [H&Y] OFF: 3 [range 1–5]; mean Unified Parkinson's Disease Rating Scale (UPDRS) III ON score: 27.4 ± 14.9).

Acceptability of the scale was adequate. There were floor effects in 8/28 items. Cronbach's alpha was 0.894. While eight items had "item-to-total" correlations below the cutoff of 0.4, removing these items did not improve Cronbach's alpha. Test-retest reliability was acceptable (intraclass correlation coefficient [ICC] 0.73; 95% confidence interval, 0.64–0.80). Concurrent validity was adequate with all Spearman's rho values comparing NoMoFA score to other measures of parkinsonian severity showing significance and in the expected direction. A final Delphi panel eliminated one item to avoid redundancy.

Conclusions: The final 27-item self-administered NoMoFA is a valid and reliable questionnaire, capturing both static and fluctuating non-motor symptoms in PD. © 2021 International Parkinson and Movement Disorder Society

Key Words: non-motor fluctuations; non-motor symptoms; questionnaire; validation; Parkinson's disease

*Correspondence to: Dr. Galit Kleiner, Jeff and Diane Ross Movement Disorders Clinic/ATC, Baycrest Center for Geriatric Health, 3560 Bathurst Street, Toronto, Ontario M6A 2E1, Canada; E-mail: galit.kleiner@assistivetechneologyclinic.ca

Relevant conflicts of interest/financial disclosures: Nothing to report.

Funding agencies: Sunovion Pharmaceuticals sponsorship through the Parkinson Study Group.

Received: 24 August 2020; **Revised:** 29 December 2020; **Accepted:** 3 January 2021

Published online 15 February 2021 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/mds.28507

Non-motor fluctuations (NMF) are the dynamic subset of non-motor symptoms (NMS). NMF have increasingly been recognized as important and disabling symptoms of Parkinson's disease (PD).^{1,2} NMS may fluctuate in parallel with motor symptoms and in relationship to plasma dopamine concentration, although the exact mechanism of NMF remains speculative.^{3,4} It is important to differentiate static from fluctuating NMS as they may have different pathophysiology and response to treatment.

Recent recognition of the impact of NMS on quality of life has driven efforts to create valid, reliable questionnaires to capture and quantify NMS.⁵⁻⁷ Some of these questionnaires have included items probing selected non-motor OFF symptoms but none capture or quantify the entire spectrum of NMF in both the ON- and the OFF-medication conditions. We sought to create and validate a patient-rated questionnaire with the goal of reliably capturing and quantifying NMF to meet clinical and research needs.

Methods

Initial Scale Development

An initial version of the Non-Motor Fluctuation Assessment (NoMoFA) questionnaire⁸ was created following the US Food and Drug Administration (FDA) guidance for development of patient-reported outcome (PRO) instruments.⁹ We used qualitative research methodology of patient-based nominal group technique and focus groups with framework analyses¹⁰ to identify NMS susceptible to fluctuations impacting the patient's function. A Delphi panel¹¹ composed of members from the Quality Standards Subcommittee on Non-Motor Symptoms in PD of the American Academy of Neurology¹² reviewed all available information (nominal group, focus group, literature review, and professional experience) to develop an initial list of 28 items for the scale. The scoring of this initial scale identified a given NMS and asked if it was related to when levodopa was working (ON) or not (OFF), and rating the severity of the symptom (mild, moderate, severe) if ON or OFF.

Initial Examination of Clinimetric Properties

The clinimetric properties of this first version of NoMoFA were initially examined in a sample of 129 patients with PD, recruited from three specialized clinics, who completed the questionnaire on their own or with the assistance of a care-partner if needed. The study cohort's mean age was 65.2 ± 10.1 years, had a disease duration of 6.3 ± 4.2 years, were 58% men, and took 812.2 ± 418.9 mg levodopa-equivalent daily dose (LEDD).¹³

Basic clinimetric properties of this scale were examined including internal consistency (Cronbach's

alpha)¹⁴ and item-to-total correlation. The thresholds for acceptable alpha was ≥ 0.80 ¹⁵ and for acceptable item-to-total correlation ≥ 0.40 .¹⁶ Multiple deficiencies in clinimetrics were noted, with ON items particularly poor for Cronbach's alpha (0.733) and low item-to-total correlations (< 0.40).

It was determined that the item scaling and the scoring scheme for the scale were contributing to these poor clinimetric properties. The Delphi panel then modified the item scaling and the scoring scheme to first ask subjects if the behavior was present, then rated the severity of the behavior if present, and then indicated if the behavior was worse when ON, when OFF, or if there was no difference between ON and OFF. This modification was piloted in a sample of 60 patients with the corresponding analysis demonstrating improved Cronbach's alpha and item-to-total correlations.

The scale was then subjected to cognitive interviewing utilizing six new patient/care-partner dyads. During the cognitive interviews, participants were asked to review the questionnaire for item relevance, ease of understanding, and whether response choices were both exhaustive and mutually exclusive. Using the results of the cognitive interview, a penultimate questionnaire was finalized in preparation for further validation.

This penultimate NoMoFA was composed of 28 items: loss of train of thought, distraction, disorientation, difficulty planning, confusion, word-finding difficulty, excessive worry, fear, restlessness, hopelessness, loneliness/isolation, hallucinations, poor decision making, impulsivity, compulsiveness, poor short-term memory, difficulty handling stressful situations, apathy, low energy/fatigue, excessive daytime sleepiness, pain, altered sensations (numbness, tingling), shortness of breath, changes in vision, excess sweating, palpitations, urinary symptoms, and constipation.

Sample Size

Based on a sample size of 5 to 10 subjects for each scale item¹⁷ it was determined that a sample size of between 140 and 280 was necessary for examining the 28 items of the NoMoFA. Feasibility permitted target recruitment of 200, well above the suggested minimum subject-to-item ratio of 5.

Patients

English-speaking patients, with a diagnosis of idiopathic PD as per Movement Disorder Society (MDS) clinical diagnostic criteria,¹⁸ and an available care-partner, were invited to participate from five Movement Disorders Centers across the US and Canada between September 2018 and September 2019. We sought patients with mild, moderate, and advanced symptoms to include the entire spectrum of PD severity.

Exclusion criteria included atypical parkinsonism and lack of English language fluency of either the patient or care-partner. While we did not impose a cognitive cut-off so as to not systematically exclude a significant portion of the population affected by NMF, we obtained cognitive scores using the Montreal Cognitive Assessment (MoCA)¹⁹ battery for reference. If a subject had cognitive impairment but a care-partner was able to assist in responding to the questionnaire, the combined response from the subject and care-partner dyad was included. If disagreement arose in those questionnaires with joint subject and care-partner responding, the subject's answer was prioritized.

Study Procedures

The study was approved by the institutional review boards of each participating center and all participants provided signed informed consent. The consent process included obtaining consent from the caregiver and assent from the patient if they were deemed unable to provide consent due to cognitive impairment. Consecutive idiopathic PD patients presenting to each of the participating sites for their regular clinic appointments were screened for use of levodopa and fluctuating PD symptoms (both motor and non-motor) and invited to participate.

Subjects were recruited across the spectrum of disease severity (mild, Hoehn and Yahr [H&Y] stages 1–2; moderate, H&Y stage 3; severe, H&Y stages 4–5). Periodic review of enrolled study subjects facilitated the prioritization of underrepresented severity categories.

Identified subjects were scheduled to return in their best “ON” state to optimally respond to questions (~1 hour after taking a dose of levodopa). Following provision of informed consent, the study team administered/collected the following data: the Movement Disorder Society-sponsored Unified Parkinson's Disease Rating Scale (MDS-UPDRS part III)²⁰ and the H&Y staging,²¹ baseline demographic information (age at diagnosis, gender), ON and OFF Schwab and England (S&E) activities of daily living scale, disease duration, medication dosage quantified as LEDD,¹³ and the MoCA¹⁹ battery. The treating movement disorders physician was queried as to their impression regarding the presence and severity of the subject's motor fluctuations (MF) and NMF (mild, moderate, severe).

The package of questionnaires included the NoMoFA questionnaire, MDS-UPDRS Part Ia Ib, II, and IV, Scales for Outcomes in Parkinson's Disease-Autonomic Dysfunction (SCOPA-AUT),²² Parkinson's Disease Questionnaire-8 (PDQ-8),²³ and Wearing-Off Questionnaire-9 (WOQ-9).²⁴ Questionnaires were packaged in random order that varied across administrations to nullify any potential order-effect. A subgroup of subjects were asked to complete the NoMoFA

questionnaire a second time 1 week following the baseline visit and were provided with a self-addressed envelope to send it back to the investigators. On this second NoMoFA questionnaire, the subjects were asked to rate if they had experienced any change in their PD symptoms (motor and non-motor) and/or in their PD medications since the initial assessment 1 week earlier to ensure stability of symptoms and medications using the Clinical Global Impression Scale.²⁵

Scoring of NoMoFA

Endorsement of any item as present required scoring its severity as mild, moderate, or severe (range 1–3). The *total NoMoFA score* was calculated as the sum of the severity scores for all items endorsed (total possible score: 28 items \times 3 = 84). If the endorsed NMS item was worse in the ON or OFF state, that severity score was recorded in the “NMF ON” or “NMF OFF” column, respectively. If the endorsed NMS item did not fluctuate, it was recorded in the “NMS no difference column”. The *total NMF subscore ON* was the sum of all the “NMF ON” scores; the *total NMF subscore OFF* the sum of all the “NMF OFF” scores. The “total NMF subscore” was the sum of “NMF ON” and “NMF OFF” subscores. The total (static) NMS score was the sum of all “NMS no difference” scores. The total NoMoFA score was calculated using the following formula: Total NoMoFA = Total NMF (ON + OFF) + Total (static) NMS.

Statistical Analysis

Subject demographics and disease-related characteristics were examined using parametric and non-parametric analyses, as appropriate. Testing of clinimetric properties was conducted using both classical test theory (CTT)²⁶ and item response theory (IRT).²⁷ CTT analyses examined *data quality* for missing values and potential floor and ceiling effects defined as skewness outside of the range -2.00 to $+2.00$,²⁸ *internal consistency* as determined by Cronbach's alpha,¹⁴ with a minimum required alpha of 0.85,¹⁵ item-to-total correlation ≥ 0.40 as minimal acceptable correlation,¹⁶ and *construct validity* using exploratory factor analyses to determine the number and types of constructs with a minimum loading of 0.40²⁹ used as a criterion for factor relevance. Item redundancy was assessed by item loading on multiple factors. Dual loading criteria was set at 0.40. IRT analyses using maximum likelihood parameter estimation³⁰ examined item discrimination (criterion of ≥ 1.00)³¹ and item threshold.³² *Test-retest reliability* was assessed in a sample of 160 patients tested over 1 week (± 2.3 days) using intraclass correlation coefficients (ICC), with a criterion of ≥ 0.70 to indicate adequate stability.²⁸ *Concurrent validity* was assessed using Spearman's rank-order

correlation coefficient (rho) for assessing the relationship between the NoMoFA and MDS-UPDRS I, II, III, IV, WOQ-9, PDQ, MF score (as rated by physician), NMF score (also as rated by physician), SCOPA-AUT, S&E ON and OFF, and H&Y ON and OFF. We chose Spearman’s rank-order correlation coefficient due to the underlying ordinal nature of the data. Frequency of endorsed NMS and NMF was calculated as the percentage of total cases.

All statistical analyses were conducted using Mplus 8.2 (www.statmodel.com), R “mrit” package for the IRT analyses, and the remainder of the analyses were conducted in SPSS Version 24 (IBM).

Results

Subjects

The packet of questionnaires was completed by 145 subjects without assistance, 35 subjects with the assistance of their care-partner, and two care-partners without assistance from the subject. Data on who was the respondent were missing from 18 questionnaires. On average it took ~10–20 minutes to complete the NoMoFA questionnaire. Study subject baseline information is included in Table 1.

The proportion of static and NMS and NMF are included in Table 2.

Examination of Clinimetric Properties

Data quality and acceptability characteristics indicated that the scale was well tolerated, with eight subjects missing one item and five subjects missing two items when answering the questionnaire. Eight items demonstrated potential floor effect as evidenced by skewness values >2.00. Examination of internal consistency indicated an acceptable level of reliability (Cronbach’s alpha = 0.89) and 20 items had adequate item-to-total correlations ($r \geq 0.40$) (Table 3). Eight items fell below this cutoff, including Disorientation, Pain, Altered Sensations, Shortness of Breath, Changes in Vision, Excess Sweating, Urinary Symptoms, and Constipation. Removing these items did not increase the overall Cronbach’s alpha.

Construct validity resulted in parsimonious two-factor solution, with adequate sampling and sphericity (KMO adequacy = 0.856, Bartlett’s test = $\chi^2 = 1809.62$, $P < 0.0005$), accounting for approximately 37% of the variance. There were no items with dual loadings ≥ 0.40 and two items that that did not load on either factor (Altered Sensations and Excess Sweating).

IRT discrimination and threshold values demonstrated adequate discrimination by most items. Five items fell below the criterion of ≤ 1.00 (Disorientation, Hallucinations, Pain, Altered Sensations, and Shortness

TABLE 1. Baseline study subject characteristics

Subjects (N)	200
Age (mean [y], SD)	66.4 ± 9.6
Gender	54% men, 46% women
Disease duration (mean [y], SD)	9 ± 5.5
LEDD (mean, SD)	1102.18 ± 650.16 mg
MoCA score (mean, SD)	25 ± 4
H&Y ON	Mild (H&Y 0–2) 58.3% Moderate (H&Y 2.5, 3) 30.1% Severe (H&Y 4, 5) 11.6%
H&Y OFF	Mild 43.4% Moderate 35.7% Severe 21%
S&E ON (median, range)	90%, 20–100
S&E OFF (median, range)	80% 10–100
MDS-UPDRS 1 (mean, SD)	14.62 ± 7.18
MDS-UPDRS 2 (mean, SD)	14.93 ± 8.20
MDS-UPDRS 3 (ON) (mean, SD)	27.39 ± 14.85
MDS-UPDRS 4 (mean, SD)	7.2 ± 3.45
NoMoFA total (mean, SD)	19.57 ± 11.97

Abbreviations: SD, standard deviation; LEDD, levodopa equivalent daily dose; MoCA, Montreal Cognitive Assessment battery; H&Y, Hoehn and Yahr; S&E, Schwab and England activities of daily living scale; MDS-UPDRS, Movement Disorder Society-Unified Parkinson’s Disease Rating Scale.

of Breath), although none of these fell below a discrimination value of 0.50.

Test-retest reliability for the total NoMoFA score was adequate (ICC = 0.73 [95% CI 0.64-0.79]). Temporal stability for the individual items as assessed using a weighted Kappa coefficient with linear weighting demonstrated stability with values >0.3 for all measures except three (Fear/Feeling Scared, K = 0.294; Impulsiveness, K = 0.297; Apathy/Loss of Interest, K = 0.298).

Concurrent validity. Spearman’s rho correlation coefficient compared the NoMoFA score to other measures of parkinsonian severity (Table 4). Correlation values were all in expected direction (negative for S&E, positive for all others) with higher correlation with other NMS surveys and measures of quality of life.

Final Questionnaire Designation

Following data analysis, investigators reconvened the Delphi panel to review results and discuss possible item reduction. Further analysis of four items that performed suboptimally demonstrated that the Spearman’s correlation between the “Disorientation” and “Confusion” items was 0.294 and the correlation between “Poor Decision Making” and “Impulsiveness” was 0.481. The relatively low correlation between “Disorientation” and “Confusion” was hypothesized to be due to the low frequency of endorsement of “Disorientation” as an item. The panel decided that “Disorientation” would be removed from the final questionnaire as the item related to “Confusion” sufficiently overlapped and would capture this construct. The Panel also decided to retain both “Poor Decision Making” and “Impulsiveness” in

TABLE 2. Frequency of non-motor symptoms (NMS) with proportion of static and fluctuating NMS (NMF)

NMS	Frequency of NMS (%)	Static NMS (%)	Fluctuating NMS (NMF) (%)
Low energy/fatigue	86.4	50.5	49.5
Word finding difficulty	76.0	64.8	35.2
Loss of train of thought	73.0	62.5	37.5
Pain	69.7	64.3	35.7
Restlessness	66.0	61.5	38.5
Poor short-term memory	65.5	84.9	15.1
Urinary symptoms	62.8	84.8	15.2
Distraction (difficulty completing task)	60.5	71.0	29.0
Excessive daytime sleepiness	59.6	65.3	34.7
Constipation	54.0	87.4	12.6
Difficulty handling stressful situations	51.8	77.0	23.0
Altered sensations	47.7	77.4	22.6
Changes in vision	46.7	85.5	14.5
Difficulty planning an activity	44.5	73.5	26.5
Excessive worry	41.5	82.9	17.1
Shortness of breath	38.5	87.0	13.0
Loneliness/isolation	34.5	86.4	13.6
Sadness/helplessness	31.2	84.9	15.1
Apathy/loss of interest	30.2	86.8	13.2
Excess sweating	29.0	88.0	12.0
Poor decision making	24.1	92.5	7.5
Palpitations	22.5	90.0	10.0
Impulsiveness	22.1	92.0	8.0
Hallucinations	21.0	88.0	12.0
Fear/feeling scared	19.5	88.5	11.5
Confusion	19.0	87.5	12.5
Compulsions/uncontrollable urges	18.5	94.5	5.5
Disorientation	17.5	88.5	11.5

the final questionnaire despite clinimetric testing inconsistencies, as these were clinically important and distinct constructs. The panel agreed that all other items were clinically relevant and that while some may have shown floor effects, this could have been due to the fact that the sample was not completely represented by those of greater disease severity.

The final validated NoMoFA questionnaire is provided in Appendix S1. The final scoring table is provided in Appendix S2. Based on removal of one item from NoMoFA there were a total of 27 items in the final questionnaire with a total maximum possible score for the NoMoFA of 81 (27×3).

Discussion

The NoMoFA, developed as a patient-derived and self-administered questionnaire using qualitative research methods that impart substantial face and construct validity, was found to be valid and reliable in capturing both static and fluctuating non-motor symptoms in PD. In creating the NoMoFA questionnaire there was deliberate intention to integrate patient input in the development of the scale from its inception, to include only items relevant to patients' own functional abilities and not easily rated by outside observers. The

self-rated framework was aimed at reducing the burden of administration as well providing a questionnaire that could be accessed remotely in an era where both clinical care and research are increasingly shifting to remote applications with the assistance of technology.

An additional priority when conceptualizing the NoMoFA was to be able to identify the proportion of NMS that fluctuate ("true" NMF) as well as the proportion of those that do not ("static" NMS). By providing a comprehensive survey of an individual's complete experience with NMS both static and in the ON and OFF states, NoMoFA has relevant application for measuring the effect of therapeutic interventions designed to reduce both NMS and NMF.

Other NMS instruments have been developed that include a subset of NMFs as part of the overall score. The recently validated Movement Disorder Society Non-Motor Rating Scale (MDS-NMS)³³ is a 52-item rater-administered survey of NMS with an eight-item NMF subscale. The items are not differentiated into their ON versus OFF presence, reflecting only the time spent in the OFF state. Item selection for the MDS-NMS was based on the non-motor symptoms scale (NMSS),³⁴ expert opinion, literature review, and included patient input through cognitive pretesting and administration of the preliminary version of the MDS-NMS. The NMSS³⁴ developed in 2007 and

TABLE 3. Clinimetric properties of Non-Motor Fluctuation Assessment (NoMoFA) items

Item	Missing	Skewness	80th %tile	Corrected item-total correlation	Cronbach's alpha if item deleted	Factor 1 loading	Factor 2 loading	Item response theory (IRT) discrimination	Test-retest weighted kappa
Loss of train of thought	0	0.266	2	0.418	0.893		0.685	1.094	0.488
Distraction (difficulty completing task)	0	0.651	2	0.539	0.891		0.639	1.063	0.446
Difficulty planning an activity	0	1.054	2	0.510	0.892		0.422	1.443	0.393
Disorientation	1	2.649	1	0.362	0.895		0.433	0.720	0.370
Confusion	0	2.399	1	0.495	0.892	0.387		1.508	0.415
Word finding difficulty	0	0.200	2	0.520	0.892		0.655	1.049	0.464
Excessive worry	0	0.927	2	0.670	0.888	0.637		2.314	0.559
Fear (feeling scared)	0	2.236	1	0.527	0.892	0.590		2.216	0.294
Restlessness	0	0.468	2	0.574	0.890	0.517		1.401	0.461
Sadness/hopelessness	1	1.474	2	0.574	0.890	0.732		2.738	0.395
Loneliness/isolation	0	1.459	1	0.586	0.890	0.691		2.592	0.439
Hallucinations	0	2.269	1	0.431	0.893		0.497	0.755	0.673
Poor decision making	1	2.274	1	0.504	0.892		0.430	2.234	0.440
Impulsiveness	1	2.268	1	0.411	0.894		0.375	1.342	0.297
Compulsions/uncontrollable urges	0	2.427	1	0.434	0.893		0.398	1.325	0.455
Poor short-term memory	0	0.345	2	0.496	0.892		0.644	1.072	0.532
Difficulty handling stressful situations	2	0.670	2	0.633	0.889	0.543		1.567	0.454
Apathy/loss of interest	2	1.605	2	0.535	0.891	0.582		1.915	0.298
Low energy/fatigue	1	-0.128	3	0.513	0.891	0.468		1.499	0.413
Excessive daytime sleepiness	2	0.549	2	0.453	0.893	0.362		1.112	0.480
Pain	2	0.212	3	0.334	0.896	0.362		0.528	0.431
Altered sensations	1	0.811	2	0.286	0.896			0.648	0.370
Shortness of breath	0	1.319	2	0.383	0.894	0.440		0.924	0.621
Changes in vision	1	0.710	2	0.364	0.895		0.314	1.917	0.524
Excess sweating	0	1.757	1	0.303	0.896			1.260	0.546
Palpitations	0	2.304	1	0.411	0.894	0.485		1.493	0.401
Urinary symptoms	1	0.331	2	0.356	0.895		0.315	1.447	0.545
Constipation	2	0.538	2	0.327	0.896	0.310		2.012	0.576

subsequently modified by Storch and colleagues³⁵ indirectly captures NMF. The NMF score is derived as the difference between NMSS scores in the ON versus the OFF state. While this instrument likely identifies NMF, the need to apply the 30-item clinician-administered scale in both the ON state and the OFF state is labor-intensive and may not lend itself readily to research and clinical practice. The self-administered Wearing-Off Questionnaire (WOQ)²⁴ has gone through several iterations initially starting as 33 items, then reduced to 19 items, and finally a nine-item questionnaire (WOQ-9),

which was found to be as valid and reliable as the longer versions. The WOQ-9 combines five motor and four non-motor items limited to items in the OFF state and was developed based on expert opinion with no patient input.

Two commonly used scales that include ratings of NMS do not include NMF. The validated non-motor experiences of daily living subscale of the Unified Parkinson Disease Rating Scale (MDS-UPDRS Ia and Ib)²⁰ includes items on cognitive, mood and behavioral symptoms, autonomic symptoms, sleep dysfunction,

TABLE 4. Concurrent validity: correlations with other instruments

	Correlation coefficient rho
MDS-UPDRS part 1 non-motor experiences of daily living	0.753
PDQ-8	0.686
MDS-UPDRS part 2 motor experiences of daily living	0.568
NMF severity (physician rating)	0.458
S&E OFF	-0.442
MF severity (physician rating)	0.399
H&Y OFF	0.376
MDS-UPDRS part 4 complications of therapy	0.359
S&E ON	-0.326
H&Y ON	0.302
WOQ-9	0.295
MDS-UPDRS part 3 motor examination	0.277

Abbreviations: MDS-UPDRS, Movement Disorder Society-Unified Parkinson's Disease Rating Scale; PDQ-8, Parkinson's Disease Questionnaire; NMF, non-motor fluctuation; S&E, Schwab and England; MF, motor fluctuation; H&Y, Hoehn and Yahr; WOQ-9, Wearing-Off Questionnaire-9.

and sensory symptoms, but it does not address fluctuations in these symptoms. The Scales for Outcomes in Parkinson Disease-Autonomic Dysfunction (SCOPA-AUT)²² is a reliable and valid measure of autonomic NMS in PD, but it too does not address fluctuations in symptoms.

We intended to recruit a representative sample to maximize generalizability of the application of NoMoFA in diverse patient populations. We elected not to exclude those with moderate or severe cognitive impairment from the sample (MoCA score was not exclusionary), despite the concern that reduced cognition would jeopardize the quality of the data by not providing reliable responses. We intentionally included these subjects in recognition that those with more advanced illness were also more likely to experience greater fluctuations in their symptoms.³⁶ We accommodated this limitation by including patient/care-partner dyads where the care-partner was able to reliably report the symptoms as a surrogate if the patient was struggling to answer independently. We did not examine differences in clinimetric properties between scales completed by patients alone versus scales completed by patients and their care-partner due to the limited sample size of the latter group and the potential for truncated variance. This examination would be interesting to conduct in a planned analysis in future studies.

Our examination of the clinimetric properties of the NoMoFA demonstrated adequate results. Most items met our criteria for acceptance. Those few that did not were either removed from the scale or were determined by the Delphi panel to be of sufficient clinical importance to be included even with suboptimal clinimetric properties. This is an important consideration in developing clinical rating scales; the inclusion of items

should not be driven solely by consideration of the statistics.³⁷ In addition, all items demonstrated adequate test-retest stability, an important marker of reliability. This is significant in that reliability is the rate-limiting factor in a scale's validity.³⁸

In assessing concurrent validity, we sought to include as many questionnaires as possible to reflect the wide spectrum of non-motor symptoms while balancing this with the concern for respondent fatigue. As such, we limited the number of questionnaires at the expense of potentially excluding relevant additional questionnaires addressing constructs such as pain or fatigue. Nonetheless, the direction of all correlations with respect to the questionnaires we included was appropriate, with strongest correlations between MDS-UPDRS I (non-motor symptoms), PDQ-8, MDS-UPDRS II, NMF-severity (physician-rated), as well as negative correlations with the ON and OFF S&E, as predicted. The correlation between the NoMoFA and the PDQ-8 suggests that greater NMF burden correlates with worse quality of life. The concurrent validity results suggest that the NoMoFA correlates to the largest extent with other questionnaires capturing NMS more than with those measuring motor symptoms, and with those questionnaires that capture fluctuating symptoms more than static symptoms. Interestingly, while MF-severity (physician-rated), WOQ-9, and UPDRS IV were positively correlated with NMF, the strength of correlation was less robust, further suggesting that MF and NMF may not be of equal magnitude and may have a different impact in an individual, with some studies suggesting NMF drives a greater reduction in quality of life than MF.¹

The NoMoFA uniquely captured static and fluctuating symptoms within individuals. Interestingly, fatigue was endorsed as both the most frequent overall NMS in a vast majority of subjects, and the NMS that most frequently fluctuates among all subjects. Likewise, word-finding difficulty, loss of train of thought, excessive daytime sleepiness, pain, and restlessness frequently fluctuated. This has direct treatment implications for people with PD, as these NMS are often managed as non-dopamine-sensitive symptoms, with significant treatment failures and negative impact on quality of life. The recognition within an individual patient of which symptom is static and which fluctuates would allow a more precise and likely effective approach to symptom management. Indeed, cognitive symptoms like *word-finding difficulties* and *loss of train of thought* frequently herald a diagnosis of dementia with limited treatment options; recognition that these symptoms could be improved with modification of dopamine-mediated strategies has significant implications on function and independence.

Our study has several limitations. Interim review of enrollment approximately halfway into the study

revealed that a disproportionate number of subjects were either mild or moderate in disease severity. As such, efforts increased to preferentially enroll subjects classified as severe, however, despite this, only 21% of the subjects enrolled were ultimately classified as severe. Our challenge with recruiting advanced patients was similar to other studies, notably the recent MDS-NMS validation study³³ that also had a predominant mild-moderate cohort. This is not surprising given the increased challenge of more severely affected people with PD to attend outpatient subspecialty clinics. Patients in later stage of illness may be placed in long-term care given their high care-needs³⁹ and therefore excluded from the pool of willing and capable participants involved in research.

Conclusions

The NoMoFA is the first valid and reliable comprehensive patient-derived and patient-administered questionnaire that captures and quantifies NMF. It has been created through a methodologically rigorous process with focus group and cognitive interviewing input, Delphi panel deliberations, and two large-scale validations, to produce a final survey of static and fluctuating NMS. Integration of NoMoFA into clinical practice and research protocols will be expected to facilitate efficient and effective customization of treatment strategies and augment the value of future research endeavors particularly for therapies that may improve NMS/NMF. Future work will include determination of the minimal clinically important difference in NoMoFA scores and its sensitivity to change (responsivity) with interventions. This will inform our understanding of the natural history of NMF, its relationship to other symptoms, and its impact on quality of life. ■

References

1. Witjas T, Kaphan E, et al. Nonmotor fluctuations in Parkinson's disease: frequent and disabling. *Neurology* 2002;59(3):408-413.
2. Riley DE, Lang AE. The spectrum of levodopa-related fluctuations in Parkinson's disease. *Neurology* 1993;43:1459-1464.
3. Quinn N. Classification of fluctuations in patients with Parkinson's disease. *Neurology* 1998;51:S25-S29.
4. Martínez-Fernández R, Schmitt E, Martínez-Martin P, Krack P. The hidden sister of motor fluctuations in Parkinson's disease: a review on nonmotor fluctuations. *Mov Disord* 2016;31(8):1080-1094.
5. Chaudhuri KR, Martínez-Martin P, Schapira AHV, et al. International multicenter pilot study of the first comprehensive self-completed nonmotor symptoms questionnaire for Parkinson's disease: the NMSQuest study. *Mov Disord* 2006;21(7):916-923.
6. Goetz CC. The Unified Parkinson's Disease Rating Scale (UPDRS): status and recommendations. *Mov Disord* 2003;18(7):738-750.
7. Martínez-Martin P, Schrag A, Weintraub D, Rizos A, Rodríguez-Blázquez C, Chaudhuri KR. Pilot study of the International Parkinson and Movement Disorder Society-sponsored Non-motor Rating Scale (MDS-NMS). *Mov Disord Clin Pract* 2019;6:227-34.6 (3):227-234.
8. Kleiner-Fisman G, Martine R, Lang AE, Stern MB. Development of a non-motor fluctuation assessment instrument for Parkinson disease. *Parkinsons Dis* 2011;292719.
9. U.S. Department of Health and Human Services: Food and Drug Administration. Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. *Clin Fed Regist* 2009.
10. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multidisciplinary health research. *BMC Med Res Methodol* 2013.
11. Taylor WJ, Schumacher HR, Baraf HSB, et al. A modified Delphi exercise to determine the extent of consensus with OMERACT outcome domains for studies of acute and chronic gout. *Ann Rheum Dis* 2008;67(6):888-891.
12. Miyasaki JM, Shannon K, Voon V, et al. Practice parameter: evaluation and treatment of depression, psychosis, and dementia in Parkinson disease (an evidence-based review): report of the Quality Standards Subcommittee of the American Academy of Neurology. *Neurology* 2006;66:996-1002.
13. Tomlinson CL, Stowe R, Patel S, Rick C, Gray R, Clarke CE. Systematic review of levodopa dose equivalency reporting in Parkinson's disease. *Mov Disord* 2010;25(15):2649-2653.
14. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297-334.
15. Loundsbury JW, Gibson LW, Saudargas RA. Scale development. In: Leong FTL, Austin JT, eds. *The Psychology Research Handbook: A Guide for Graduate Students and Research Assistants*. Thousand Oaks, CA: Sage Publications; 2006.
16. Stevens JP. *Applied Multivariate Statistics for the Social Sciences*, 5th Edition. New York: Taylor & Francis Group; 2009.
17. Maxwell SE, Kelley K, Rausch JR. Sample size planning for statistical power and accuracy in parameter estimation. *Annu Rev Psychol* 2008;59 537-563.
18. Postuma RB, Berg D, Stern M, et al. MDS clinical diagnostic criteria for Parkinson's disease. *Mov Disord* 2015;30(12):1591-601.
19. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* 2005;53(4):695-699.
20. Goetz CG, Tilley BC, Shaftman SR, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov Disord* 2008;23(15):2129-2170.
21. Hoehn MMYM. Parkinsonism: onset, progression and mortality. *Neurology* 1967;17:427-442.
22. Visser M, Marinus J, Stiggelbout AM, van Hilten JJ. Assessment of autonomic dysfunction in Parkinson's disease: the SCOPA-AUT. *Mov Disord* 2004;19(11):1306-1312.
23. 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 Heal* 1997;12(6):805-814.
24. Stacy MA, Murphy JM, Greeley DR, Stewart RM, Murck H, Meng X. The sensitivity and specificity of the 9-item Wearing-off Questionnaire. *Park Relat Disord* 2008;14(3):205-212.
25. Guy W. *ECDEU Assessment Manual for Psychopharmacology: Revised* (DHEW Publication No. ADM 76-338. Rockville, MD: US Department of Health, Education and Welfare; 1976:383-406.
26. DeVellis RF. *Classical test theory*. *Med Care* 2006;44(11 Suppl 3):S50-S59.
27. Lord FM. *Applications of item response theory to practical testing problems*. Publisher: Lawrence Erlbaum Associates, Hillsdale NJ, 1980.
28. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol Assess* 1994;6:284-290.
29. Curran PJ, West SG, Finch JF. The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychol Methods* 1996;1:16-29.
30. Forero CG, Maydeu-Olivares A. Estimation of IRT graded response models: limited versus full information methods. *Psychol Methods* 2009;14:275-299.

31. Hays RD, Morales LS, Reise SP. Item Response Theory and health outcomes measurement in the 21st century. *Med Care* 2000;38:1128–1142.
32. Kang T, Chen TT. Performance of the generalized S-X2 item fit index for the graded response model. *Asia Pacific Educ Rev* 2011;45:391–406.
33. Chaudhuri KR, Schrag A, Weintraub D, et al. The Movement Disorder Society Nonmotor Rating Scale: initial validation study. *Mov Disord* 2020;35(1):116–133.
34. Chaudhuri KR, Martinez-Martin P, Brown RG, et al. The metric properties of a novel non-motor symptoms scale for Parkinson's disease: results from an international pilot study. *Mov Disord* 2007;22(13):1901–11.
35. Storch A, Schneider CB, Klingelhöfer L, et al. Quantitative assessment of non-motor fluctuations in Parkinson's disease using the Non-Motor Symptoms Scale (NMSS). *J Neural Transm* 2015;122(12):1673–84.
36. Freitas ME, Hess CW, Fox SH. Motor complications of dopaminergic medications in Parkinson's disease. *Semin Neurol* 2017.
37. Stebbins GT. Clinical rating scale development. In: C. Sampaio CGG& AS, ed. *Park Dis Rat Scale*. New York, NY: Oxford University Press, Inc.; 2012.
38. Nunnally JC, Bernstein IH. *Psychometric Theory*. 3rd ed. New York, NY: McGraw Hill; 2010.
39. Safarpour D, Thibault DP, Desanto CL, et al. Nursing home and end-of-life care in Parkinson disease. *Neurology* 2015;85(5):413–19.

Supporting Data

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.