Using Movement Disorder Society Unified Parkinson's Disease Rating Scale Parts 2 and 3 Simultaneously: Combining the Patient Voice with Clinician Ratings

Yuanyuan Guo, PhD[1], **Christopher G. Goetz, MD**[2], **Glenn T. Stebbins, PhD**[2], **Tiago A. Mestre, MD, PhD**[3], **Sheng Luo, PhD**[1,*]

[1]Department of Biostatistics & Bioinformatics, Duke University, Durham, North Carolina, USA

[2]Department of Neurological Sciences, Rush University Medical Center, Chicago, Illinois, USA

[3]Division of Neurology, Department of Medicine, Ottawa Hospital Research Institute, University of Ottawa Brain and Mind Research Institute, University of Ottawa, Ottawa, Ontario, Canada

Abstract

Background: Regulatory recommendations favor outcomes combining objective and patient input. The Movement Disorder Society Unified Parkinson's Disease Rating Scale (MDS-UPDRS), the most commonly used scale in Parkinson's disease (PD), includes patient and investigator ratings in distinct parts, but original clinimetric analyses failed to confirm the validity of combining parts by simple summing.

Objectives: The aim was to develop clinimetrically valid constructs for combining patient-reported Part 2 and investigator-rated Part 3 MDS-UPDRS scores.

Methods: Using 7888 MDS-UPDRS scores, we assessed construct validity of combined Part 2 and Part 3 items using exploratory factor analysis (EFA) and graded item response theory (IRT) with threshold criteria: comparative fit index ≥ 0.9 (EFA) and discrimination parameters ≥ 0.65 (IRT).

Results: The direct sum of Parts 2 + 3 failed to meet the threshold for a valid outcome of PD severity (comparative fit index, CFI = 0.855). However, a two-domain construct combining item scores for tremor and non-tremor domains from Parts 2 and 3 confirmed validity, meeting both EFA and IRT criteria as distinct but correlated indices of disease severity (CFI = 0.923; discrimination mean 2.197 ± 0.480 [tremor] and 1.737 ± 0.344 [non-tremor] domains).

Conclusions: The sum of Parts 2 + 3 is not clinimetrically sound. However, considering tremor and non-tremor items of both Parts 2 and 3 as two outcomes results in a valid summary of PD motor severity that leverages simultaneous patient- and investigator-derived measures.

*Correspondence to: Dr. Sheng Luo, Department of Biostatistics & Bioinformatics, Duke University, Durham, NC 27708, USA; sheng.luo@duke.edu.

This analytic application addresses regulatory prioritizations and retains the well-validated MDS-UPDRS items. In future interventional trials, we suggest that tremor and non-tremor components of PD motor severity from Parts 2 + 3 be monitored and analyzed to accurately detect objective changes that integrate the patient's voice.

Keywords

total score; item response theory; factor analysis; Parkinson's disease

Introduction

Parkinson's disease (PD) is one of the most common neurodegenerative brain disorders characterized by distinctive motor and nonmotor manifestations.[1] It is a complex and heterogeneous disease, and an important research question is how to describe and measure its severity and progression. Among rating scales, the Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS) is the most widely used one in PD.[2]

Since its development, MDS-UPDRS has been utilized as either the primary or secondary outcome in many clinical trials.[3] Based on its clinimetric development, the developers recommended that each part of the MDS-UPDRS be considered separately.[4] To this end, most prior studies analyzed each part of the scale individually.[5,6] Nevertheless, some studies have used the combined score of all or selected parts as an outcome.[7] In particular, a sum score of the two parts focusing on motor functional impact and severity, the patient-based Part 2 "Motor Experiences of Daily Living" and the rater-based Part 3 "Motor Examination," respectively, has been used as an efficacy outcome in clinical trials, both for primary[8,9] and secondary outcomes.[2,10] Furthermore, recent FDA (U.S. Food and Drug Administration) recommendations highlight the importance of including the voice of the patient in clinical outcome assessments, in addition to objective ratings.[11] The challenge is to find an outcome that respects the face validity and construct validity of item combination to achieve the twofold goal of utilizing the patient's voice and the rater examination.

The objectives of this research were first to assess the validity of combining MDS-UPDRS patient-reported motor disability (Part 2) with examiner-rated motor severity (Part 3), and model new methods to address regulatory recommendations to incorporate the patient's voice in clinical outcomes. We evaluated if the overall motor sections of MDS-UPDRS Parts 2 + 3 scores would be better captured by score combinations than the single summed score. We used a large sample size of MDS-UPDRS scores covering all Hoehn and Yahr (HY) stages to re-examine the validity of a combined Parts 2 and 3 score to assess the impact of motor severity in PD. Further, because we have previously demonstrated that Part 3 has a two-domain construct with distinct tremor and non-tremor domains,[5] we investigated the structural characteristics of a combined Parts 2 and 3, hypothesizing that the combined outcome would also reflect the same dichotomy of distinct but statistically correlated domains.[6,12]

Patients and Methods

Study Population

The MDS-UPDRS translation program sponsored by the International Parkinson and Movement Disorder Society (MDS) is an ongoing cross-sectional, multinational, multicenter study designed to develop and validate the translation of the MDS-UPDRS.[5] The original data set included 8931 complete MDS-UPDRS ratings from PD patients (UK Brain Bank Criteria[13]) representing all Hoehn and Yahr stages with assessments performed in the patient's native language (24 international languages, not including English). We excluded 1043 patients with missing scores in MDS-UPDRS Part 2 or 3 items resulting in a total of 7888 patients included for analysis. All PD patients participating in the MDS-UPDRS translation program provided informed consent, and this research was approved by institutional review board.

Statistical Analyses

We conducted a two-step analysis. In step 1, we evaluated the validity of a unidimensional construct of the MDS-UPDRS Part 2 (13 items) and Part 3 (33 items) scores. In particular, we used exploratory factor analysis (EFA) and item response theory model (IRT) to determine the factor loading and discrimination profile. In step 2, we evaluated the validity of a multidimensional construct (a two-factor structure, tremor vs. non-tremor) found in prior studies. In particular, we used EFA to determine best factor structure and employed IRT to validate this structure by fitting 35 non-tremor items (Part 2 items 2.1–2.9, 2.11–2.13, and Part 3 items 3.1–3.14) and 11 tremor items (Part 2 item 2.10 and Part 3 items 3.15a–3.18) separately in two independent IRT models. We compared the discrimination differences and the model-fitting statistics from these two steps to check the model performance. The EFA and IRT analyses were conducted using the R package *mirt*.[14] We computed the Pearson's correlations to assess the relationship between tremor and non-tremor domains.

Construct Validity Criteria

Categorical EFA was performed on MDS-UPDRS Parts 2 + 3 to evaluate factor structures. To assess the model fit, the goodness-of-fit indices were used, including the comparative fit index (CFI), the Tucker-Lewis Index (TLI), and the standardized root mean square residual (SRMSR). Based on prior statistical literature, we prespecified the thresholds for validity with a CFI and TLI equal to or above 0.90,[15] and an SRMSR lower than 0.08[16] as reflective of a good fit. A factor loading cutoff of larger than 0.40 was used to determine whether items were loaded in a factor.[17] We also compared different models using the likelihood ratio test (χ^2). Moreover, Akaike information criterion (AIC) was used as estimators of the relative quality of statistical models for a given set of data, where lower values indicate better fit to the model.[18]

The graded-response IRT model was used to analyze data from MDS-UPDRS Parts 2 + 3.[12] In IRT, we considered the latent variable, termed *theta*, to capture the underlying latent trait of PD severity and to establish the relationship between the hidden latent variable and the measure response.[19] The model generates five parameters: one discrimination

parameter and four location parameters. The location parameter, also referred to as the "difficulty" parameter, describes where the item functions best along the trait scale.[20] The discrimination parameter describes how well the corresponding item can differentiate individuals at different trait levels and provides information about the differential capability of the item. A higher value of discrimination parameter of the item suggests that the item contributes more in precision measurement of the disease, compared to those with lower values.[21] The magnitude of the discrimination parameter can be determined with the following thresholds: none = 0; very low = 0.01 to 0.34; low = 0.35 to 0.64; moderate = 0.65 to 1.34; high = 1.35 to 1.69; very high 1.70.[22] We consider the score would be valid if the discrimination parameters are in or above the moderate range. IRT approaches identify subdomains by looking at the scales of its discrimination parameters.[20] In addition, the correlation between two latent variables assesses the strength of the association between tremor and non-tremor impairments.

Overall, the construct validity was defined by acceptable CFI 0.9 from EFA[15] and all discrimination parameters 0.65 from IRT analysis.[22]

Results

Study Sample

Of the total 7888 subjects, 6161 (78.1%) were on medication "ON" state, and 3439 (43.8%) were female. The mean time since PD diagnosis year was 7.64 years and 28.8% of subjects had dyskinesia. Patients were diverse of Hoehn and Yahr stages, ranged from stage 0 to stage 5 (with the proportion of 0.6%, 13.5%, 49.0%, 26.4%, 8.3%, and 2.2%, respectively). The mean scores for MDS-UPDRS Parts 1, 2, and 3 were 11.95, 14.77, and 33.52, respectively. The mean education level was 11.27 years (Table 1).

Construct Validity of a Unidimensional Structure of MDS-UPDRS Parts 2 + 3

We conducted the EFA and IRT analyses considering all 46 Parts 2 + 3 items in a unidimensional manner. The one-factor EFA results (Table 2) suggest that all 35 non-tremor items (Part 2 items 2.1–2.9, 2.11–2.13, and Part 3 items 3.1–3.14) had factor loading larger than 0.4, whereas all 11 tremor items (Part 2 item 2.10 and Part 3 items 3.15a–3.18) had factor loading smaller than 0.4. Moreover, the one-factor EFA model had CFI = 0.855 and TLI = 0.847, which were lower than the prespecified validity threshold of 0.9. Also, the SRMSR = 0.116, which was higher than the threshold of 0.08.

We then analyzed all 46 Parts 2 + 3 items in an IRT model. The results in Table 3 suggest that the discrimination parameters for 35 non-tremor items were from moderate to very high, ranging from 0.870 to 2.243 with a mean of 1.731 ± 0.342. In contrast, the discrimination parameters for the 11 tremor items were low to moderate (range: 0.365–0.674, mean: 0.529 ± 0.081). Given the fact that one-factor EFA rendered CFI < 0.9 and the discrimination parameters of most tremor items from IRT analysis were smaller than 0.65, the direct sum of Parts 2 + 3 failed to meet the validity threshold for a valid outcome of PD severity, corroborating the conclusion in the original article that the single factor structure for the combination of Parts 2 + 3 could not be confirmed.

Construct Validity of a Tremor and Non-Tremor Structure of MDS-UPDRS Parts 2 + 3

We conducted the EFA and IRT analyses considering all 46 Parts 2 + 3 items in a multidimensional manner. We first fitted a two-factor EFA model to all 46 Parts 2 + 3 items. EFA results in Table 4 suggest two explicit factors, with the 35 non-tremor items loading on one factor (factor loading mean of 0.691 ± 0.083) and the 11 tremor items loading on another factor (factor loading mean of 0.751 ± 0.098). No items had cross-loading (item simultaneously loaded on two factors). Moreover, the two-factor EFA model had CFI = 0.923 and TLI = 0.914, which meets the threshold of 0.9. In addition, the SRMSR = 0.075, satisfying the threshold of 0.08. The multidimensional model with two factors (AIC = 745,981.1) had a superior fit than the unidimensional model with one factor (AIC = 775,081.6), in addition to a significant improvement ($\chi^2 = 29190.55$; df = 45; $P < 0.005$).

We then fitted two IRT models to 35 non-tremor items and 11 tremor items separately. The results displayed in Table 5 suggest that the low to moderate discrimination parameters associated with 11 tremor items shown in Table 3 increased remarkably, with the mean of 2.197 ± 0.480 and a range of 1.479 to 3.076, being "high" to "very high." The Pearson's correlation coefficient between non-tremor and tremor latent constructs estimated from two separate IRT models was 0.240 (95% CI: 0.219, 0.260), suggesting that two latent constructs describe unique but not completely independent aspects of PD and further proved the superiority of the multidimensional structure. The results from EFA and IRT modeling were statistically consistent, and they confirmed the clinically meaningful two-domain construct with distinct tremor and non-tremor domains for MDS-UPDRS Parts 2 + 3.

Investigation of Three-Factor Structure of MDS-UPDRS Parts 2 + 3 and All Combinations

We also examined a three-factor structure for Parts 2 + 3, but several items (Part 3 items 3,1, 3.2, 3.8a, 3.9, 3.10, 3.12–3.14) had salient loadings on more than one factor and did not add substantial clinical clarity compared to the two-factor solution (Table S1). Moreover, because nonmotor (Part 1) and motor complications of fluctuations and dyskinesia (Part 4) may be of clinical interest, we examined all combinations (Parts 1 + 2 + 3, Parts 2 + 3 + 4, and Parts 1 + 2 + 3 + 4), and they failed to meet validity threshold of CFI 0.9 (0.870 in Parts 1 + 2 + 3, 0.842 in Parts 2 + 3 + 4, and 0.864 in Parts 1 + 2 + 3 + 4).

Discussion

The MDS-UPDRS is used widely in cross-sectional and longitudinal studies of PD. The original clinimetric analyses justified that each part score was best analyzed independently because a total MDS-UPDRS or any combination of individual parts (Parts 2 and 3 included) did not meet validation criteria.[4] However, the sample size of the original analysis was only a few hundred subjects, and the distinct tremor and non-tremor structure was not investigated. Moreover, given new FDA recommendations of adopting ratings that not only matter to investigators but also reflect the patient's voice and recent clinical studies using the Parts 2 + 3 combined scores as primary outcomes, a reasonable research question is to comprehensively re-examine the validity of a combined Parts 2 + 3 score that has both the patient- and rater-based information. Part 2 was originally designed with fewer items than Part 3 to alleviate patient burden. Part 2 has been established to be a useful

PD patient-reported instrument for assessment of disability, and it performs according to hypotheses enunciated into the theoretical framework in which the scale was designed.[23]

Our current study enjoys the advantages of a larger sample size and more advanced statistical applications provided by IRT modeling. It allows us to verify our own original recommendations and to corroborate that the combined Parts 2 + 3 represent a strong clinimetric measure of motor severity combining the patient's voice with objective ratings with distinct domains. Our analysis included EFA and IRT, which are both strong statistical tools for scale evaluation.[5] The factor loadings are coefficients that express the relationship between each item and the underlying factor. The discrimination parameters generated by IRT measure the differential capability of an item where a higher value suggests the item has a high ability to differentiate subjects. By examining the scale of those parameters, we can identify whether items function together or separately and allow clinicians to determine if these function clusters fall into components that represent clinically relevant domains.

Regarding the cutoff point of CFI, some reports[24–26] accept 0.80 for an adequate fit, whereas some recommend the use of CFI 0.95 as the threshold.[27] On the contrary, multiple studies,[28–30] especially those considering a variety of simulation settings with different sample sizes and item numbers, recommend CFI 0.90 as the threshold. The point is particularly relevant to analyses where there are more than 30 items and the sample size is larger than 250. Following this recommendation and accepting that there is no absolute rule, we set CFI 0.90 as the threshold of pass versus no pass. These were the same criteria used in the original validation of the MDS-UPDRS.[4]

We applied the factor analysis and the graded response model to all 46 Parts 2 + 3 items. The factor loadings of the 11 tremor items (Part 2 item 2.10 and Part 3 items 3.15a–3.18) were below 0.4 with the unsatisfactory CFI. Their discrimination parameters were below the construct validity threshold. Those results suggest that the tremor items contribute minimally when modeling the overall disease severity compared to those non-tremor items assessing bradykinesia, rigidity, gait, and posture. In contrast, when we modeled by considering tremor and non-tremor items separately, the CFI for the factor analysis rose to meet the acceptable validity threshold and the discrimination parameters improved to high and very high values. A similar two-domain factor structure can be found in our prior work of only Part 3,[5] suggesting the robustness of the structure in both Part 3 and Parts 2 + 3. In addition, the multidimensional constructs of the motor section of MDS-UPDRS had a superior fit over the unidimensional construct, supported by the contrasting CFI comparisons and discrimination scales. These results suggest that it is feasible to combine the objective ratings of PD motor severity based on the clinician's examination along with the patient-reported functional impact of motor severity. However, tremor and non-tremor domains must be treated as distinct outcomes, each with the potential to respond differently in terms of natural history, progression over time, and responses to interventions. Combining data from Parts 2 + 3 into two discrete outcomes properly addresses regulatory concerns about incorporating patients' voice in efficacy outcomes of clinical trials of people with PD. Given that tremor and non-tremor signs of PD may respond differently to medication (on vs. off states), our prior work confirmed the two-domain construct in Part 3 is retained in both conditions.[31] This finding needs to be confirmed in Parts 2 + 3.

In our view, it is also important to emphasize strategies and formulas that fail validity testing, so that the MDS-UPDRS does not have inappropriate or indefensible applications. Our two-domain solution was a clear best fit for handing the full item bank of Parts 2 + 3, and this finding was not only confirmed by our analysis but also amplified by the failure of both the unidimensional and the three-dimensional solutions. Moreover, even with the very large data set of this study, the failure of the basic summing strategy for other combinations of MDS-UPDRS Parts (eg, Parts 1 + 2 + 3, Parts 2 + 3 + 4, and Parts 1 + 2 + 3 + 4) should convince the field that such outcomes cannot be utilized in a valid manner in research or clinical practice.[32]

This study builds on our previous work that identified the two-domain factor structure, tremor and non-tremor, of Part 3 items of MDS-UPDRS.[5] The current study confirmed the stable structure clinimetrically sound regardless of the data collection method. Moreover, it would also be reasonable to assume that tremor and non-tremor function areas would not respond equally to treatment. For this reason, we have previously re-evaluated both the SURE-PD3[33] and STEADY-PD III[34] studies with the Part 3 divided into tremor and non-tremor domains, specifically documenting that tremor and non-tremor elements of the MDS-UPDRS did not behave in the same way.[6,12] We plan to test these hypotheses with the combined Parts 2 + 3 tremor and non-tremor scores in future research and clinical trial analyses. To investigate different progression patterns and treatment responses in the related but distinct tremor and non-tremor domains of diseases, we recommend the use of multidimensional, longitudinal IRT models with multiple latent variables as detailed in our prior work.[6,12,35]

MDS-UPDRS Part 2 is a patient-reported outcome (PRO) reporting motor disability in experiences of daily living. The selection of an appropriate PRO to reflect patients' voice should align with the research objectives. The goal of this work is to demonstrate that the two-domain construct of distinct but correlated tremor and non-tremor domains identified in Part 3[5,6,12] also exists in a combined Parts 2 and 3. We do not intend to suggest to replace other PROs (eg, PDQ-39) or outcomes from wearable devices (eg, Parkinson's KinetiGraph) by MDS-UPDRS Part 2.

Our analysis had limitations. We did not have an extensive clinical sample size representing advanced disease (Stage 5 Hoehn and Yahr), so we are not able to model the advanced stage of the disease in full confidence. Only 2.2% were Stage 5, that very advanced patients rarely participate in clinical studies. We also did not have information on patients' status on palliative care and PD dementia. The inclusion of these patients and other subset in future analyses may increase the heterogeneity of the study population and the generalizability of our findings across the full clinical spectrum of PD. Further, our analysis is limited by its cross-section design, and further research could be planned to validate those findings in a longitudinal setting. Nevertheless, our findings break new ground for further exploration of the combined use of MDS-UPDRS Parts 2 + 3.

In future interventional trials, we suggest that tremor and non-tremor components of PD motor severity from Parts 2 + 3 be analyzed to accurately detect objective changes that are relevant to patients. The items of the current MDS-UPDRS remain unchanged but are

simply rearranged categorically to address domains of impairment and disability combining both the patient's feedback and the investigator's observations. This novel but historically anchored view of tremor and non-tremor components of PD can serve as a model for future studies but also the basis for a re-evaluation of prior studies where salient observations may emerge once data are re-organized and presented in clinically relevant divisions that converge patient- and investigator-based data.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

1. Lima MM, Martins EF, Delattre AM, et al. Motor and non-motor features of Parkinson's disease—a review of clinical and experimental studies. CNS Neurol Disord Drug Targets 2012;11:439–449. [PubMed: 22483309]

2. Regnault A, Boroojerdi B, Meunier J, Bani M, Morel T, Cano S. Does the MDS-UPDRS provide the precision to assess progression in early Parkinson's disease? Learnings from the Parkinson's progression marker initiative cohort. J Neurol 2019;266:1927–1936. [PubMed: 31073716]

3. Holden SK, Finseth T, Sillau SH, Berman BD. Progression of MDS-UPDRS scores over five years in De novo Parkinson disease from the Parkinson's progression markers initiative cohort. Mov Disord Clin Pract 2018;5:47–53. [PubMed: 29662921]

4. 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;23:2129–2170. [PubMed: 19025984]

5. de Siqueira Tosin MH, Goetz CG, Luo S, Choi D, Stebbins GT. Item response theory analysis of the MDS-UPDRS motor examination: tremor vs nontremor items. Mov Disord 2020;35:1587–1595. [PubMed: 32469456]

6. Luo S, Zou H, Goetz C, et al. Novel approach to MDS-UPDRS monitoring in clinical trials: longitudinal item response theory models. Mov Disord Clin Pract 2021;8:1083–1091. [PubMed: 34631944]

7. Makkos A, Kovacs M, Aschermann Z, et al. Are the MDS-UPDRS-based composite scores clinically applicable? Mov Disord 2018;33: 835–839. [PubMed: 29488318]

8. Hattori N, Takeda A, Takeda S, et al. Rasagiline monotherapy in early Parkinson's disease: a phase 3, randomized study in Japan. Parkinsonism Relat Disord 2019;60:146–152. [PubMed: 30205936]

9. Hattori N, Takeda A, Takeda S, et al. Long-term, open-label, phase 3 study of rasagiline in Japanese patients with early Parkinson's disease. J Neural Transm (Vienna) 2019;126:299–308. [PubMed: 30689042]

10. Arrington L, Ueckert S, Ahamadi M, Macha S, Karlsson MO. Performance of longitudinal item response theory models in shortened or partial assessments. J Pharmacokinet Pharmacodyn 2020;47: 461–471. [PubMed: 32617833]

11. McLeod LD, Coon CD, Martin SA, Fehnel SE, Hays RD. Interpreting patient-reported outcome results: US FDA guidance and emerging methods. Expert Rev Pharmacoecon Outcomes Res 2011; 11:163–169. [PubMed: 21476818]

12. Luo S, Zou H, Stebbins GT, et al. Dissecting the domains of Parkinson's disease: insights from longitudinal item response theory modeling. Mov Disord 2022;37:1904–1914. [PubMed: 35841312]

13. Gibb WR, Lees AJ. The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. J Neurol Neurosurg Psychiatry 1988;51:745–752. [PubMed: 2841426]

14. Chalmers RP. Mirt: a multidimensional item response theory package for the R environment. J Stat Softw 2012;48:1–29.

15. Kline RB. Principles and Practice of Structural Equation Modeling. 4th ed. New York: The Guilford Press; 2016.

16. Hair JF, Hult GTM, Ringle CM, Sarstedt M. A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM). Newbury Park, California: SAGE; 2021.

17. Guadagnoli E, Velicer WF. Relation of sample size to the stability of component patterns. Psychol Bull 1988;103:265–275. [PubMed: 3363047]

18. Wagenmakers EJ, Farrell S. AIC model selection using Akaike weights. Psychon Bull Rev 2004;11:192–196. [PubMed: 15117008]

19. Baker FB, Hambleton RK, Swaminathan H. Item response theory—principles and applications. Appl Psychol Meas 1985;9:337–339.

20. Nguyen TH, Han HR, Kim MT, Chan KS. An introduction to item response theory for patient-reported outcome measurement. Patient 2014;7:23–35. [PubMed: 24403095]

21. Ul Hassan M, Miller F. Discrimination with unidimensional and multidimensional item response theory models for educational data. Commun Stat Simul Comput 2022;51:2992–3012.

22. The Basics of Financial Econometrics: Tools, Concepts, and Asset Management Applications [Computer Program]. 1st ed. Hoboken, New Jersey: John Wiley & Sons; 2014.

23. Rodriguez-Blazquez C, Rojo-Abuin JM, Alvarez-Sanchez M, et al. The MDS-UPDRS part II (motor experiences of daily living) resulted useful for assessment of disability in Parkinson's disease. Parkinsonism Relat Disord 2013;19:889–893. [PubMed: 23791519]

24. Planing P Innovation Acceptance: The Case of Advanced Driver-Assistance Systems. Wiesbaden: Springer Fachmedien Wiesbaden; 2014.

25. Akku A Developing a scale to measure students' attitudes toward science. Int J Assess Tool Educ 2019;6:706–720.

26. Meyers LS, Gamst G, Guarino AJ. Applied Multivariate Research: Design and Interpretation. Newbury Park, California: Sage Publications; 2016.

27. Lt H, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Model Multidiscip J 1999;6:1–55.

28. Nunnally JC, Nunnaly JC. Psychometric Theory. McGraw-Hill,New York City, NY, USA; 1978.

29. Horn JL, McArdle JJ. A practical and theoretical guide to measurement invariance in aging research. Exp Aging Res 1992;18: 117–144. [PubMed: 1459160]

30. Hair J, Anderson R, Black B, Babin B. Multivariate Data Analysis. Pearson Education,London, England; 2016.

31. Guo Y, Stebbins GT, Mestre TA, Goetz CG, Luo S. MDS-UPDRS motor examination retains its two-domain profile in both ON and OFF. Mov Disord Clin Pract 2022;9:1149–1151. [PubMed: 36339308]

32. Goetz CG, Choi D, Guo Y, Stebbins GT, Mestre TA, Luo S. It is as it was: MDS-UPDRS part 3 scores cannot be combined with other parts to give a valid sum. Mov Disord 2022; In press.

33. Schwarzschild MA, Ascherio A, Casaceli C, et al. Effect of urateelevating inosine on early Parkinson disease progression: the SURE-PD3 randomized clinical trial. JAMA 2021;326:926–939. [PubMed: 34519802]

34. Investigators PSGS-PI. Isradipine versus placebo in early Parkinson disease: a randomized trial. Ann Intern Med 2020;172:591–598. [PubMed: 32227247]

35. Zou H, Aggarwal V, Stebbins GT, et al. Application of longitudinal item response theory models to modeling Parkinson's disease progression. CPT Pharmacometrics Syst Pharmacol 2022;11: 1382–1392. [PubMed: 35895005]

TABLE 1

Baseline demographic and disease characteristics of the data set

	ALL
Sample size, N	7888
Medication states, N	N = 7888
OFF	1727 (21.9%)
ON	6161 (78.1%)
Age[a]	N = 6280
Mean (SD)	65.65 (10.41)
Education years, N	N = 6356
Mean (SD)	11.27 (5.85)
Sex, N	N = 7843
Female, N (%)	3439 (43.8%)
Male, N (%)	4404 (56.2%)
Language	N = 7888
Arabic	360 (4.56%)
Chinese (simplified)	350 (4.44%)
Chinese (traditional)	344 (4.36%)
Czech	215 (2.73%)
Dutch	302 (3.83%)
Estonian	282 (3.58%)
French	345 (4.37%)
German	392 (4.97%)
Greek	317 (4.02%)
Hebrew	215 (2.73%)
Hindi	356 (4.51%)
Hungarian	357 (4.53%)
Italian	340 (4.31%)
Japanese	293 (3.71%)
Kazakh	362 (4.59%)
Korean	349 (4.42%)
Polish	333 (4.22%)
Portuguese	367 (4.65%)
Romanian	368 (4.67%)
Russian	252 (3.19%)
Slovakian	309 (3.92%)
Spanish	374 (4.74%)
Thai	354 (4.49%)
Turkish	352 (4.46%)
PD diagnosis years, N	N = 7379
Mean (SD)	7.64 (5.74)
Dyskinesias presence, N	N = 7476

	ALL
Yes (%)	2153 (28.8%)
Hoehn and Yahr stage, N (%)	N = 7814
0[b]	46 (0.6%)
1[c]	1055 (13.5%)
2[c]	3832 (49.0%)
3[c]	2060 (26.4%)
4[c]	650 (8.3%)
5[c]	171 (2.2%)
MDS-UPDRS Part 1 sum, mean (SD)	11.95 (7.35)
MDS-UPDRS Part 2 sum, mean (SD)	14.77 (9.67)
MDS-UPDRS Part 3 sum, mean (SD)	33.52 (19.16)

Abbreviations: SD, standard deviation; PD, Parkinson's disease; MDS-UPDRS, The Movement Disorder Society Unified Parkinson's Disease Rating Scale.

[a]The age is calculated as the time difference between birth year and exam year.

[b]On treatment.

[c]On treatment and not on treatment.

TABLE 2

Exploratory factor analysis (EFA) results from one-factor model showing item factor loading for MDS-UPDRS Parts 2 + 3 items

Items	Item Names	Factor 1
2.1	Speech	0.587
2.2	Saliva and drooling	0.455
2.3	Chewing and swallowing	0.568
2.4	Eating tasks	0.693
2.5	Dressing	0.762
2.6	Hygiene	0.748
2.7	Handwriting	0.601
2.8	Doing hobbies and other activities	0.662
2.9	Turning in bed	0.710
2.10	Tremor	
2.11	Getting out of bed	0.760
2.12	Walking and balance	0.726
2.13	Freezing	0.631
3.1	Speech	0.692
3.2	Facial expression	0.687
3.3a	Rigidity—neck	0.625
3.3b	Rigidity—RUE	0.607
3.3c	Rigidity—LUE	0.619
3.3d	Rigidity—RLE	0.658
3.3e	Rigidity—LLE	0.663
3.4a	Finger tapping—right hand	0.726
3.4b	Finger tapping—left hand	0.749
3.5a	Hand movements—right hand	0.764
3.5b	Hand movements—left hand	0.753
3.6a	Pronation-supination—right hand	0.739
3.6b	Pronation-supination—left hand	0.734
3.7a	Toe tapping—right foot	0.753
3.7b	Toe tapping—left foot	0.754
3.8a	Leg agility—right leg	0.789
3.8b	Leg agility—left leg	0.794
3.9	Arising from chair	0.794
3.10	Gait	0.797
3.11	Freezing of gait	0.710
3.12	Postural stability	0.741
3.13	Posture	0.739
3.14	Global spontaneity of movement	0.791
3.15a	Postural tremor—right hand	
3.15b	Postural tremor—left hand	

Items	Item Names	Factor 1
3.16a	Kinetic tremor—right hand	
3.16b	Kinetic tremor—left hand	
3.17a	Rest tremor amplitude—RUE	
3.17b	Rest tremor amplitude—LUE	
3.17c	Rest tremor amplitude—RLE	
3.17d	Rest tremor amplitude—LLE	
3.17e	Rest tremor amplitude—lip/jaw	
3.18	Constancy of rest tremor	

Note: Factor loadings 0.40 are not displayed for clarity purpose.

Abbreviations: MDS-UPDRS, the Movement Disorder Society Unified Parkinson's Disease Rating Scale; RUE, right upper extremity; LUE, left upper extremity; RLE, right lower extremity; LLE, left lower extremity.

TABLE 3

Discrimination (Discrim) and item location parameters of all MDS-UPDRS Parts 2 + 3 items from fitting one IRT model

Items	Item Names	Discrim	0 to 1	1 to 2	2 to 3	3 to 4
2.1	Speech	1.235	0.668	0.588	2.017	4.146
2.2	Saliva and drooling	0.870	0.292	0.717	2.332	4.432
2.3	Chewing and swallowing	1.175	0.495	1.994	3.001	5.951
2.4	Eating tasks	1.634	0.470	0.890	2.329	3.641
2.5	Dressing	2.003	0.911	0.490	1.606	2.625
2.6	Hygiene	1.916	0.636	0.837	1.775	2.710
2.7	Handwriting	1.280	1.279	0.392	1.568	2.872
2.8	Doing hobbies and other activities	1.504	0.812	0.404	1.445	2.436
2.9	Turning in bed	1.714	0.701	0.722	1.718	2.588
2.10	Tremor	**0.492**	2.259	0.871	3.735	6.712
2.11	Getting out of bed	1.990	0.880	0.379	1.359	2.254
2.12	Walking and balance	1.798	1.077	0.461	1.250	2.374
2.13	Freezing	1.385	0.009	0.927	1.801	3.132
3.1	Speech	1.633	−0.87	0.71	2.127	3.646
3.2	Facial expression	1.610	1.392	0.244	1.864	3.441
3.3a	Rigidity—neck	1.363	0.559	0.780	2.220	3.788
3.3b	Rigidity—RUE	1.302	1.113	0.451	2.294	4.388
3.3c	Rigidity—LUE	1.340	0.978	0.486	2.176	4.133
3.3d	Rigidity—RLE	1.488	0.596	0.698	2.219	3.865
3.3e	Rigidity—LLE	1.508	0.551	0.658	2.100	3.756
3.4a	Finger tapping—right hand	1.796	1.300	0.137	1.412	3.029
3.4b	Finger tapping—left hand	1.922	1.221	0.003	1.204	2.795
3.5a	Hand movements—right hand	2.017	1.008	0.281	1.530	2.991
3.5b	Hand movements—left hand	1.947	1.025	0.285	1.514	3.065
3.6a	Pronation-supination—right hand	1.867	1.029	0.298	1.548	2.979
3.6b	Pronation-supination—left hand	1.841	1.084	0.152	1.349	2.804
3.7a	Toe tapping—right foot	1.945	0.987	0.321	1.459	2.738
3.7b	Toe tapping—left foot	1.953	1.041	0.146	1.243	2.524
3.8a	Leg agility—right leg	2.184	0.735	0.511	1.600	2.811

Items	Item Names	Discrim	0 to 1	1 to 2	2 to 3	3 to 4
3.8b	Leg agility—left leg	2.224	0.776	0.382	1.410	2.662
3.9	Arising from chair	2.220	0.003	0.974	1.666	2.441
3.10	Gait	2.243	1.102	0.332	1.472	2.509
3.11	Freezing of gait	1.717	0.524	1.289	2.214	2.938
3.12	Postural stability	1.878	0.220	0.608	1.237	2.529
3.13	Posture	1.868	1.023	0.370	1.585	2.906
3.14	Global spontaneity of movement	2.202	1.221	0.079	1.222	2.629
3.15a	Postural tremor—right hand	**0.532**	0.525	3.832	6.903	11.149
3.15b	Postural tremor—left hand	**0.601**	0.517	3.573	6.505	9.943
3.16a	Kinetic tremor—right hand	**0.517**	1.168	4.782	8.418	12.139
3.16b	Kinetic tremor—left hand	**0.523**	0.970	4.315	7.949	11.264
3.17a	Rest tremor amplitude—RUE	**0.453**	0.884	3.096	6.056	10.336
3.17b	Rest tremor amplitude—LUE	**0.542**	0.910	2.883	5.664	9.312
3.17c	Rest tremor amplitude—RLE	**0.529**	2.630	4.774	7.785	11.998
3.17d	Rest tremor amplitude—LLE	**0.599**	2.655	4.482	7.154	10.553
3.17e	Rest tremor amplitude—lip/jaw	**0.674**	3.083	5.089	7.330	9.221
3.18	Constancy of rest tremor	**0.365**	0.627	1.901	4.213	6.762

Note: 11 tremor items with low and moderate discrimination parameters are in boldface.

Abbreviations: MDS-UPDRS, The Movement Disorder Society Unified Parkinson's Disease Rating Scale; IRT, item response theory; RUE: right upper extremity; LUE, left upper extremity; RLE, right lower extremity; LLE, left lower extremity.

TABLE 4

Exploratory factor analysis (EFA) results from two-factor model showing item factor loading for MDS-UPDRS Parts 2 + 3 items

Items	Item Names	Factor 1	Factor 2
2.1	Speech		0.641
2.2	Saliva and drooling		0.485
2.3	Chewing and swallowing		0.585
2.4	Eating tasks		0.695
2.5	Dressing		0.794
2.6	Hygiene		0.778
2.7	Handwriting		0.620
2.8	Doing hobbies and other activities		0.695
2.9	Turning in bed		0.764
2.10	Tremor	0.532	
2.11	Getting out of bed		0.809
2.12	Walking and balance		0.784
2.13	Freezing		0.700
3.1	Speech		0.709
3.2	Facial expression		0.651
3.3a	Rigidity—neck		0.574
3.3b	Rigidity—RUE		0.524
3.3c	Rigidity—LUE		0.543
3.3d	Rigidity—RLE		0.598
3.3e	Rigidity—LLE		0.612
3.4a	Finger tapping—right hand		0.664
3.4b	Finger tapping—left hand		0.693
3.5a	Hand movements—right hand		0.705
3.5b	Hand movements—left hand		0.696
3.6a	Pronation-Supination—right hand		0.677
3.6b	Pronation-Supination—left hand		0.685
3.7a	Toe tapping—right foot		0.711
3.7b	Toe tapping—left foot		0.723
3.8a	Leg agility—right leg		0.745
3.8b	Leg agility—left leg		0.762
3.9	Arising from chair		0.803
3.10	Gait		0.802
3.11	Freezing of gait		0.736
3.12	Postural stability		0.758
3.13	Posture		0.725
3.14	Global spontaneity of movement		0.744
3.15a	Postural tremor—right hand	0.765	
3.15b	Postural tremor—left hand	0.729	

Items	Item Names	Factor 1	Factor 2
3.16a	Kinetic tremor—right hand	0.707	
3.16b	Kinetic tremor—left hand	0.661	
3.17a	Rest tremor amplitude—RUE	0.853	
3.17b	Rest tremor amplitude—LUE	0.838	
3.17c	Rest tremor amplitude—RLE	0.793	
3.17d	Rest tremor amplitude—LLE	0.785	
3.17e	Rest tremor amplitude—lip/jaw	0.728	
3.18	Constancy of rest tremor	0.875	

Note: Factor loadings 0.40 are not displayed for clarity purpose.

Abbreviations: MDS-UPDRS, the Movement Disorder Society Unified Parkinson's Disease Rating Scale; RUE, right upper extremity; LUE, left upper extremity; RLE, right lower extremity; LLE, left lower extremity.

TABLE 5

Discrimination (Discrim) and item location parameters of all MDS-UPDRS Parts 2 + 3 items from fitting two separate IRT models to 35 non-tremor items (Part 2 items 2.1–2.9, 2.11–2.13, and Part 3 items 3.1–3.14) and 11 tremor items (Part 2 item 2.10 and Part 3 items 3.15a–3.18)

Items	Item Names	Discrim	0 to 1	1 to 2	2 to 3	3 to 4
2.1	Speech	1.271	0.659	0.575	1.980	4.080
2.2	Saliva and drooling	0.884	0.291	0.706	2.302	4.379
2.3	Chewing and swallowing	1.180	0.493	1.987	2.994	5.943
2.4	Eating tasks	1.626	0.472	0.891	2.336	3.667
2.5	Dressing	2.037	0.905	0.484	1.593	2.610
2.6	Hygiene	1.947	0.632	0.829	1.761	2.694
2.7	Handwriting	1.286	1.278	0.390	1.564	2.867
2.8	Doing hobbies and other activities	1.524	0.807	0.400	1.434	2.418
2.9	Turning in bed	1.765	0.692	0.709	1.692	2.554
2.10	Tremor	**1.479**	0.929	0.467	1.639	2.766
2.11	Getting out of bed	2.047	0.870	0.372	1.341	2.226
2.12	Walking and balance	1.852	1.063	0.453	1.231	2.344
2.13	Freezing	1.426	0.005	0.908	1.768	3.082
3.1	Speech	1.663	0.863	0.700	2.106	3.630
3.2	Facial expression	1.603	1.397	0.241	1.866	3.460
3.3a	Rigidity—neck	1.349	0.564	0.782	2.233	3.823
3.3b	Rigidity—RUE	1.271	1.131	0.457	2.326	4.468
3.3c	Rigidity—LUE	1.310	0.992	0.491	2.205	4.205
3.3d	Rigidity—RLE	1.471	0.601	0.701	2.232	3.904
3.3e	Rigidity—LLE	1.498	0.554	0.658	2.105	3.781
3.4a	Finger tapping—right hand	1.772	1.309	0.137	1.420	3.056
3.4b	Finger tapping—left hand	1.899	1.228	0.003	1.209	2.813
3.5a	Hand movements—right hand	1.991	1.013	0.282	1.537	3.014
3.5b	Hand movements—left hand	1.923	1.031	0.286	1.522	3.090
3.6a	Pronation-Supination—right hand	1.840	1.036	0.299	1.556	3.004
3.6b	Pronation-Supination—left hand	1.826	1.088	0.152	1.353	2.820
3.7a	Toe tapping—right foot	1.941	0.988	0.320	1.459	2.745
3.7b	Toe tapping—left foot	1.957	1.040	0.145	1.241	2.523

Items	Item Names	Discrim	0 to 1	1 to 2	2 to 3	3 to 4
3.8a	Leg agility—right leg	2.170	0.737	0.511	1.602	2.822
3.8b	Leg agility—left leg	2.229	0.775	0.381	1.408	2.663
3.9	Arising from chair	2.252	0.005	0.963	1.652	2.430
3.10	Gait	2.274	1.097	0.327	1.462	2.499
3.11	Freezing of gait	1.745	0.517	1.275	2.194	2.918
3.12	Postural stability	1.908	0.221	0.599	1.224	2.513
3.13	Posture	1.869	1.024	0.366	1.582	2.911
3.14	Global spontaneity of movement	2.184	1.226	0.077	1.222	2.640
3.15a	Postural tremor—right hand	**2.168**	0.249	1.450	2.435	3.714
3.15b	Postural tremor—left hand	**2.008**	0.274	1.547	2.640	3.865
3.16a	Kinetic tremor—right hand	**1.845**	0.506	1.840	3.050	4.255
3.16b	Kinetic tremor—left hand	**1.641**	0.449	1.797	3.125	4.307
3.17a	Rest tremor amplitude—RUE	**2.704**	0.337	0.990	1.745	2.736
3.17b	Rest tremor amplitude—LUE	**2.683**	0.397	1.076	1.911	2.917
3.17c	Rest tremor amplitude—RLE	**2.264**	1.020	1.726	2.632	3.852
3.17d	Rest tremor amplitude—LLE	**2.296**	1.126	1.785	2.679	3.786
3.17e	Rest tremor amplitude—lip/jaw	**2.003**	1.485	2.312	3.192	3.931
3.18	Constancy of rest tremor	**3.076**	0.099	0.568	1.079	1.551

Note: 11 tremor items with high and very high discrimination parameters are in boldface.

Abbreviations: MDS-UPDRS Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale; IRT, item response theory; RUE, right upper extremity; LUE, left upper extremity; RLE, right lower extremity; LLE, left lower extremity.