Remote Monitoring of Treatment Response in Parkinson's Disease: The Habit of Typing on a Computer

Matarazzo Michele (Orcid ID: 0000-0002-2907-8667)
Martinez-Castrillo Juan Carlos (Orcid ID: 0000-0001-7744-6850)
Sanchez-Ferro Alvaro (Orcid ID: 0000-0003-2461-2485)

Manuscript type: Original Article

Remote Monitoring of Treatment Response in Parkinson´s disease: The Habit of Typing on a Computer

Michele Matarazzo, MD,[1,2,3,4] Teresa Arroyo-Gallego, MSc,[5,6,7] Paloma Montero, MD,[2,8] Verónica Puertas-Martín, PhD,[2] Ian Butterworth, MSc,[9] Carlos S. Mendoza, PhD,[9] María J. Ledesma-Carbayo, PhD,[5] María José Catalán, MD,[8] José Antonio Molina, MD,[2] Félix Bermejo-Pareja, MD,[2] Juan Carlos Martínez-Castrillo, MD,[10] Lydia López-Manzanares, MD,[11] Araceli Alonso-Cánovas, MD,[10] Jaime Herreros Rodríguez, MD,[12] Ignacio Obeso, PhD,[1,3] Pablo Martínez-Martín, MD,[3,13] José Carlos Martínez-Ávila, PhD,[14] Agustín Gómez de la Cámara, MD,[14] Martha Gray, PhD,[6,9] José A. Obeso, MD,[1,3] Luca Giancardo, PhD,[6,15*] Álvaro Sánchez-Ferro, MD,[1,2,3,6*]

[1]HM-CINAC, Hospital Universitario HM Puerta del Sur, Móstoles and Medical School, CEU-San Pablo University, Madrid, Spain

[2]Neurology Department, Instituto de Investigación del Hospital 12 de Octubre, Madrid, Spain

[3]Centro de Investigación Biomédica en Red, Enfermedades Neurodegenerativas (CIBERNED), Madrid, Spain

[4]Pacific Parkinson's Research Centre, University of British Columbia, Vancouver, BC, Canada

[5]Biomedical Image Technologies, Universidad Politécnica de Madrid and CIBERBBN Madrid, Spain

[6]Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, MA, USA

[7]nQ Medical Inc., Cambridge, MA, USA

[8]Movement disorders unit, Hospital Clínico San Carlos, Madrid, Spain

[9]Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA, USA

[10]Movement disorders unit, Hospital Ramón y Cajal, Madrid, Spain

[11]Movement disorders unit, Hospital de la Princesa, Madrid, Spain

[12]Neurology Department, Hospital Infanta Leonor, Madrid, Spain

[13]Area of Applied Epidemiology, National Centre of Epidemiology, Carlos III Institute of Health, Madrid, Spain

[14]Clinical Research Unit, Instituto de Investigación Hospital 12 de Octubre, Consorcio de Investigación Biomédica en Red de Epidemiología y Salud Pública (CIBERESP), Hospital Universitario 12 de Octubre, Madrid, Spain.

[15]Center for Precision Health, School of Biomedical Informatics, University of Texas Health Science Center at Houston, Houston, TX, USA

* These authors contributed equally

Correspondence to: Álvaro Sánchez-Ferro, MD, PhD, Centro Integral de Neurociencias A.C., Avenida del Emperador Carlos V, 70, E-28938 Madrid, Spain. Tel.: +34 912673200, fax: +34 912673197, e-mail: aferro.hmcinac@hmhospitales.com

Word count for the text: 3292

Word count for the abstract: 234

Running title: Remotely Monitoring Drug Response in Parkinson's Disease

References: 33

Figures: 3

Tables: 1

Keywords: Technology; Neural Network; Drug Monitoring; Parkinson's disease; Machine Learning.

Funding sources

This study was conducted with research funds from the Michael J. Fox Foundation for Parkinson's Research (grant number 10860).

ABSTRACT

Objective: The recent advances in technology are opening a new opportunity to remotely evaluate motor features in people with Parkinson's disease (PD). We hypothesized that typing on an electronic device, a habitual behavior facilitated by the nigrostriatal dopaminergic pathway, could allow for objectively and non-obtrusively monitoring parkinsonian features and response to medication in an at-home setting.

Methods: We enrolled thirty-one subjects recently diagnosed with PD, who were due to start dopaminergic treatment, and thirty age-matched controls. We remotely monitored their typing pattern over a six-month (24 weeks) follow-up period, before and while dopaminergic medications were being titrated. The typing data were used to develop a novel algorithm based on recursive neural networks ("nQRNN") and detect participants' response to medication. The latter was defined by the UPDRS-III minimal clinically important difference (MCID). Furthermore, we tested the accuracy of the algorithm to predict the final response to medication as early as 21 weeks prior to the final six-month clinical outcome.

Results: The nQRNN score had an overall moderate kappa agreement and fair area under the ROC curve with the time-coincident UPDRS-III MCID. The subjects classified as responders at the final visit (based on the UPDRS-III MCID) had higher nQRNN score compared to subjects with stable UPDRS-III, from the third week of the study onwards.

Conclusions: This preliminary study suggests that remotely-gathered unsupervised typing data allows for accurate detection and prediction of drug response in PD.

INTRODUCTION

Current evaluation standards in Parkinson's disease (PD), such as the Unified Parkinson's Disease Rating Scale (UPDRS) [1], are very useful, but have important limitations. As recently pointed out, these scales report a semi-quantitative and subjective score, non-sensitive to subtle motor changes [2–4]. In addition, these assessments typically require the patient to travel to the clinic and need to be performed by a trained specialist, representing an additional burden for the patient and hence being time- and cost-consuming.

For these reasons, several attempts are underway to complement traditional standards with more objective, quantitative and continuous outcome measures [5,6].

Notably, in the last decades we are witnessing an exponential adoption of smart technologies, such as computers, smartphones and tablets. This natural interaction with keyboards and touch screens is probably driven by habitual-directed movements, whose control is regulated by nigro-striatal activity [7–9]. Hence, we set to ascertain whether a natural interaction with keyboards would enable a new method to remotely detect and monitor parkinsonian motor signs non-obtrusively by analyzing the characteristics of free-text typing. Such approach could have advantages over existing solutions [10–12], because (i) it can extracts motor information from the natural interaction of the patients or study participants with their devices, without requiring active collaboration, (ii) it could virtually reach any person who is typing with an

internet-connected device, opening a window on the motor skills and parkinsonian signs of an enormous number of individuals, (iii) it can acquire data remotely, without requiring to attend a clinic and (iv) subjects could be monitored longitudinally in a quasi-continuous manner.

We have previously shown that data collected from an in-clinic typing task accurately differentiated early PD subjects from sex- and age-matched healthy controls [13], and replicated this results using at-home, unsupervised data [14]. Recently, we demonstrated similar performance with data acquired during typing on a touch-screen smartphone [15]. In the present study, our aim is to detect response to medication in PD by using remotely-gathered unsupervised typing data in an at home everyday life setting as an additional step to this new digital care model. Thus, we designed a prospective naturalistic study enrolling early PD subjects that were going to start dopaminergic medication, and followed them for six months. The specific goal of this study was. to validate a novel approach applied to remotely-gathered typing data for (i) detecting response to medication in an early PD population and (ii) predicting which PD subjects will respond to the drugs at the final visit based on the typing data collected at home up to 21 weeks in advance.

MATERIALS AND METHODS

Study Participants

Between March 2015 and June 2016, 31 consecutive early PD subjects were recruited from seven hospitals in Madrid (Spain) (figure 1), according to prespecified inclusion and exclusion criteria that are detailed in the Supplementary Materials. Thirty age- and sex-matched healthy controls (HC) were enrolled after ruling out the existence of a parkinsonism, hand deformities, cognitive impairment, sleep problems, or any other potential confounders (e.g., use of psychoactive medication, drug abuse or a serious medical condition).

The sample size was prespecified to detect at least fifteen subjects with response to medication, according to a previous definition of response (i.e. decrease of at least five points in the total UPDRS-III score) [16]. We a priori estimated a responder rate of 50% based on previous information from various randomized clinical trials[17–19] For this, we targeted subjects that were prescribed dopamine agonist or L-Dopa. We did not exclude participants already on rasagiline, but evaluated them off medication. We expected a neglectable confounding motor effect of this drug based on the main pivotal trials were the reported UPDRS change at 12-36 weeks was ≤ 0.11 points [20].

Besides the participants who met the exclusion criteria, two additional PD participants were excluded from the final analyses, in one case because typing data was insufficient to generate a score, and in the other case because the subject's laptop had an

operating system that was not compatible with our software (figure 1). Finally, five participants who got worse based on the definition of response described below were also not included in the model training and typing analyses.

Post-hoc evaluation of typing consistency was done to define the minimum amount of data needed to obtain a reliable score. For this reason, a typing day was defined as at least 10 valid windows per day, where a valid window was represented by a data sequence of at least 30 keystrokes within 90-second time interval. Then, we defined as "consistent typers" participants with at least one typing day in 80% of every possible 15 days rolling windows over the entire follow-up period (an overview of typing activity and consistency is available in the supplementary figure 1). The analyses were conducted in both consistent and non-consistent typers to evaluate the impact of typing frequency on the method's diagnostic performance.

Study design

A summary of the study design can be found in the Supplementary Materials, supplementary figure 2 and in clinicaltrials.gov (NCT02522065). Succinctly, all the subjects included in the study received a complete evaluation by a movement disorder specialist (MM, PME) at baseline that included the Unified Parkinson's Disease Rating Scale section III (UPDRS-III), the Purdue Pegboard test, and other standards.

At the baseline visit, the neuroQWERTY software was installed in the subjects' laptop. Subjects were invited to freely type for at least 20 minutes per day during the whole duration of the study. The software ran in the background of the laptop, capturing the

typing data – press/release timestamps of keystrokes –, that was automatically sent to a remote server located at MIT (Boston, USA). The privacy of the typing data was assured by encryption of keystroke information, anonymization of the subjects and authentication for accessing the data [14].

To obtain at-home (off) baseline typing data before the participants started the newly prescribed dopaminergic medications, PD subjects were instructed to delay the start of the new drug seven days after the initial baseline visit. Further follow-up visits were scheduled flexibly at week 4, 8, 16 and 24 after starting the medication with the same assessments that were conducted at baseline (for further information see Supplementary Materials).

Standard Protocol Approvals, and Patient Consents

All the experimental protocols were approved by the Massachusetts Institute of Technology, USA (no. 1412006804), "HM Puerta del Sur" University Hospital, Spain (no. 15.05.796-GHM), "12 de Octubre" University Hospital, Spain (no. CEIC:14/090) and "Clínico San Carlos" University Hospital, Spain (no.14/136-E). All subjects provided written informed consent prior to study enrollment.

Definition of drug response: UPDRS-III Minimal Clinically Important Difference

To classify participants as "improved", "not-changed" or "worsened", we calculated the Minimal Clinically Important Difference (MCID) of UPDRS-III for this study [21]. The relevant cut-off for our cohort was established in $\pm$ 5 points. There were only five

participants who worsened, and they have not been included in the results (see limitations in the discussion). In terms of response, we compared those subjects that did not change (UPDRS-III change ranging from minus 5 to plus 5 points) to those participants that improved (UPDRS-III scores that were lower by more than 5 points at follow up).

Classification Modelling: nQRNN

We used a machine learning model (nQRNN) which receives as input typing features derived from the hold time, i.e. the time required to press and release each key on the subject's laptop, regardless of the text typed. The typing features are encoded as "Key Hold Time Distribution" matrices [22] joined with the encoding previously described [13,14]. nQRNN outputs the probability of each patient of being a "responder" or "non-responder" and were employed to generate the plots in figures 2 and 3. nQRNN architecture is based on hierarchical layers of long short-term memory units (a type of recurrent neural network) trained using a nested cross-validation approach to avoid overfitting and a previously described optimization algorithm known as RMSprop. This type of software architecture is known to be an effective predictive model for complex time series data [23]. More details are available in the Supplementary Materials.

Data analyses

Two different types of analysis were performed: (i) the agreement of the nQRNN-based with the time-coincident UPDRS-III MCID-based classification of response and

the (ii) the prediction of whether the patients would be classified as responders or non-responder at the final visit (according to UPDRS-III MCID) based on the at-home nQRNN score obtained from previous weekly time-points throughout the study. The score was calculated from the third week after the medication was prescribed onwards, when there was sufficient data to conduct the described analyses. The analyses were replicated in both consistent typers and in the whole study cohort. More details on the analyses conducted are provided in the Supplementary Materials.

Evaluation of cognition as possible confounder

To rule out the possible confounding effect of cognition, we computed the Spearman correlation between Montreal Cognitive Assessment test (MoCA) score and nQRNN score [24,25]. Moreover, we analyzed the response classification of participants who fulfilled MDS criteria of PD mild cognitive impairment (PD-MCI) according to the results of a complete neuropsychological battery [26].

Data collection, database processing and statistical analysis

Software characteristics have been previously described [13]. Clinical data were collected using the Research Electronic Data Capture software (REDCap) [27]. The RNN was developed in Python [28] and Keras framework. Database processing and statistical analysis were performed using R 3.3.2 [29]. Baseline characteristics were compared using a nonparametric approach (Mann-Whitney *U* test for continuous variables and Fisher's exact test for categorical variables; multiple group comparisons

at baseline were done using Kruskal-Wallis test with post-hoc pairwise Mann-Whitney *U* test in the continuous variables). The comparisons between nQRNN medians of the "improved" and "not-changed" groups of the prediction analysis were performed using Mann-Whitney *U* test. Receiver Operating Characteristics (ROC) curve analysis was used for the agreement of at-home nQRNN score with UPDRS-III-based response classification, as well as for the prediction analysis of final response. Cohen's kappa was used as measure of agreement, and Cohen's effect size was calculated for agreement and prediction analyses. Significance was defined as a two-sided type I error below the 5% probability.

RESULTS

Comparability and characterization of the studied cohort

Baseline demographic characteristics were similar between the PD (N=29) and HC (N=30) groups, as shown in the table 1. As expected, statistically significant differences were observed in the motor performance of the PD and controls (e.g., UPDRS-III).

The Levodopa Equivalent Daily Dose (LED) was also statistically different between the 2 groups.

The median MoCA score was 28 (IQR: 27-29) in the HC group and 27 (IQR: 26-28) in the PD group, being this 1-point difference statistically significant (p= 0.049).

Based on the participants' final response to medication (HC, PD who improved and PD who did not change), further comparisons were done between them at baseline,

(supplementary table 1). Overall, statistically significant differences were observed only in motor performance-dependent variables. Pairwise comparisons found that those differences were due to difference between PD participants and HC, as expected. The only statistically significant difference between the two PD groups was in the Purdue Assembly task ($p = 0.01$).

The median disease duration at recruitment of the PD group was 13.9 months (IQR: 10.4-32.4). Most of the PD participants were in stage 2 of HY scale (n= 19/29, 65.5%). The rest were in stage 1 (n= 7/29, 24.1%) or 2.5 (n= 3/29, 10.3%); none of them were in stage 4 or 5. The tremor-dominant phenotype of PD was the most frequent (n= 15/29, 51.7%) [30]. Six PD participants (20.7%) were already receiving rasagiline at recruitment.

Ninety-five percent of the sample had a six month follow up. Two of the PD participants and one HC dropped out before the 24-week visit for personal reasons, not related to the study. During the 6-month follow-up, there was a progressive increase in the median LED of the PD group from 0 at baseline to 340 (IQR= 220-400) at week 24 (supplementary figure 3).

The number of PD participants who responded to medication, according to the 5-point UPDRS-III MCID, increased during the study as the medication was titrated (supplementary figure 4). At the final visit (i.e. week 24) the PD subjects who responded to medication were 51.9%. Of the whole sample 37 (66.1%) subjects were "consistent typers", including 20 (69.0%) controls and 17 (63.0%) PD patients.

Evaluating the concurrent and Discriminant Validity of nQRNN

There was a moderate significant correlation (Spearman Rho 0.33. P = 0.02) between nQRNN and the time coincident UPDRS-III delta at the final endpoint visit. The correlation with different non-motor measures such as MoCA was non-significant (Spearman Rho = 0.10. P 0.49).

Analyzing the agreement and accuracy of nQRNN to detect drug response

The Area Under the ROC Curve (AUC) for all the aggregated data (i.e. all time-points) to classify the subjects as "improved" or "not changed" using the nQRNN was 0.77 (95% Confidence Interval [CI]: 0.68-0.87) in consistent typers, and 0.75 (95% CI: 0.67-0.84) for the whole sample (figure 2). Kappa agreement was moderate for both consistent typers and the whole cohort (0.55 and 0.47, respectively). A large Cohen's d effect size was observed (1.26 for the consistent typers and 0.92 for the whole sample). The overall balanced accuracy was 76.5% in the whole sample and 77.4% in the consistent typers. Supplementary table 2 shows the results for each time-point and the results of all available data aggregated in the whole sample and in consistent typers.

Predicting response to medication with nQRNN score at home

Considering only consistent typers, the nQRNN achieved a good prediction of the final classification since week 3 after the treatment was started, showing stable median scores in both groups from week 7 onwards (figure 3). The nQRNN score was higher in

responders compared to non responders for every week analyzed (i.e. from week 3 to week 24), with p values < 0.005 for consistent typers (supplementary table 3). When adjusted for multiple comparisons with the Bonferroni correction, the weeks 20-24 did not reach statistical significance.

The longitudinal ROC curve analysis for predicting response to medication showed AUCs > 0.80 for the entire period analyzed (from week 3 to week 24) in consistent typers, while the AUCs considering the whole cohort oscillated between 0.69 and 0.75 (0.73-0.75 after the 6^{th} week of treatment). The nQRNN threshold, calculated on a weekly basis using the Youden's method, was stable from the week 7 onward (figure 3). Supplementary data, including AUCs and Cohen's d effect sizes are available in supplementary table 4.

The AUC of other baseline characteristics (age, computer use, UPDRS-III and PDQ-39) were not statistically significant for the prediction of the final classification, confirming that the results of the nQRNN are not due to baseline group differences (supplementary table 5).

Evaluation of cognition as confounding

Statistically significant difference was observed between PD participants and HC in the MoCA score, but such difference was not noted between the three groups analyzed (PD subjects who improved, PD subject who did not change and HC). Spearman ρ between MoCA and nQRNN score showed non-significant correlation (see results above). Moreover, PD subjects classified as PD-MCI were evenly distributed between

the three groups (1 in the “improved” group, 2 in “not changed” and 2 in the “worsened”).

DISCUSSION

Medicine and neurology are moving towards a new model of care, based on objective data collected remotely (i.e. ecologically valid) and non-intrusively (i.e. not requiring the active cooperation of the patients) [5,6,10,31,32]. This approach will allow doctors or drug makers to take informed decisions on PD diagnosis or therapy remotely.

In this new scenario, we investigated the preliminary validity of free unconstrained typing at home as a proxy of drug response in PD. In a longitudinal prospective naturalistic study with a lengthy follow up, we have shown that a recurrent neural network algorithm accurately detected (i.e. AUC 0.75) the response to dopaminergic therapy in an early PD population, with a moderate kappa agreement and large Cohen’s d effect size, compared to time-coincident in-clinic UPDRS-III classification. Further, we showed that remote monitoring of motor signs of PD using non-intrusive, free typing information is feasible and with good compliance, considering that only two subjects (3.3%) were excluded because of insufficient data.

The possibility of remotely monitoring the response to medication and the motor status of the PD subjects can be a major step for improving the management of the disease in clinical practice and take decisions on further changes of treatment or on the planning of future follow-up visits. Moreover, our score predicted from the third

week after starting the drug which PD participants responded to medication at the final visit. The classification became stable at week seven with a nQRNN threshold of 0.28 that remained the same until the study completion. Therefore, we were able to anticipate the clinical response to medication as early as the 21 weeks in advance, using uniquely remotely-gathered typing data. These findings suggest that our tool may be sensitive to subtle motor changes, being able to detect people that are responding to medication at an earlier stage and using remote, objective data. This could be crucial for example in supporting go/no-go decisions in early intervention trials reducing the cost of developing new compounds and also potentially being helpful to adjust treatments in clinical practice.

Our study has some limitations that should be considered. First, even though we used a nested cross-validation approach that allowed us to test the generalizability of our model in a limited dataset, our cohort does not provide a complete representation of all nuances of PD progression, cognitive states, coexisting conditions and typing habits. However, a machine learning model such as nQRNN is able to learn from new examples by design [33]. Therefore, there is potential for fine tuning the nQRNN performance by increasing the dataset or even adding other data modalities (e.g. touch screen or mouse clicks, among other possibilities). Secondly, due to the size of our dataset, we could not train the model to predict three distinct types of progressions and we focused on “improved” and “not changed”. We are confident that future studies, including a higher number of subjects with heterogeneous types of PD

progressions, could overcome this issue that is critical for the translation of our approach to everyday practice and clinical trials. In our model, the "not changed" group included the HC. HC are indeed a good sample of stable motor status (as it has been confirmed by a UPDRS-III change that was always below UPDRS-III MCID), however in a drug trial or in clinical practice HC do not necessarily need to be included.

As expected, the accuracy of agreement and prediction were higher in the consistent typers sub-group. Consistent typers are more likely to produce typing data evenly during the disease progression, leading to a more accurate prediction. However, the results obtained including non-consistent typers are still significant, which was an unexpected result in light of the limited amount of typing data available for the non-consistent typers sub-group. This is particularly important as, currently, the age group of people affected by Parkinson's disease might be less active users of technology. Finally, subjects with other medical conditions, such as hand deformities or other neurological issues, were excluded from our study. The impact of these possible confounders on our outcome score, still needs to be assessed.

In conclusion, we report on a pilot study on a novel technological approach to monitor motor features of PD and drug response remotely and ecologically in an accurate way, reflecting the underlying effects of basal ganglia neurodegeneration on a habitual task, such as typing. We show that this approach is feasible and suggest that it could be useful in the everyday clinical practice and could complement the current standard outcomes for improving the efficacy of clinical trials in PD, helping to reduce the

burden for participants and investigators and to assess in a more time- and cost-efficient way the response to medication.

AUTHORS' ROLE

Name	Role	Contribution
Michele Matarazzo, MD	First Author	Conceptualized, organized and executed the research project. Participated in the design and execution of the statistical analyses and wrote the first draft of the manuscript.
Teresa Arroyo-Gallego, MSc	Author	Conceptualized, organized and executed the research project. Participated in the design and execution of the statistical analyses and wrote the first draft of the manuscript.
Paloma Montero, MD	Author	Organized and executed the research project. Participated in the design of the statistical analyses. Reviewed the manuscript for intellectual content
Verónica Puertas-Martín, PhD	Author	Organized and executed the research project. Participated in the design of the statistical analyses. Reviewed the manuscript for

		intellectual content
Ian Butterworth, MSc	Author	Conceptualized the research project. Participated in the design of the statistical analyses. Reviewed the manuscript for intellectual content
Carlos S. Mendoza, PhD	Author	Conceptualized the research project. Participated in the design of the statistical analyses. Reviewed the manuscript for intellectual content
María J. Ledesma-Carbayo, PhD	Author	Conceptualized the research project. Participated in the design and execution of the statistical analyses. Reviewed the manuscript for intellectual content
María José Catalán, MD, PhD	Author	Conceptualized and executed the research project. Participated in the design and execution of the statistical analyses. Reviewed the manuscript for intellectual content

José Antonio Molina, MD, PhD	Author	Conceptualized the research project. Participated in the design of the statistical analyses. Reviewed the manuscript for intellectual content
Félix Bermejo-Pareja	Author	Conceptualized the research project. Participated in the design of the statistical analyses. Reviewed the manuscript for intellectual content
Juan Carlos Martínez-Castrillo	Author	Executed the research project. Reviewed the manuscript for intellectual content
Lydia López-Manzanares	Author	Executed the research project. Reviewed the manuscript for intellectual content
Araceli Alonso-Cánovas	Author	Executed the research project. Reviewed the manuscript for intellectual content
Jaime Herreros Rodríguez	Author	Executed the research project. Reviewed the manuscript for

		intellectual content
Ignacio Obeso	Author	Conceptualized and executed the research project. Reviewed the manuscript for intellectual content
Pablo Martínez-Martín	Author	Conceptualized the research project. Participated in the design and execution of the statistical analyses. Reviewed the manuscript for intellectual content
José Carlos Martínez-Ávila	Author	Conceptualized the research project. Participated in the design and execution of the statistical analyses. Reviewed the manuscript for intellectual content
Agustín Gómez de la Cámara	Author	Conceptualized the research project. Participated in the design and execution of the statistical analyses. Reviewed the manuscript for intellectual content
Martha Gray	Author	Conceptualized and organized the research project. Participated in

		the design and execution of the statistical analyses. Reviewed the manuscript for intellectual content.
José A. Obeso	Author	Conceptualized and organized the research project. Participated in the design and execution of the statistical analyses. Reviewed the manuscript for intellectual content.
Luca Giancardo	Co-corresponding Author	Conceptualized, organized and executed the research project. Participated in the design and execution of the statistical analyses and wrote the first draft of the manuscript.
Álvaro Sánchez-Ferro	Corresponding Author	Conceptualized, organized and executed the research project. Participated in the design and execution of the statistical analyses and wrote the first draft of the manuscript.

FINANCIAL DISCLOSURES OF ALL AUTHORS

Michele Matarazzo is currently supported by the Parkinson Canada Clinical Research Fellowship; no other financial disclosures. Teresa Arroyo-Gallego is currently employed by nQ Medical Inc., a company that has licensed the neuroQWERTY technology. Carlos Sánchez-Mendoza, Ian Butterworth, Luca Giancardo and Álvaro Sánchez-Ferro are co-inventors of a Method and Apparatus for Motor function Characterization (PCT/US2015/022581). Álvaro Sánchez-Ferro has received funding from "Consejería de Educación, Juventud y Deporte of Comunidad de Madrid" and the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA Grant 291820 and has also received speaker and travel honoraria from Teva, Zambon, Abbvie and Novartis pharmaceutical. All other authors have nothing to disclose.

REFERENCES

1 Fahn S, Elton R. Unified Parkinson's Disease Rating Scale (UPDRS). In: Fahn S, Marsden CD, Calne DB, *et al.*, eds. *Recent Developments in Parkinson's Disease, Vol 2.* Florham Park, NJ: 1987. 153–63.

2 Kieburtz K, Olanow CW. Advances in clinical trials for movement disorders. *Mov Disord* 2015;**30**:1580–7. doi:10.1002/mds.26371

3 Schneider SA, Drude L, Kasten M, *et al.* A study of subtle motor signs in early Parkinson's disease. *Mov Disord* 2012;**27**:1563–6. doi:10.1002/mds.25161

4 Maetzler W, Domingos J, Srulijes K, *et al.* Quantitative wearable sensors for objective assessment of Parkinson's disease. *Mov Disord* 2013;**28**:1628–37. doi:10.1002/mds.25628

5 Sánchez-Ferro Á, Elshehabi M, Godinho C, *et al.* New methods for the assessment of Parkinson's disease (2005 to 2015): A systematic review. *Mov Disord* 2016;**31**:1283–92. doi:10.1002/mds.26723

6 Ferreira JJ, Godinho C, Santos AT, *et al.* Quantitative home-based assessment of Parkinson's symptoms: the SENSE-PARK feasibility and usability study. *BMC Neurol* 2015;**15**:89. doi:10.1186/s12883-015-0343-z

7 Redgrave P, Rodriguez M, Smith Y, *et al.* Goal-directed and habitual control in the basal ganglia: implications for Parkinson's disease. *Nat Rev Neurosci* 2010;**11**:760–72. doi:10.1038/nrn2915

8 Catalan MJ, Ishii K, Honda M, *et al.* A PET study of sequential finger movements of varying length in patients with Parkinson's disease. *Brain* 1999;**122 (Pt 3**:483–95. doi:10.1093/brain/122.3.483

9 Lehéricy S, Bardinet E, Tremblay L, *et al.* Motor control in basal ganglia circuits using fMRI and brain atlas approaches. *Cereb Cortex* 2006;**16**:149–61. doi:10.1093/cercor/bhi089

10 Giovannoni G, van Schalkwyk J, Fritz VU, *et al.* Bradykinesia akinesia inco-ordination test (BRAIN TEST): an objective computerised assessment of upper limb motor function. *J Neurol Neurosurg Psychiatry* 1999;**67**:624–9. doi:10.1136/jnnp.67.5.624

11 Hasan H, Athauda DS, Foltynie T, *et al.* Technologies Assessing Limb Bradykinesia in Parkinson's Disease. *J Parkinsons Dis* 2017;**7**:65–77. doi:10.3233/JPD-160878

12 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**:2129–70. doi:10.1002/mds.22340

13 Giancardo L, Sánchez-Ferro Á, Arroyo-Gallego T, *et al.* Computer keyboard interaction as an indicator of early Parkinson's disease. *Sci Rep* 2016;**6**:34468. doi:10.1038/srep34468

14 Arroyo-Gallego T, Ledesma-Carbayo MJ, Butterworth I, *et al.* Detecting Motor

Impairment in Early Parkinson's Disease via Natural Typing Interaction With Keyboards: Validation of the neuroQWERTY Approach in an Uncontrolled At-Home Setting. *J Med Internet Res* 2018;**20**:e89. doi:10.2196/jmir.9462

15 Arroyo-Gallego T, Ledesma-Carbayo MJ, Sanchez-Ferro A, *et al.* Detection of Motor Impairment in Parkinson's Disease Via Mobile Touchscreen Typing. *IEEE Trans Biomed Eng* 2017;**64**:1994–2002. doi:10.1109/TBME.2017.2664802

16 Schrag A, Sampaio C, Counsell N, *et al.* Minimal clinically important change on the unified Parkinson's disease rating scale. *Mov Disord* 2006;**21**:1200–7. doi:10.1002/mds.20914

17 Jankovic J, Watts RL, Martin W, *et al.* Transdermal rotigotine: double-blind, placebo-controlled trial in Parkinson disease. *Arch Neurol* 2007;**64**:676–82. doi:10.1001/archneur.64.5.676

18 Poewe W, Rascol O, Barone P, *et al.* Extended-release pramipexole in early Parkinson disease: a 33-week randomized controlled trial. *Neurology* 2011;**77**:759–66. doi:10.1212/WNL.0b013e31822affb0

19 Kieburtz K, Parkinson Study Group PramiBID Investigators. Twice-daily, low-dose pramipexole in early Parkinson's disease: a randomized, placebo-controlled trial. *Mov Disord* 2011;**26**:37–44. doi:10.1002/mds.23396

20 Olanow CW, Rascol O, Hauser RA, *et al.* A double-blind, delayed-start trial of rasagiline in Parkinson's disease. *N Engl J Med* 2009;**361**:1268–78. doi:10.1056/NEJMoa0809335

21 Sánchez-Ferro Á, Matarazzo M, Martínez-Martín P, *et al.* Minimal Clinically Important Difference for UPDRS-III in Daily Practice. *Mov Disord Clin Pract* Published Online First: 24 April 2018. doi:10.1002/mdc3.12632

22 Giancardo L, Sánchez-Ferro A, Butterworth I, *et al.* Psychomotor Impairment Detection via Finger Interactions with a Computer Keyboard During Natural Typing. *Sci Rep* 2015;**5**:1–8. doi:10.1038/srep09678

23 LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;**521**:436–44. doi:10.1038/nature14539

24 Wissler C. THE SPEARMAN CORRELATION FORMULA. *Science* 1905;**22**:309–11. doi:10.1126/science.22.558.309

25 Spearman C. The proof and measurement of association between two things. By C. Spearman, 1904. *Am J Psychol* 2012;**100**:441–71.

26 Litvan I, Goldman JG, Tröster AI, *et al.* Diagnostic criteria for mild cognitive impairment in Parkinson's disease: Movement Disorder Society Task Force guidelines. *Mov Disord* 2012;**27**:349–56. doi:10.1002/mds.24893

27 Harris PA, Taylor R, Thielke R, *et al.* Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;**42**:377–81. doi:10.1016/j.jbi.2008.08.010

28 van Rossum G, de Boer J. Interactively Testing Remote Servers Using the Python Programming Language. *CWI Q* 1991;**4**:283–303.

29 R Core Team. R: A Language and Environment for Statistical Computing. 2016.

30 Stebbins GT, Goetz CG, Burn DJ, *et al.* How to identify tremor dominant and postural instability/gait difficulty groups with the movement disorder society unified Parkinson's disease rating scale: comparison with the unified Parkinson's disease rating scale. *Mov Disord* 2013;**28**:668–70. doi:10.1002/mds.25383

31 Taylor Tavares AL, Jefferis GSXE, 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**:1286–98. doi:10.1002/mds.20556

32 Noyce AJ, Nagy A, Acharya S, *et al.* Bradykinesia-akinesia incoordination test: validating an online keyboard test of upper limb function. *PLoS One* 2014;**9**:e96260. doi:10.1371/journal.pone.0096260

33 Bishop CM. *Pattern Recognition and Machine Learning*. 1st ed. New York, NY: : Springer 2006.

Table 1. Demographic and baseline characteristics

	Group		p-value
	Healthy Controls (n = 30)	PD subjects (n = 29)	
Age	63.00 [56.48, 69.44]	59.78 [54.19, 68.60]	0.476
Sex (woman)	16 (53.3)	14 (48.3)	0.797
Handedness (right)	28 (93.3)	29 (100.0)	0.492
Alcohol	8 (26.7)	4 (13.8)	0.333
Tobacco	3 (10.0)	3 (10.3)	1.000
Hypertension	10 (33.3)	9 (31.0)	1.000
Diabetes Mellitus	5 (16.7)	2 (6.9)	0.424
Dyslipemia	8 (26.7)	7 (24.1)	1.000
Computer Use (years)	20.00 [10.00, 25.00]	20.00 [12.00, 20.00]	0.982
Weekly Computer Use (days)	7.00 [5.00, 7.00]	7.00 [4.00, 7.00]	0.402
Education (years)	15.00 [12.00, 18.00]	18.00 [12.00, 20.00]	0.322
MoCA	28.00 [27.00, 29.00]	27.00 [26.00, 28.00]	0.049*
LED	0.00 [0.00, 0.00]	0.00 [0.00, 0.00]	0.009*
UPDRS-III	1.00 [0.00, 2.00]	19.00 [17.00, 26.00]	<0.001*
Purdue Right	14.83 [14.00, 16.00]	12.33 [11.33, 15.33]	0.003*
Purdue Left	13.50 [12.67, 15.25]	11.33 [10.67, 12.67]	<0.001*
Purdue Both	11.17 [10.67, 12.67]	9.00 [8.00, 11.00]	<0.001*
Purdue Assembly	27.67 [23.25, 30.00]	24.67 [17.67, 29.67]	0.075

MoCA, Montreal Cognitive Assessment test; LED, Levodopa Equivalent Dose; UPDRS-III, Unified Parkinson's Disease Rating Scale part III.

Quantitative variables are represented as "median [IQR]" and qualitative ones as "n (%)".

*Statistically significant difference.

FIGURE TITLES AND LEGENDS

Figure 1. Flowchart of the study sample

Figure 2. Receiver Operating Characteristic Curves of nQRNN for UPDRS-III MCID-based classification of responders

The plot shows the ROC curves of the nQRNN for the binary classification of the subjects as "improved" (i.e. responders) and "not changed" according to the Minimal Clinically Important Difference of the UPDRS-III. The blue line is obtained plotting the whole sample of the study. The red line is obtained plotting only the subjects classified as "consistent typers". The shaded areas represent the 95% confidence interval.

Figure 3. Longitudinal nQRNN change in consistent typers

The thick lines and shaded areas represent the longitudinal median nQRNN scores and their interquartile ranges for the Parkinson's disease (PD) subjects that were finally classified as "improved" (green line) and for the PD and control subjects that were classified as "not changed" (yellow line). The thin lines represent the single subjects and their score over time. The grey line represents the best possible threshold according to the Youden method, computed for each week-interval, which maintained a stable value of 0.28 since the week 7 after starting the medication.

SUPPLEMENTARY FIGURES

Supplementary figure 1. Activity map and typing consistency definition

The figure shows the active typing days of the participants, and the classification in "consistent" and "non-consistent" typers. Each colored tile represent an active typing day of a single study participant.

Supplementary figure 2. Study design

The figure shows the design of the study. Participants were evaluated at baseline, when also our software was installed in their personal computers. The typing information was collected during the rest of the study (i.e. 24 weeks) from the at-home setting. PD participants started the dopaminergic drug 7 days after the baseline visit. Further in-clinic visits were flexibly scheduled at week 4, 8, 16 and 24.

Supplementary figure 3. Levodopa Equivalent Dose (LED) change over time

The blue line and the grey area represent the locally weighted non-parametric regression (LOESS) curve with 95% confidence interval of the Levodopa Equivalent Daily Dose (LED) of the PD subjects during the study. Further boxplots are also plotted with the LED at the time of each in-clinic visit. There was a progressive increase of the median LED, that was 0 at baseline (IQR= 0-0), 150 at week 4 (IQR= 100-267), 240 at week 8 (IQR= 125-320), 280 at week 16 (IQR= 150-340) and 340 at week 24 (IQR= 220-400).

Supplementary figure 4. Percentage of responders to medication over time

The blue area represents the percentage of PD participants that were classified as responders at each visit, according to the Minimal Clinically Important Difference for UPDRS-III. There was a progressive increase of responders throughout the study, until the final percentage of 51.7% fulfilling the criteria of response to medication at the last visit, at week 24 (i.e. change of UPDRS-III from baseline ≤ - 5 points).

Supplementary table 1. Demographic and baseline characteristics of the groups categorized MCID classification

	Group			p-value			
					Pairwise comparisons		
	PD Not changed (n = 8)	PD Improved (n = 14)	Controls (n = 29)	Overall	Not changed vs Improved	Improved vs Controls	Not changed vs Controls
Age	64.20 [57.06, 77.16]	58.32 [51.75, 62.08]	63.00 [56.52, 69.65]	0.212	0.219	0.097	0.658
Sex (woman)	5 (62.5)	7 (50.0)	15 (51.7)	0.858	0.675	1.000	0.701
Handedness (right)	8 (100.0)	14 (100.0)	27 (93.1)	1.000	1.000	1.000	1.000
Alcohol	0 (0.0)	3 (21.4)	8 (27.6)	0.327	0.273	1.000	0.160
Tobacco	0 (0.0)	3 (21.4)	3 (10.3)	0.413	0.273	0.373	1.000
Hypertension	3 (37.5)	4 (28.6)	10 (34.5)	1.000	1.000	1.000	1.000
Diabetes Mellitus	1 (12.5)	0 (0.0)	5 (17.2)	0.216	0.364	0.156	1.000
Dyslipemia	2 (25.0)	3 (21.4)	8 (27.6)	1.000	1.000	1.000	1.000
Computer Use (years)	16.00 [7.00, 20.00]	20.00 [15.00, 25.00]	20.00 [10.00, 25.00]	0.291	0.104	0.353	0.371
Weekly Computer Use (days)	7.00 [6.00, 7.00]	6.00 [2.25, 7.00]	7.00 [5.00, 7.00]	0.288	0.218	0.165	0.739
Education (years)	16.50 [12.75, 18.50]	16.50 [11.00, 20.00]	15.00 [12.00, 18.00]	0.695	0.973	0.482	0.515
MoCA	28.00 [27.00, 28.00]	27.50 [26.25, 29.00]	28.00 [27.00, 29.00]	0.583	0.944	0.405	0.406
LED	0.00 [0.00, 0.00]	0.00 [0.00, 0.00]	0.00 [0.00, 0.00]	0.046*	0.610	0.011*	0.057
UPDRS-III	20.00 [16.25, 25.25]	24.00 [16.50, 28.00]	1.00 [0.00, 2.00]	<0.001*	0.321	<0.001*	<0.001*
Purdue Right	11.83 [11.08, 13.25]	14.00 [12.58, 15.25]	14.67 [14.00, 16.00]	0.028*	0.218	0.122	0.013*
Purdue Left	11.00 [10.67, 11.33]	12.00 [10.08, 13.92]	13.33 [12.67, 15.00]	0.001*	0.411	0.018*†	<0.001*
Purdue Both	8.33 [7.92, 9.33]	9.67 [8.42, 11.25]	11.00 [10.67, 12.67]	0.001*	0.205	0.015*	0.001*

Purdue Assembly	18.17 [17.58, 19.83]	28.67 [25.00, 31.00]	27.67 [23.00, 30.00]	0.004*	0.009*	0.604	0.001*

MoCA, Montreal Cognitive Assessment test; LED, Levodopa Equivalent Daily Dose; UPDRS-III, Unified Parkinson's Disease Rating Scale part III. Quantitative variables are represented as "median [IQR]" and qualitative ones as "n (%)". Participants who got worse (N=5) were not included in these comparisons.
*Statistically significant difference.
†Loses statistical significance after Bonferroni adjustment for multiple comparisons.

Supplementary table 2. Longitudinal agreement analysis of nQRNN compared with UPDRS-III-based "improvement" vs. "no-change" classification

Week	AUC (95% CI)		Kappa* (95% CI)		Percentage agreement		Effect size† (95% CI)	
	Whole sample	Consistent typers	Whole sample	Consistent typers	Whole sample	Consistent typers	Whole sample	Consistent typers
4	0. 82 (0.71-0.93)	0.92 (0.82-1.00)	0.44 (0.17-0.72)	0.72 (0.43-1.00)	81.63%	91.18%	1.05 (-0.09-2.03)	2.04 (0.93-3.15)
8	0.83 (0.68-0.97)	0.77 (0.56-0.99)	0.47 (0.19-0.74)	0.42 (0.10-0.75)	80.39%	78.38%	1.31 (0.53-2.10)	1.17 (0.27-2.07)
16	0.72 (0.54-0.89)	0.68 (0.45-0.90)	0.44 (0.17-0.72)	0.47 (0.17-0.77)	77.78%	76.47%	0.83 (-0.02-1.51)	0.86 (0.04-1.68)
24	0.74 (0.58-0.89)	0.80 (0.65-0.96)	0.49 (0.24-0.75)	0.58 (0.31-0.85)	78.43%	81.08%	0.83 (0.18-1.48)	1.27 (0.48-2.06)
Aggregated	0.75 (0.67-0.84)	0.77 (0.68-0.87)	0.47 (0.33-0.60)	0.55 (0.40-0.70)	79.59%	81.69%	0.92 (0.55-1.30)	1.26 (0.83-1.68)

AUC, Area Under the Receiver Operating Curve; UPDRS-III, Unified Parkinson's Disease Rating Scale Part III.

*Cohen's Kappa value; †Cohen's d effect size.

Supplementary table 3. nQRNN score difference in consistent typers

Week	Group		p-value
	Not changed (n = 26)	Improved (n = 11)	
3	0.28 [0.27, 0.30]	0.34 [0.32, 0.36]	0.002*
4	0.27 [0.25, 0.28]	0.36 [0.32, 0.37]	0.001*
5	0.26 [0.24, 0.28]	0.36 [0.31, 0.39]	0.001*
6	0.26 [0.24, 0.27]	0.37 [0.32, 0.39]	0.001*
7	0.27 [0.24, 0.27]	0.37 [0.33, 0.39]	0.001*
8	0.27 [0.24, 0.27]	0.37 [0.33, 0.39]	0.001*
9	0.27 [0.24, 0.27]	0.37 [0.33, 0.39]	0.001*
10	0.26 [0.24, 0.27]	0.37 [0.33, 0.39]	0.001*
11	0.26 [0.24, 0.27]	0.37 [0.33, 0.39]	0.001*
12	0.26 [0.24, 0.27]	0.37 [0.33, 0.39]	0.001*
13	0.26 [0.24, 0.27]	0.37 [0.33, 0.39]	0.001*
14	0.26 [0.24, 0.27]	0.37 [0.33, 0.39]	0.001*
15	0.26 [0.24, 0.27]	0.37 [0.33, 0.39]	0.002*
16	0.27 [0.24, 0.27]	0.37 [0.33, 0.39]	0.002*
17	0.27 [0.24, 0.27]	0.37 [0.33, 0.39]	0.002*
18	0.27 [0.24, 0.27]	0.37 [0.33, 0.39]	0.002*
19	0.27 [0.24, 0.27]	0.37 [0.33, 0.39]	0.002*
20	0.27 [0.24, 0.27]	0.37 [0.33, 0.39]	0.003
21	0.27 [0.24, 0.27]	0.37 [0.33, 0.39]	0.003
22	0.27 [0.24, 0.27]	0.37 [0.33, 0.39]	0.004
23	0.27 [0.24, 0.27]	0.37 [0.33, 0.39]	0.004
24	0.27 [0.24, 0.27]	0.37 [0.33, 0.39]	0.004

Score is represented as “median [IQR]"

*Statistically significant difference after Bonferroni correction for multiple comparisons

Supplementary table 4. Longitudinal analysis of nQRNN accuracy predicting UPDRS-III-based "improvement" vs. "no-change" classification at last visit

Week	AUC (95% CI)		Effect size* (95% CI)	
	Whole sample	Consistent typers	Whole sample	Consistent typers
3	0.72 (0.55-0.88)	0.84 (0.66-1.00)	0.50 (-0.16-1.15)	1.48 (0.65-2.32)
4	0.71 (0.53-0.88)	0.85 (0.69-1.00)	0.67 (0.02-1.31)	1.50 (0.69-2.31)
5	0.69 (0.50-0.87)	0.85 (0.70-1.00)	0.67 (0.02-1.31)	1.56 (0.74-2.38)
6	0.74 (0.57-0.92)	0.85 (0.70-1.00)	0.96 (0.30-1.62)	1.63 (0.80-2.45)
7	0.75 (0.57-0.92)	0.86 (0.71-1.00)	0.95 (0.29-1.61)	1.61 (0.78-2.43)
8	0.74 (0.56-0.92)	0.85 (0.70-1.00)	0.94 (0.28-1.59)	1.57 (0.75-2.39)
9	0.74 (0.56-0.92)	0.85 (0.70-1.00)	0.91 (0.26-1.57)	1.51 (0.70-2.33)
10	0.74 (0.57-0.92)	0.85 (0.70-1.00)	0.89 (0.24-1.55)	1.46 (0.65-2.26)
11	0.74 (0.57-0.91)	0.84 (0.69-0.99)	0.89 (0.23-1.54)	1.44 (0.63-2.24)
12	0.74 (0.57-0.91)	0.84 (0.69-0.98)	0.89 (0.23-1.54)	1.43 (0.63-2.24)
13	0.74 (0.57-0.91)	0.84 (0.69-0.98)	0.89 (0.24-1.55)	1.43 (0.63-2.24)
14	0.74 (0.56-0.91)	0.84 (0.69-0.98)	0.89 (0.24-1.55)	1.43 (0.63-2.24)
15	0.74 (0.56-0.91)	0.83 (0.68-0.99)	0.89 (0.23-1.54)	1.43 (0.62-2.23)
16	0.74 (0.57-0.91)	0.83 (0.67-0.98)	0.88 (0.23-1.54)	1.42 (0.62-2.22)
17	0.74 (0.57-0.90)	0.82 (0.67-0.98)	0.87 (0.22-1.53)	1.39 (0.59-2.19)
18	0.74 (0.58-0.90)	0.82 (0.67-0.98)	0.86 (0.21-1.52)	1.36 (0.56-2.16)
19	0.74 (0.58-0.90)	0.82 (0.67-0.98)	0.85 (0.20-1.51)	1.33 (0.53-2.12)
20	0.74 (0.57-0.90)	0.81 (0.65-0.97)	0.85 (0.19-1.50)	1.31 (0.52-2.10)
21	0.74 (0.57-0.90)	0.81 (0.65-0.97)	0.84 (0.19-1.50)	1.30 (0.51-2.09)
22	0.73 (0.57-0.89)	0.80 (0.64-0.96)	0.84 (0.19-1.50)	1.30 (0.51-2.09)
23	0.73 (0.57-0.89)	0.80 (0.64-0.96)	0.84 (0.19-1.49)	1.30 (0.50-2.09)
24	0.73 (0.57-0.89)	0.80 (0.64-0.96)	0.84 (0.19-1.49)	1.30 (0.50-2.09)

AUC, Area Under the Receiver Operating Characteristic Curve; UPDRS-III, Unified Parkinson's Disease Rating Scale Part III.
*Cohen's d effect size.

Supplementary table 5. ROC analysis of other baseline characteristics to predict final outcome

Predictor	AUC (95% CI)
Age	0.66 (0.499-0.817)
Computer use	0.61 (0.448-0.780)
PDQ-39*	0.64 (0.366-0.920)
UPDRS-III*	0.63 (0.381-0.878)

AUC, Area Under the Receiver Operating Characteristic Curve; PDQ-39, 39-Item Parkinson's Disease Questionnaire, UPDRS-III, Unified Parkinson's Disease Rating Scale Part III.
*Only PD subjects were included in the analysis of PDQ-39 and UPDRS-III