Published in final edited form as:
Mov Disord. 2011 August 15; 26(10): 1859–1863. doi:10.1002/mds.23740.

The Modified Bradykinesia Rating Scale for Parkinson's Disease: Reliability and Comparison with Kinematic Measures

Dustin A. Heldman, PhD[1], **Joseph P. Giuffrida, PhD**[1], **Robert Chen, MBBChir, MSc, FRCPC**[2], **Megan Payne, MS, MBA**[3], **Filomena Mazzella, BScN, RN**[2], **Andrew P. Duker, MD**[3], **Alok Sahay, MD**[3], **Sang Jin Kim, MD**[2], **Fredy J. Revilla, MD**[3], and **Alberto J. Espay, MD, MSc**[3,*]

[1]Division of Movement Disorders, Cleveland Medical Devices, Inc., Cleveland, Ohio, USA

[2]Toronto Western Research Institute and Division of Neurology, Department of Medicine, University of Toronto, Toronto, Ontario, Canada

[3]UC Neuroscience Institute, Department of Neurology, Gardner Center for Parkinson's Disease and Movement Disorders, University of Cincinnati, Cincinnati, Ohio, USA

Abstract

Bradykinesia encompasses slowness, decreased movement amplitude, and dysrhythmia. Unified Parkinson's Disease Rating Scale–based bradykine-sia-related items require that clinicians condense abnormalities in speed, amplitude, fatiguing, hesitations, and arrests into a single score. The objective of this study was to evaluate the reliability of a modified bradykinesia rating scale, which separately assesses speed, amplitude, and rhythm and its correlation with kinematic measures from motion sensors. Fifty patients with Parkinson's disease performed Unified Parkinson's Disease Rating Scale–directed finger tapping, hand grasping, and pronation–supination while wearing motion sensors. Videos were rated blindly and independently by 4 clinicians. The modified bradykinesia rating scale and Unified Parkinson's Disease Rating Scale demonstrated similar inter- and intrarater reliability. Raters placed greater weight on amplitude than on speed or rhythm when assigning a Unified Parkinson's Disease Rating Scale score. Modified bradykinesia rating scale scores for speed, amplitude, and rhythm correlated highly with quantitative kinematic variables. The modified bradykinesia rating scale separately captures bradykinesia components with interrater and intrarater reliability similar to that of the Unified Parkinson's Disease Rating Scale. Kinematic sensors can accurately quantify speed, amplitude, and rhythm to aid in the development and evaluation of novel therapies in Parkinson's disease.

Keywords

Parkinson's disease; bradykinesia; Unified Parkinson's Disease Rating Scale; modified bradykinesia rating scale; KinetiSense

Bradykinesia is the defining motor symptom in Parkinson's disease (PD). The term *bradykinesia* has been variably applied to delays or hesitations in initiating movements and

*Correspondence to: Dr. Alberto J. Espay, Department of Neurology, University of Cincinnati, 260 Stetson St., Suite 2300 (PO Box 670525), Cincinnati, OH 45267-0525, USA; alberto.espay@uc.edu.

Additional Supporting Information may be found in the online version of this article.

Relevant conflicts of interest/financial disclosures: Full financial disclosures and author roles may be found in the online version of this article.

slowness in executing movements and may be used interchangeably with the terms *akinesia* and *hypokinesia*.[1–3] From a purist perspective, bradykinesia only refers to slowness of movement; however, poverty of spontaneous movement (akinesia) and smaller amplitude of movement (hypokinesia) tend to be grouped within the same construct. Although nominally referred to as bradykinesia in the literature (and henceforth in this article), distinct movement features often coexist in a given patient[1,4] and may deserve separate measurement, given potentially distinct correlations with disease severity and response to treatment.[5]

The multiple components included in bradykinesia add greater complexity to the rating task than a monolithic manifestation such as tremor.[6–8] The standard rating for bradykinesia is a qualitative clinician assessment and score assignment (0–4) based on tasks 23–25 from the Unified Parkinson's Disease Rating Scale (UPDRS-III) motor subscale. The extraction and separate scoring of these items is a common, albeit never validated practice but is assumed to properly capture the construct of bradykinesia as a stand-alone end point in clinical trials. Unfortunately, bradykinesia-related items have the lowest reliability among all UPDRS items. Interrater reliability for finger tapping yielded Kappa coefficients below 0.50[9,10] and led to poor to fair agreement within raters in 2 of 3 studies[11,12] for the "bradykinesia motoric domain" (items 23–26 and 31).[13] One study showed fair to good agreement (κ_w = 0.61–0.69), but early, untreated PD patients were rated without blinding.[14] These studies suggest that UPDRS-III may be insufficient in its evaluative properties for bradykinesia. A major source of rating variability within and between clinician raters may be the need to convert speed impairment, amplitude impairment, fatiguing, hesitations, and arrests in movement into a single score. Clinicians may place variable weight on various components. Combining multiple movement features into a single score, the UPDRS not only dilutes the power of finding true changes but may result in a differential response becoming unnoticed when evaluating the overall "bradykinesia" outcome of clinical trials. Certain therapies may differentially improve specific components of movement impairment. For example, levodopa may normalize bradykinesia to a greater extent than hypokinesia.[5,15]

Previous attempts to objectively quantify bradykinesia have used electromyography,[4] accelerometers,[2,16] gyroscopes,[7] electromagnetic tracking,[5] magnetic coils,[17,18] and timed motor activities[19] to measure movement rate and/or time to complete a task, but rarely separate bradykinesia into subcomponents or relate the measurement to a standardized rating scale. In addition, quantitative digitography using repetitive alternating finger tapping on a keyboard correlated moderately well with both the overall UPDRS-III and UPDRS-III bradykinesia subscore, but was not compared with individual UPDRS tasks.[20] Recently, the modified bradykinesia rating scale (MBRS) was introduced to independently rate speed, amplitude, and rhythm components of bradykinesia.[15] We aimed to evaluate inter- and intrarater reliability of the MBRS in patients with PD and to assess validity by comparing it to UPDRS scores and kinematic data recorded using motion sensors.

Patients and Methods

We recruited 50 patients with idiopathic PD meeting research diagnostic criteria[21] (Supporting Information Table 1). Subjects were videoed performing UPDRS-directed finger tapping, hand grasping, and pronation–supination tasks in the OFF (12–15 hours after dopaminergic drug withdrawal) and ON states while wearing wireless 6-degree-of-freedom motion sensors (KinetiSense, CleveMed, Cleveland, OH) on the index finger and thumb (Supplementary Fig. 1). Patients were asked to perform each of the 3 tasks by the more affected limb for 15 seconds with as large an amplitude and as fast movements as possible.

The videos were randomized for independent evaluation by 4 movement disorders neurologists who used the UPDRS and MBRS to score each task. The MBRS was developed by Kishore et al[15] for scoring speed, amplitude, and rhythm separately (Supporting Information Table 2). Approximately 4 weeks after scoring the videos, the same clinicians rescored the videos (rerandomized) to examine intrarater reliability. After the second scoring, the clinicians held a group training session using 15 videos containing each of the tasks (not included in the study data) in an attempt to normalize severity ratings across clinicians. Approximately 2 weeks after this training session, the videos were rerandomized and scored a third and final time by all 4 clinicians.

We assessed both agreement between clinicians (interrater reliability) as well as agreement of repetitions of ratings by each individual clinician (intrarater reliability). Scores for each MBRS subtask were correlated with their corresponding UPDRS scores to determine which movement components were given greater subjective weight when assigning a UPDRS score. MBRS scores were compared with several quantitative features extracted from the 2 motion sensors in order to examine their validity (extent to which they measure what they intend to measure).

Further details of the methods are available in the online Supporting Material.

Results

Reliability and UPDRS Correlation

Each clinician's scores were compared to the average of the other 3 to measure interrater reliability (Fig. 1, Table 1). Correlation coefficients and RMS errors between clinicians for MBRS and UPDRS scores were comparable on average. However, there were tendencies for worse MBRS speed and better MBRS amplitude interrater reliability. Training did not improve correlation between raters; however, RMS errors decreased for speed and amplitude ($P < .05$). Intrarater reliability between the first and second scoring sessions were similar and comparable between UPDRS and MBRS items (Table 1).

To examine how raters considered various movement features when assigning UPDRS scores, MBRS subscores were correlated with UPDRS scores for each task. Clinicians on average gave greater weight to amplitude than to speed or rhythm when scoring the finger-tapping task (Fig. 2A–C). There were, however, individual clinicians who placed greater emphasis on speed and/or rhythm for other tasks (not shown).

Quantitative Assessment

The average MBRS subscores for each task were correlated with quantitative variables extracted from kinematic data recorded on the motion sensor units. Figure 2D–F plots the quantitative variables versus the average MBRS subscores for the finger-tapping task. The log of RMS angular velocity was found to correlate best to speed scores, RMS excursion angle correlated best with amplitude scores, and coefficient of variation correlated best with rhythm scores (Supporting Information Table 3). Rhythm scores correlated poorly with other variables, including speed fatigue (mean $r = 0.36$), amplitude fatigue (mean $r = 0.38$), and arrests in movement (mean $r = 0.28$), all of which are reflected in the coefficient of variation.

Discussion

The MBRS demonstrated inter- and intrarater reliabilities similar to those of the UDPRS but affording greater discrimination for the 3 main components of movement impairment in PD: slowness, low amplitude, and dysrhythmia. Although training did not improve the

correlation coefficient between raters, it reduced RMS errors for speed and amplitude subscores, indicating a reduction in systematic biases across raters. That is, baselines and slopes (rate of increase in score with increase in symptom severity) were normalized across raters.

The MBRS may have several advantages over the UPDRS. The MBRS separately rates speed, amplitude, and rhythm, thus providing increased sensitivity in identifying different components of bradykinesia. On average, clinicians give greater weight to amplitude impairment when determining a UPDRS score, whereas speed is hardly considered. This may partially explain why interrater reliability was highest for amplitude scores and lowest for speed scores. Also, if speed is more difficult to rate than amplitude, interrater reliably suffers, and the clinician may be less likely to rely on it when assigning a UPDRS score. However, not all clinicians placed the highest weight on amplitude. Differential weighting of various components of movement impairment among raters when assigning a UPDRS score could dilute the power of finding true changes (type II error) when evaluating the "bradykinesia" outcome of clinical trials. For example, studies have shown that dopaminergic medications normalize speed to a greater extent than amplitude.[5] However, if clinicians are only concerned with amplitude when performing a UPDRS evaluation, improvements in speed may go unnoticed. Furthermore, in multicenter clinical trials, UPDRS evaluations are performed by multiple clinicians. Because UPDRS scores are often used as primary outcome measures, clinicians weighing various aspects of bradykinesia differently could greatly increase variability and confound results.

The MBRS also has some of the limitations that the UPDRS has because it relies on subjective clinical judgment. In contrast, quantitative features extracted from motion sensors yielded objective and reliable kinematics of hand movements, which were highly correlated with average clinician MBRS subscores. Although it produced higher correlation with the rhythm score than did variables representing fatigue or arrests in the movement, the coefficient of variation was less correlated with the rhythm subscore than were the speed and amplitude variables with their respective subscores. This lower correlation could be because hesitations, arrests in movement, and fatigue are all represented by the MBRS rhythm subscore. Overall, correlations between the quantitative features and average MBRS scores were similar to correlations achieved when comparing each clinician to the others, which is the best that could be expected because of variability between clinicians. In addition, our kinematic correlations to MBRS subscores were higher than others have achieved using quantitative digitography measurements of key-strike velocity ($r = 0.63$), time between key strikes ($r = 0.67$), and the coefficient of variation of duration of key strikes ($r = 0.67$) compared with the overall UPDRS-III bradykinesia sections.[20] Our results suggest that motion sensors can objectively measure speed, amplitude, and rhythm without reliability concerns associated with clinical rating scales, as was done previously for rating rest, postural, and kinetic tremor.[6]

This study demonstrated that bradykinesia manifestations of speed, amplitude, and rhythm can be evaluated independently using the MBRS and that objective features extracted from kinematic data are highly correlated with clinician scores. The primary utility of the MBRS may be in clinical trials, in which it could be important to determine which specific aspects of movement are responsive to the intervention in question. In addition, independent quantification of speed, amplitude, and rhythm could aid in the basic understanding of neurological pathways and specific drug mechanisms. Distinct neural mechanisms may underlie different motor manifestations of movement. Separately quantifying the subcomponents of movement impairment may enable the effect of novel therapies to be more accurately measured and possibly better targeted.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding agencies: This study was supported by the Davis Phinney Foundation for Parkinson's Disease.

The Davis Phinney Foundation for Parkinson's Disease provided unrestricted support and had no role in the oversight or review of the research data or reporting.

References

1. Berardelli A, Rothwell JC, Thompson PD, Hallett M. Pathophysiology of bradykinesia in Parkinson's disease. Brain. 2001; 124:2131. [PubMed: 11673316]
2. Dunnewold RJ, Jacobi CE, van Hilten JJ. Quantitative assessment of bradykinesia in patients with Parkinson's disease. J Neurosci Methods. 1997; 74:107–112. [PubMed: 9210580]
3. Poluha PC, Teulings HL, Brookshire RH. Handwriting and speech changes across the levodopa cycle in Parkinson's disease. Acta Psychol (Amst). 1998; 100:71–84. [PubMed: 9844557]
4. Evarts EV, Teravainen H, Calne DB. Reaction time in Parkinson's disease. Brain. 1981; 104:167–186. [PubMed: 7470841]
5. Espay AJ, Beaton DE, Morgante F, Gunraj CA, Lang AE, Chen R. Impairments of speed and amplitude of movement in Parkinson's disease: a pilot study. Mov Disord. 2009; 24:1001–1008. [PubMed: 19230031]
6. Giuffrida JP, Riley D, Maddux B, Heldman DA. Clinically deployable Kinesia technology for automated tremor assessment. Mov Disord. 2009; 24:723–730. [PubMed: 19133661]
7. Salarian A, Russmann H, Wider C, Burkhard PR, Vingerhoets FJ, Aminian K. Quantification of tremor and bradykinesia in Parkinson's disease using a novel ambulatory monitoring system. IEEE Trans Biomed Eng. 2007; 54:313–322. [PubMed: 17278588]
8. Boraud T, Tison F, Gross C. Quantification of motor slowness in Parkinson's disease: correlations between the tapping test and single joint ballistic movement parameters. Parkinsonism Relat Disord. 1997; 3:47–50. [PubMed: 18591054]
9. Martínez-Martín P, Gil-Nagel A, Gracia LM, Gómez JB, Martínez-Sarriés J, Bermejo F. Unified Parkinson's Disease Rating Scale characteristics and structure. The Cooperative Multicentric Group. Mov Disord. 1994; 9:76–83. [PubMed: 8139608]
10. Henderson L, Kennard C, Crawford TJ, et al. Scales for rating motor impairment in Parkinson's disease: studies of reliability and convergent validity. J Neurol Neurosurg Psychiatry. 1991; 54:18–24. [PubMed: 2010754]
11. Bennett DA, Shannon KM, Beckett LA, Goetz CG, Wilson RS. Metric properties of nurses' ratings of parkinsonian signs with a modified Unified Parkinson's Disease Rating Scale. Neurology. 1997; 49:1580–1587. [PubMed: 9409350]
12. Camicioli R, Grossmann SJ, Spencer PS, Hudnell K, Anger WK. Discriminating mild parkinsonism: methods for epidemiological research. Mov Disord. 2001; 16:33–40. [PubMed: 11215590]
13. Kirshner B, Guyatt G. A methodological framework for assessing health indices. J Chronic Dis. 1985; 38:27–36. [PubMed: 3972947]
14. Siderowf A, McDermott M, Kieburtz K, Blindauer K, Plumb S, Shoulson I. Test-retest reliability of the unified Parkinson's disease rating scale in patients with early Parkinson's disease: results from a multicenter clinical trial. Mov Disord. 2002; 17:758–763. [PubMed: 12210871]
15. Kishore A, Espay AJ, Marras C, et al. Unilateral versus bilateral tasks in early asymmetric Parkinson's disease: differential effects on bradykinesia. Mov Disord. 2007; 22:328–333. [PubMed: 17216641]
16. Dunnewold RJ, Hoff JI, van Pelt HC, Fredrikze PQ, Wagemans EA, van Hilten BJ. Ambulatory quantitative assessment of body position, bradykinesia, and hypokinesia in Parkinson's disease. J Clin Neurophysiol. 1998; 15:235–242. [PubMed: 9681561]

17. Kandori A, Yokoe M, Sakoda S, et al. Quantitative magnetic detection of finger movements in patients with Parkinson's disease. Neurosci Res. 2004; 49:253–260. [PubMed: 15140567]
18. Kandori A, Miyashita T, Hosono N, et al. Motion analysis of grip and release with fingers using simple magnetic detection system. Rev Sci Instrum. 2007; 78 034302–034306.
19. Baas H, Stecker K, Fischer PA. Value and appropriate use of rating scales and apparative measurements in quantification of disability in Parkinson's disease. J Neural Transm Park Dis Dement Sect. 1993; 5:45–61. [PubMed: 8439392]
20. Tavares ALT, Jefferis GS, Koop M, et al. Quantitative measurements of alternating finger tapping in Parkinson's disease correlate with UPDRS motor disability and reveal the improvement in fine motor control from medication and deep brain stimulation. Mov Disord. 2005; 20:1286–1298. [PubMed: 16001401]
21. Ward CD, Gibb WR. Research diagnostic criteria for Parkinson's disease. Adv Neurol. 1990; 53:245–249. [PubMed: 2239463]

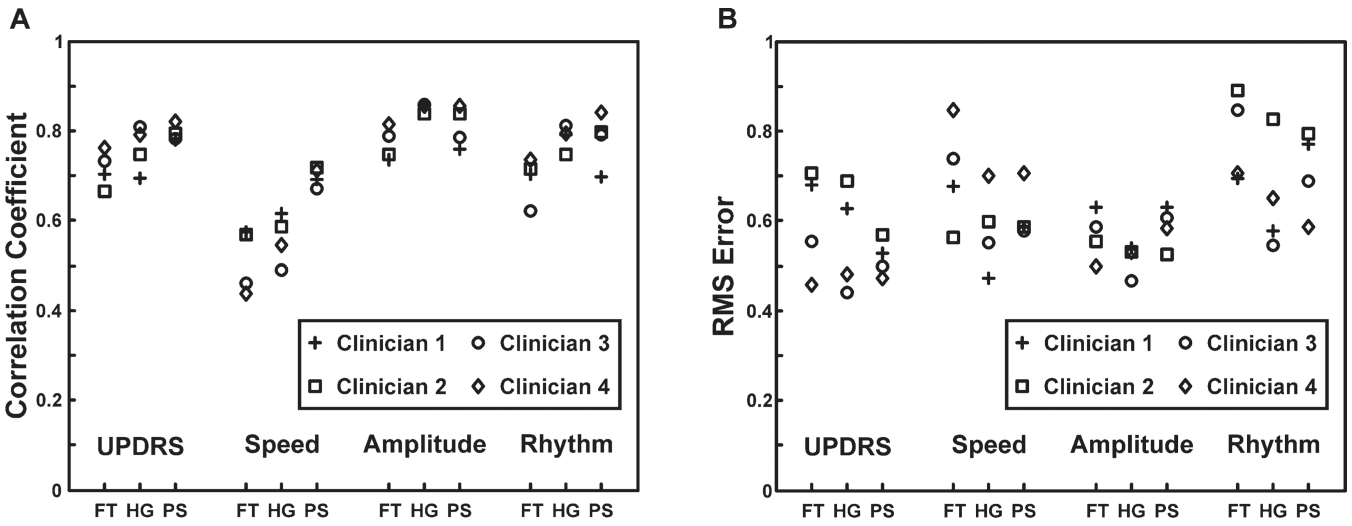

FIG. 1.

The correlation coefficients (**A**) and RMS errors (**B**) for each clinician compared with the average of the other 3 are shown for each score type during the finger-tapping (FT), hand-grasp (HG), and pronation–supination (PS) tasks.

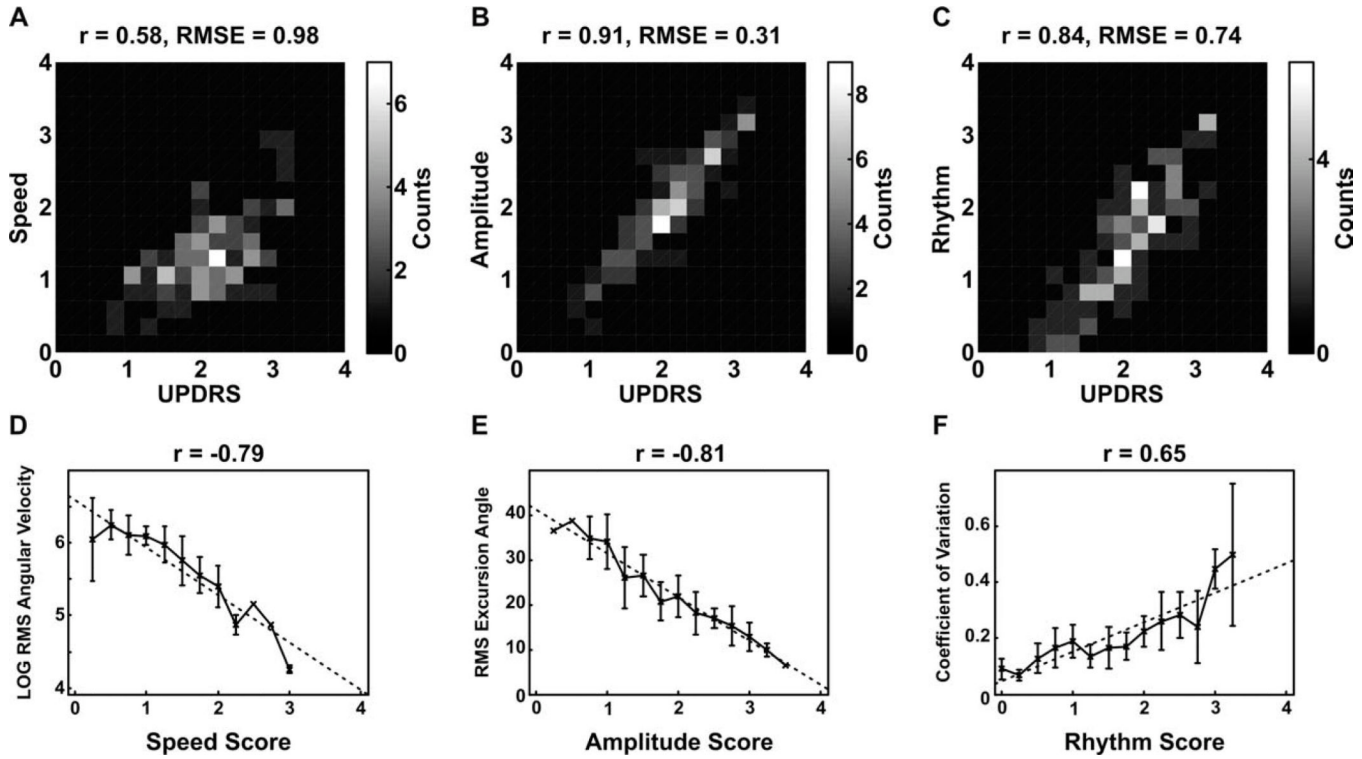

FIG. 2.

Grayscale histograms show counts of the average UPDRS compared with the average speed (**A**), amplitude (**B**), and rhythm (**C**) MBRS subscores for the finger-tapping task, along with their corresponding correlation coefficient (*r*) and root-mean-squared error (RMSE). On average, amplitude was weighted highest, followed by rhythm, then speed, when assigning a UPDRS score. Quantitative variables representing (**D**) speed (log RMS angular velocity), (**E**) amplitude (RMS excursion angle), and (**F**) rhythm (coefficient of variation) are plotted versus the average clinician MBRS subscores for the finger-tapping task. The dotted line is the least-squares fit, and the error bars equal 1 standard deviation.

TABLE 1

Inter- and intrarater reliability

	UPDRS		MBRS					
			Speed		Amplitude		Rhythm	
	r	RMSE	*r*	RMSE	*r*	RMSE	*r*	RMSE
Interrater reliability								
Finger taps	0.72	0.6	0.51	0.71	0.77	0.57	0.69	0.78
Hand grasps	0.76	0.56	0.56	0.58	0.85	0.52	0.79	0.65
Pronation–supination	0.79	0.52	0.7	0.61	0.81	0.59	0.78	0.71
Intrarater reliability								
Finger taps	0.75	0.55	0.61	0.66	0.79	0.56	0.72	0.82
Hand grasps	0.72	0.58	0.56	0.63	0.83	0.54	0.73	0.68
Pronation–supination	0.71	0.64	0.66	0.6	0.79	0.66	0.79	0.69

Each clinician's ratings were compared with the average ratings of the other 3. The averages of all 4 clinicians' correlation coefficients (*r*) and root-mean-squared errors (RMSE) are shown for the UPDRS and MBRS subscores for each task (top). Each clinician's ratings during the second rating session were compared with his ratings during the first rating session. The averages of all 4 clinicians' intrarater *r* and RMSE are shown for the UPDRS and MBRS subscores for each task (bottom).