

# Resolving Missing Data from the Movement Disorder Society Unified Parkinson's Disease Rating Scale: Implications for Telemedicine

Sheng Luo, PhD,<sup>1</sup> Christopher G. Goetz, MD,<sup>2</sup> Dongrak Choi, MS,<sup>1</sup> Sanket Aggarwal, MD,<sup>2</sup> Tiago A. Mestre, MD PhD,<sup>3</sup> and Glenn T. Stebbins, PhD<sup>2\*</sup>

<sup>1</sup>Department of Biostatistics & Bioinformatics, Duke University, Durham, North Carolina, USA

<sup>2</sup>Department of Neurological Sciences, Rush University Medical Center, Chicago, Illinois, USA

<sup>3</sup>Ottawa Hospital Research Institute, University of Ottawa Brain and Mind Research Institute, Division of Neurology, Department of Medicine, University of Ottawa, The Ottawa Hospital, Ottawa, Ontario, Canada

**ABSTRACT: Background:** Telemedicine has become standard in clinical care and research during the coronavirus disease 2019 pandemic. Remote administration of Movement Disorder Society Unified Parkinson's Disease Rating Scale (MDS-UPDRS) Part III (Motor Examination) precludes ratings of all items, because Rigidity and Postural Stability (six scores) require in-person rating.

**Objective:** The objective of this study was to determine imputation accuracy for total-sum and item-specific MDS-UPDRS Motor Examination scores in remote administration.

**Methods:** We applied multivariate imputation by chained equations techniques in a cross-sectional dataset where patients had one MDS-UPDRS rating (International Translational Program,  $n = 8,588$ ) and in a longitudinal dataset where patients had multiple ratings (Rush Program,  $n = 396$ ). Successful imputation was stringently defined as (1) generalized Lin's concordance correlation coefficient  $>0.95$ , reflecting near-perfect agreement between total-sum score with complete data and surrogate score, calculated without patients' actual Rigidity and Postural Stability scores; and (2) perfect agreement

for item-level scores for Rigidity and Postural Stability items.

**Results:** For total-sum score when Rigidity and Postural Stability scores were withdrawn, using one or multiple visits, multivariate imputation by chained equations imputation reached near-perfect agreement with the original total-sum score. However, at the item level, the degree of perfect agreement between the surrogate and actual Rigidity items and Postural Stability scores always fell below threshold.

**Conclusions:** The MDS-UPDRS Part III total-sum score, a key clinical outcome in research and in clinical practice, can be accurately imputed without the Rigidity and Postural Stability items that cannot be rated by telemedicine. No formula, however, allows for specific item-level imputation. When Rigidity and Postural Stability item scores are of key clinical or research interest, patients with PD must be scored in person. © 2022 International Parkinson and Movement Disorder Society.

**Key Words:** Parkinson's disease; multiple imputation; rating scale

With the coronavirus disease 2019 (COVID-19) pandemic and a restricted ability to examine patients with Parkinson's disease (PD) in person, telemedicine and

virtual visits have become a necessary step to assure quality care and safe distancing.<sup>1</sup> The MDS-UPDRS (Movement Disorder Society Unified Parkinson's Disease Rating Scale) is the most widely used scale for measuring parkinsonian symptoms in clinical and research practice. It was designed to be completed in the office setting, but if it could be applied even in a telemedicine setting, accurate ongoing monitoring could ensue without interruption.<sup>2</sup> The scale consists of four parts (Parts I–IV) that measure progression in patients with PD. The item scores in each part are recorded using a Likert scale from 0 to 4 with an integer rating, and the higher scores indicate more severe status.

\*Correspondence to: Dr. Glenn T. Stebbins, Department of Neurological Sciences, Rush University Medical Center, 1725 West Harrison Street Suite 755, Chicago, IL, USA; E-mail: [glenn\\_stebbins@rush.edu](mailto:glenn_stebbins@rush.edu)

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

Full financial disclosures and author roles may be found in the online version of this article.

**Received:** 7 March 2022; **Accepted:** 30 May 2022

Published online in Wiley Online Library ([wileyonlinelibrary.com](http://wileyonlinelibrary.com)). DOI: 10.1002/mds.29129

Part III measures parkinsonian motor features with 33 items and has instructions for the rater to conduct the various motor tests. Among the items of Part III, the severity scores of Rigidity (3.3a–3.3e, five items) and Postural Stability (3.12, one item), a total of six items, require an in-person physical assessment by raters. To determine a patient's rigidity of the limbs and neck, the rater assesses the patient's resistance to passive movement when in a relaxed state by placing their hands at the patients' major joints and moving the limbs.<sup>3</sup> To assess Postural Stability, the rater produces an unexpected displacement of the patient's body by pulling backward on the patient's shoulders, thus requiring a "hands-on" examination.<sup>3</sup>

Given that completing the Part III motor examination requires the examiner to make physical contact with the patient for Rigidity and Postural Stability testing, these items create an application problem for telemedicine visits. To use MDS-UPDRS Part III in a telemedicine setting, the values for Rigidity and Postural Stability would need to be imputed in some manner. In an earlier study, we noted that a reliable proxy score for Part III could not be obtained if three or more Part III scores were missing in a consistent fashion.<sup>4</sup> The proxy method used in that study did not use newer, multiple imputation (MI) approaches that might provide an adequate estimate for a total score in spite of missing values. In this project, using both cross-sectional and longitudinal data, we removed Rigidity and Postural Stability scores from MDS-UPDRS assessments and then applied MI methods to derive surrogate Part III scores (total-sum and item-specific scores for Rigidity and Postural Stability measurements) to replicate the clinical or research setting of a telemedicine-based MDS-UPDRS rating. To qualify as an acceptable surrogate estimate, we purposefully applied a rigorous statistical criterion that could be argued at the regulatory level to be a near-perfect match to the in-person rating method.

## Subjects and Methods

### Study Populations

#### **Cross-Sectional Study Population**

This study used the cross-sectional data from the MDS-UPDRS translation program, a dataset of 8,931 complete MDS-UPDRS ratings from patients with PD (UK Brain Bank Criteria<sup>5</sup>) representing all Hoehn and Yahr stages with assessments performed in the patient's native language (24 international languages, not including English). All raters in this program were trained in the MDS-UPDRS Training Program, and all translations had undergone a five-step process including the field testing of 350 native-speaking patients with PD.<sup>6</sup> The MDS-UPDRS was administered in a single visit,

and this visit was the source rating from which the Rigidity and Postural Stability scores were systematically removed for the imputation application. We deleted 343 patients with missing scores in some MDS-UPDRS Part III items, and a total of 8,588 patients (referred to as "cross-sectional dataset") were included for analysis.

#### **Longitudinal Study Population**

This study used longitudinal clinical data from Rush Parkinson's Disease and Movement Disorders Program collected between March 5, 2007, and August 26, 2010. In this dataset, there was a total of 813 patients with complete MDS-UPDRS Part III scores with a total of 9,448 visits. The median duration between each visit was 155 days, with a minimum of 1 day and a maximum of 182 days. Only patients who had visits that were longer than 60 days apart were included, leaving a total of 396 patients (referred to as "longitudinal dataset") for further analysis. The MDS-UPDRS was measured as part of regular clinical care by one clinician (C.G.G.) at approximately 6-month intervals with the number of follow-up visits ranging from 1 to 16 (median, 5).

### Statistical Analysis

The MI was applied both in a cross-sectional setting when each patient with PD has MDS-UPDRS measured in a single visit and in a longitudinal setting when each patient with PD has MDS-UPDRS measured in multiple visits.

#### **MI in Cross-Sectional Data**

To mimic missing scores in Rigidity (3.3a–3.3e) and Postural Stability (3.12) items, we selected one patient from the cross-sectional dataset and deleted the observed scores of these items. We conducted MI for this patient with missing data using multivariate imputation by chained equations (MICE).<sup>7</sup> This process was repeated until each patient's Rigidity and Postural Stability items were deleted once and imputed by MICE. For external validation of the MI model developed using the cross-sectional dataset, we used the data from one target visit in the longitudinal dataset.

#### **MI in Longitudinal Data**

In the longitudinal dataset, where the individual's MDS-UPDRS scorable values are used in conjunction with his/her own prior visit scores, we tested whether prior complete scores including prior Rigidity and Postural Stability ratings allowed an accurate imputed value. To perform MI in the longitudinal setting, we selected visit 4 (referred to as "target visit") from the longitudinal dataset, and we based this decision on the

clinical applicability of three prior visits as a reasonable background reference and the large number of available patients ( $n = 242$ ). Nonetheless, we also used all patient data relative to visit 1 ( $n = 396$ ), visit 2 ( $n = 365$ ), and visit 3 ( $n = 289$ ). Supporting Information Table S1 displays the number of patients for each visit in the longitudinal dataset. For the longitudinal analysis, we then selected one patient from the target visit and deleted the observed scores of Rigidity and Postural Stability items. We conducted MI using a different amount of information from various other visits in the following two scenarios: (1) used the visit immediately before the target visit (i.e., visit 3 only), or (2) used all visits preceding the target visit (i.e., visits 1–3). In these two scenarios, the MI model also used this patient’s data from other Part III items in the target visit. This process was repeated until each patient’s Rigidity and Postural Stability item scores at the target visit were deleted once and imputed.

### MI Performance Assessment

In both cross-sectional and longitudinal datasets, we conducted MI with 21 replications using MICE implemented in R package *mice*.<sup>8</sup> We assessed the performance of MI in both the total-sum score of MDS-UPDRS Part III and the individual item scores of the six items (five Rigidity and Postural Stability) that cannot be acquired by telemedicine assessment. To measure the exact matching between the original Part III total-sum score (the “gold standard”) and the imputed total-sum score, we adopted the generalized concordance correlation coefficient (CCC),<sup>9,10</sup> a U-statistics-based nonparametric test that is more robust against outliers and nonnormal continuous data such as MDS-UPDRS Part III total-sum, as compared with Lin’s CCC.<sup>11,12</sup> The generalized CCC integrates accuracy and correlation, and it examines how closely the imputed total-sum score agrees with the original total-sum score. Based on 21 replications of the imputed datasets, we obtained 21 generalized CCCs, which were combined to create one multiple-imputation inference by Rubin’s MI rules,<sup>13</sup> with the overall generalized CCC being estimated by the mean of 21 generalized CCCs and 95% confidence interval (CI) being computed from the pooled variance, for example, sum of within- and between-imputation variances. Sensitive to concerns that regulatory agencies do not easily accept imputation strategies or missing data allowances, we set our critical generalized CCC level at  $>0.95$ , interpreted as near-perfect agreement between the imputed value-based and the full data-based total-sum scores. This high bar also echoes the statistical stringency of our prior work.<sup>4</sup> The generalized CCC approach cannot be used to assess the imputed individual item scores because it is applicable only to total-sum scores, not individual item scores. Instead, we computed the degree of

perfect agreement between the surrogate calculation and the actual score, that is, the percentage of results for which the imputed item score was identical to the original score before deletion. We set the critical degree of perfect agreement level at  $\geq 0.95$ . In clinical practice it is quite common to see MDS-UPDRS item scores fluctuate one level above or below a given prior score without alarm and ascribed to natural variability. Even if such an estimate would not be rigorous enough for research applications, we tested this “ $\pm 1$ ” option as a feasibility test for possible clinical use (the imputed item score being identical to or one level above or below the original score before deletion). In all instances, we set our critical degree of perfect agreement at  $>0.95$ .

## Results

### Study Sample

Our study used all 8,588 cross-sectional records of MDS-UPDRS Part III scores from the original set, with full data on all 33 items. Supporting Information Figure S1 displays a histogram of the total-sum score of 33 items (left) and a bar plot of median scores of 33 items for all patients (right), respectively. The mean of the Part III total-sum score was 33.5, with the minimum being 0 and the maximum being 132. The frequencies of the median scores of 33 items for all subjects are also displayed in Supporting Information Figure S1.

The longitudinal data used 396 subjects’ records of MDS-UPDRS Part III scores with full data on all 33 items. Supporting Information Figure S2 displays histograms of the total-sum score of 33 items and bar plots of median scores of 33 items at baseline and at the target visit. The mean of Part III total-sum score was 40.2, with the minimum being 0 and the maximum being 108 for all subjects at baseline, whereas the mean total-sum score was 40.0 with a range of 6–91 at the target visit. The frequencies of the median scores of 33 items for all subjects both at baseline and at the target visit are also displayed in Supporting Information Figure S2. Also refer to Supporting Information Table S2 for the distributions of Hoehn and Yahr stages both at baseline and at the target visit.

Table 1 displays the baseline demographic and disease characteristics of the cross-sectional dataset and the longitudinal dataset. Overall, both sets cover all Hoehn and Yahr stages, with the majority of patients being Hoehn and Yahr stages 2 and 3 reflecting patterns seen in clinical practice.<sup>14</sup>

### Total-Sum Score Imputation

#### Cross-Sectional Data

In the cross-sectional dataset, where the individual’s MDS-UPDRS scorable values were used in conjunction

with other patients' complete but single-visit scores. The generalized CCC for imputed total-sum score with the total-sum score as determined by the "hands-on" inclusion of Rigidity and Postural Stability scores by in-person examination is 0.976 (95% CI: 0.975–0.977), which meets the threshold of near-perfect agreement (Table 2, upper row). As a validation test, we applied the MI model developed from the cross-sectional

**TABLE 1** Baseline demographic and disease characteristics of the cross-sectional dataset and longitudinal dataset

Variables	Cross-sectional dataset (n = 8,588)	Longitudinal dataset (n = 396)
Male, n (%)	4,769 (56.0)	265 (66.9)
Age, mean (SD), y	65.7 (10.6)	69.5 (10.5)
MDS-UPDRS Part III score, mean (SD)	33.5 (19.1)	40.2 (15.7)
H&Y stage, <sup>a</sup> n (%)		
0	51 (0.59)	4 (1.0)
1	1,216 (14.2)	2 (0.5)
2	4,145 (48.3)	249 (62.9)
3	2,205 (25.7)	79 (19.9)
4	692 (8.1)	41 (10.4)
5	181 (2.1)	21 (5.3)

<sup>a</sup>Note that H&Y stage numbers for the cross-sectional dataset were based on 3.20 (H&Y stage) measure after deleting 96 subjects with missing data and subjects with the following H&Y scores: 1 with H&Y = 2.5 and 1 with H&Y = 6. SD, standard deviation; MDS-UPDRS, Movement Disorder Society Unified Parkinson's Disease Rating Scale; H&Y, Hoehn and Yahr.

**TABLE 2** Generalized CCC of MDS-UPDRS Part III total-sum score and degree of perfect agreement (the % of having the imputed item score being identical to the original score before deletion) in the cross-sectional dataset and longitudinal dataset visit 4

	Cross-sectional dataset (n = 8,588)		Longitudinal dataset visit 4 (n = 242)	
	Generalized CCC	Acceptance (>0.95)	Generalized CCC	Acceptance (>0.95)
Sum score	0.976 <sup>a</sup>	Yes	0.966 <sup>b</sup>	Yes
Item scores	Perfect agreement	Acceptance (>0.95)	Perfect agreement	Acceptance (>0.95)
Rigidity – Neck	0.460	No	0.438	No
Rigidity – RUE	0.515	No	0.517	No
Rigidity – LUE	0.505	No	0.603	No
Rigidity – RLE	0.492	No	0.483	No
Rigidity – LLE	0.484	No	0.554	No
Postural stability	0.564	No	0.599	No

<sup>a</sup>95% CI: 0.975–0.977.

<sup>b</sup>95% CI: 0.958–0.974.

CCC, concordance correlation coefficient; MDS-UPDRS, 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; CI, confidence interval.

dataset to a single assessment focused on target visit 4 in the longitudinal dataset. Likewise, the model met the near-perfect bar with a generalized CCC of the original versus imputed Part III total-sum scores (target visit 4) of 0.966 (95% CI: 0.958–0.974).

**Longitudinal Data**

In scenario 1, only the visit immediately before the target visit was used in addition to the target visit to perform MI. The generalized CCC is 0.970 (95% CI: 0.956–0.984), which meets the threshold of near-perfect agreement (Table 3, upper row). In scenario 2, where all visits preceding the target visit were used to perform MI, the generalized CCC is 0.976 (95% CI: 0.953–0.998), which also meets the threshold. In both longitudinal scenarios, we obtained the same near-perfect agreement standard as seen with the "big-data" approach in the cross-sectional dataset.

To facilitate the clinical use of MI when "hands-on" assessments are not feasible, we are developing a web-based interactive calculator (a beta version can be found at: <https://francisco-dongrak-choi.shinyapps.io/MISC/>), which takes as input an individual's scores on all MDS-UPDRS Part III items, except Rigidity (3.3a–3.3e) and Postural Stability (3.12) items. It conducts MI with 21 replications and produces the surrogate total-sum score for the MDS-UPDRS Part III with a 95% CI.

**Individual Item Score Imputation**

**Cross-Sectional Data.** In contrast with the successful total score generation from the MICE calculations, the degree of perfect agreement of the six individual scores covering Rigidity and Postural Stability fell below threshold in the cross-sectional analyses. For the

**TABLE 3** Generalized CCC of MDS-UPDRS Part III total-sum score and degree of perfect agreement (the percentage of having the imputed item score being identical to the original score before deletion) under scenarios 1 and 2 using the longitudinal dataset

	Scenario 1		Scenario 2	
	Generalized CCC	Acceptance(>0.95)	Generalized CCC	Acceptance(>0.95)
Sum score	0.970 <sup>a</sup>	Yes	0.976 <sup>b</sup>	Yes
Item scores	Perfect agreement	Acceptance (>0.95)	Perfect agreement	Acceptance (>0.95)
Rigidity – Neck	0.545	No	0.612	No
Rigidity – RUE	0.678	No	0.740	No
Rigidity – LUE	0.640	No	0.723	No
Rigidity – RLE	0.450	No	0.723	No
Rigidity – LLE	0.417	No	0.661	No
Postural stability	0.583	No	0.719	No

Scenario 1: used the visit immediately before the target visit 4; scenario 2: used all visits preceding the target visit 4.

<sup>a</sup>95% CI: 0.956–0.984.

<sup>b</sup>95% CI: 0.953–0.998.

CCC, concordance correlation coefficient; MDS-UPDRS, 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; CI, confidence interval.

cross-sectional dataset and single visit 4 assessment, the results suggest agreement below the perfect agreement benchmark and failure of MI strategies to recover the score for any of the individual items (Table 2, lower rows).

**Longitudinal Data.** The degree of perfect agreement of the six individual scores covering Rigidity and Postural Stability fell below threshold in both scenarios where prior data (immediate visit or scenario 1 and all prior visits or scenario 2) were accessed, suggesting agreement below perfect agreement benchmark and failure of MI strategies to consistently recover the score for any of the individual items (Table 3, lower rows).

**Perfect ± One Level Analysis**

If we allowed imputed scores to fluctuate one level above or below the true scores, the degree of agreement of some Rigidity items (e.g., Rigidity – Neck, Rigidity – Right Lower Extremity [RLE], and Rigidity – Lower Left Extremity [LLE] in the cross-sectional dataset) and Postural Stability (in the cross-sectional dataset and longitudinal dataset target visit 4) still fell below threshold (Supporting Information Table S3). In the longitudinal dataset, the degree of agreement of some Rigidity items (e.g., Rigidity – Neck, Rigidity – RLE, and Rigidity – LLE in scenario 1) and Postural Stability in both scenarios still fell below threshold (Supporting Information Table S4). Hence, whether in both cross-sectional dataset analyses or in the longitudinal analyses when the prior visit data were considered (one visit or multiple visits), accurate imputed item scores could not be reliably generated in its entirety.

**Discussion**

The MDS-UPDRS is undoubtedly the most widely used rating scale in PD and is endorsed by the MDS and the National Institute of Neurological Disorders and Stroke Common Data Elements program as the recommended rating scale to measure PD disability.<sup>15</sup> It was designed to be applied in both clinical and research settings. Whereas most of the item scores are derived by observation, six scores require that the rater be present for “hands-on” assessments of Rigidity (five items) and Postural Stability. When the MDS-UPDRS is used in telemedicine, which is increasingly common during the COVID-19 pandemic and may become more frequent in the future, the absence of “hands-on” access causes missing data. In our prior work that focused on the MDS-UPDRS Motor Examination (Part III) total-sum score,<sup>4</sup> we identified a clear threshold for how many missing items can exist in Part III and still allow near-perfect agreement between the missing value-based and the full data-based scores. Using traditional statistical analyses, a consistent loss of all Rigidity scores and Postural Stability fell outside the reliable range for a valid surrogate substituted score. Further, in another prior study,<sup>16</sup> we did not examine individual item substitutions because we failed in the goal of generating a surrogate total-sum score. This study provides a method to impute a total-sum score for MDS-UPDRS Part III even when “hands-on” assessments are not feasible. Whereas traditional methods did not permit this success, the MICE approach offers a new statistical strength that can be immediately accessed for direct total-sum score application. By demonstrating high generalized CCC, we have shown that, at the individual patient level, this

imputation approach provides an accurate surrogate total-sum score that is highly correlated to the scores of comparable patients from our databases. Although we used a threshold generalized CCC of 0.95 and a perfect matching score would result in a generalized CCC of 1, at both statistical and clinical levels, these imputations can be considered near perfect. For the first time, we have a method whereby a telemedicine strategy can potentially be used either exclusively or mixed with in-person visits in clinical trials to obtain accurate total-sum scores of MDS-UPDRS Part III. Of interest, longitudinal data of prior visits from the patient are not really needed to bolster confidence with this powerful MICE approach, and no statistical advantage is seen when using additional personalized data from prior visits.

These results are important because the US Food and Drug Administration and European Medicines Agency have focused on the total-sum score for MDS-UPDRS Part III in prior regulatory evaluations. There has never been an expressed focus on the scores of individual items or factors. Therefore, to the extent that imputation methodologies would be judged on the credibility and power of the statistical results without preconceived bias, we offer the strong recommendation that a final and statistically valid MDS-UPDRS Part III total-sum score can be generated despite select missing data, as necessarily evidenced with telemedicine examinations.

However, we cannot offer a solution, even with these high-powered tools, to accurately impute at the individual item-score level. Regardless of using cross-sectional or longitudinal data, even with current and three prior visits with complete data, we could not arrive at reliable or valid surrogate substituted item scores. We can only conclude that in the situation where there is an essential focus on Rigidity or Postural Stability, the patient must be examined “hands-on.”

We recognize that in some clinical situations, a physician may be willing to accept a slightly imperfect MDS-UPDRS item score within the  $\pm$  one level range. Normal fluctuations in parkinsonism and rating variability may explain some item score changes that do not specifically translate into clinical improvement or decline. At the narrative level, this approach might mean: “Given that the patient actually has left arm Slight Rigidity (score 1), defined as appreciable with the facilitation maneuver only, I could accept Absent (0) or Mild (2) Rigidity for my clinical decision-making, and consider such estimates acceptable.” This solution is subperfect mathematically and would not likely be applicable to any research endeavor but might still be acceptable in many clinical contexts. However, even with these compromises, the Postural Stability item, key to many decisions in terms of fall risk and patient autonomy, cannot be imputed accurately, requiring the

specific “pull test” maneuver that is not considered safe outside of professional training and supervision. Moreover, accurate imputed scores of some Rigidity items (e.g., Rigidity – Neck, Rigidity – RLE, and Rigidity – LLE) could not be reliably generated without “hands-on” assessments. Given that Rigidity and Postural Stability represent two of the cardinal features of PD, we accept that a simplified MDS-UPDRS Part III score that excludes their assessment and multiplies the 27 available item scores by a standard formula would be inappropriate. Our MI method imputes a total-sum MDS-UPDRS Part III that accommodates the representation of rigidity and postural stability function based on actual patients with similar scores from our large database.

The application of our method to telemedicine presumes that the professional ratings of two-dimensional virtual data are comparable with those generated from three-dimensional office visits. The feasibility and safety of remotely assessing PD, albeit without Rigidity or Postural Stability, has been demonstrated,<sup>2,17-19</sup> and direct comparison of in-home, video-based versus in-person administration of UPDRS Part III demonstrated moderate overall agreement.<sup>17</sup>

The cross-sectional dataset has the strong advantage of a large “big data” resource acquired internationally by raters who all had passed a certificate training program and contributed cases that spanned all Hoehn and Yahr stages. The longitudinal set was smaller, but all patients were examined and rated prospectively by the same rater who likewise was specifically trained in the MDS-UPDRS (C.G.G.). We accept that this latter database is smaller, but it has the merit of repeated observations in time and includes all patients with PD in a single practice, including those monitored without medications, those on medications, and those with advanced therapies including surgery. As a university-based practice, however, it still may not reflect a fully generalized PD population.

MICE is a commonly used method of imputing missing data because of its usefulness for dealing with large datasets and availability of software packaging.<sup>7,8</sup> MDS-UPDRS is a primary end point of many PD clinical trials (e.g., completed SURE-PD3 study<sup>20</sup> and *Helicobacter pylori* eradication study,<sup>21</sup> and the ongoing TEMPO-2 study, [ClinicalTrials.gov: NCT04223193](https://clinicaltrials.gov/ct2/show/study/NCT04223193)) and observational studies (e.g., Parkinson’s Progression Markers Initiative study,<sup>22</sup> Oxford Parkinson Disease Center discovery cohort,<sup>23</sup> and Tracking study<sup>24</sup>). Because telemedicine, once regarded as an important supplementary tool, is now becoming a nearly essential component of care, strategies to allow for a valid MDS-UPDRS administration have emerged as a high priority. MICE is particularly useful to impute total-sum scores despite the inevitable missing data in telemedicine from items that require “hands-on” assessments. In an era of COVID-19, research inclusion, and emphasis on access,

the response of regulatory agencies to analyses and applications of this type may open the possibility of research study visits by telemedicine with an acceptance of total-sum scores from Part III by surrogate MICE calculations. Still, unstudied questions concern the inclusion of patient-related data, such as disease duration, age, sex, and other factors, that could potentially impact individual item scores or score patterns and thereby allow for still unachieved imputation solutions. ■

**Acknowledgments:** The Rush Parkinson's Disease and Movement Disorders Program is a designated Clinical Center of Excellence supported by the Parkinson's Foundation. S.L.'s research was supported by the National Institute on Aging (Grants R01AG064803, P30AG072958, and P30AG028716). This research was supported by the International Parkinson and Movement Disorder Society and Consolidated Anti-Aging Foundation.

### 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

- Goetz CG, Stebbins GT, Luo S. Movement disorder society-unified Parkinson's disease rating scale use in the Covid-19 era. *Mov Disord* 2020;35(6):911. doi: <https://doi.org/10.1002/mds.28094> [published Online First: 20200511].
- Schneider RB, Myers TL, Tarolli CG, et al. Remote administration of the MDS-UPDRS in the time of COVID-19 and beyond. *J Parkinsons Dis* 2020;10(4):1379–1382. <https://doi.org/10.3233/JPD-202121>
- 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):2129–70. doi: <https://doi.org/10.1002/mds.22340> [published Online First: 2008/11/26].
- Goetz CG, Luo S, Wang L, et al. Handling missing values in the MDS-UPDRS. *Mov Disord* 2015;30(12):1632–8. doi: <https://doi.org/10.1002/mds.26153> [published Online First: 20150204].
- Gibb WR, Lees AJ. The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. *J Neurol Neurosurg Psychiatry* 1988;51(6):745–752. <https://doi.org/10.1136/jnnp.51.6.745>
- Goetz CG, Stebbins GT, Wang L, et al. MDS-sponsored scale translation program: process, format, and clinimetric testing plan for the MDS-UPDRS and UDysRS. *Mov Disord Clin Pract* 2014;1(2):97–101.
- Azur MJ, Stuart EA, Frangakis C, et al. Multiple imputation by chained equations: what is it and how does it work? *Int J Methods Psychiatr Res* 2011;20(1):40–49.
- Van Buuren S, Groothuis-Oudshoorn K. Mice: multivariate imputation by chained equations in R. *J Stat Softw* 2011;45:1–67.
- King TS, Chinchilli VM. Robust estimators of the concordance correlation coefficient. *J Biopharm Stat* 2001;11(3):83–105.
- King TS, Chinchilli VM. A generalized concordance correlation coefficient for continuous and categorical data. *Stat Med* 2001;20(14):2131–2147.
- Lawrence I, Lin K. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 1989;45(1):255–268.
- Lawrence I, Lin K. Assay validation using the concordance correlation coefficient. *Biometrics* 1992;48(2):599–604.
- Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons; Hoboken, New Jersey, 2004.
- Sato K, Hatano T, Yamashiro K, et al. Prognosis of Parkinson's disease: time to stage III, IV, V, and to motor fluctuations. *Mov Disord* 2006;21(9):1384–1395. <https://doi.org/10.1002/mds.20993>
- Regnault A, Boroojerdi B, Meunier J, et al. 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(8):1927–1936.
- de Siqueira Tosin MH, Goetz CG, Luo S, et al. Item response theory analysis of the MDS-UPDRS motor examination: tremor vs. nontremor items. *Mov Disord* 2020;35(9):1587–1595.
- Cubo E, Gabriel-Galan JM, Martinez JS, et al. Comparison of office-based versus home web-based clinical assessments for Parkinson's disease. *Mov Disord* 2012;27(2):308–11. doi: <https://doi.org/10.1002/mds.24028> [published Online First: 20111215].
- Dorsey ER, Wagner JD, Bull MT, et al. Feasibility of virtual research visits in fox trial finder. *J Parkinsons Dis* 2015;5(3):505–515. <https://doi.org/10.3233/JPD-150549>
- Dorsey ER, Deuel LM, Voss TS, et al. Increasing access to specialty care: a pilot, randomized controlled trial of telemedicine for Parkinson's disease. *Mov Disord* 2010;25(11):1652–1659. <https://doi.org/10.1002/mds.23145>
- Schwarzschild MA, Ascherio A, Casaceli C, et al. Effect of urate-elevating inosine on early Parkinson disease progression: the SURE-PD3 randomized clinical trial. *JAMA* 2021;326(10):926–39. doi: <https://doi.org/10.1001/jama.2021.10207> [published Online First: 2021/09/15].
- Tan AH, Lim SY, Mahadeva S, et al. Helicobacter pylori eradication in Parkinson's disease: a randomized placebo-controlled trial. *Mov Disord* 2020;35(12):2250–60. doi: <https://doi.org/10.1002/mds.28248> [published Online First: 20200907].
- Marek K, Jennings D, Lasch S, et al. The Parkinson progression marker initiative (PPMI). *Prog Neurobiol* 2011;95(4):629–635.
- Szewczyk-Krolikowski K, Tomlinson P, Nithi K, et al. The influence of age and gender on motor and non-motor features of early Parkinson's disease: initial findings from the Oxford Parkinson disease center (OPDC) discovery cohort. *Parkinsonism Relat Disord* 2014;20(1):99–105. doi: <https://doi.org/10.1016/j.parkreldis.2013.09.025> [published Online First: 20131012].
- Malek N, Swallow DM, Grosset KA, et al. Tracking Parkinson's: study design and baseline patient data. *J Parkinsons Dis* 2015;5(4):947–959. <https://doi.org/10.3233/JPD-150662>

### Supporting Data

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