

Predicting Parkinson's Disease Medication Regimen Using Sensor Technology

Jeremy Watts

University of Tennessee at Knoxville

Anahita Khojandi

University of Tennessee at Knoxville

Rama Vasudevan

Oak Ridge National Laboratory

Fatta B. Nahab

University of California, San Diego

Ritesh Ramdhani (✉ rramdhani@northwell.edu)

Donald and Barbara Zucker School of Medicine at Hofstra/Northwell

Research Article

Keywords: Parkinson's disease (PD), Personal KinetiGraph™ (PKG), longitudinal dopaminergic regimens

Posted Date: February 18th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-198765/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Predicting Parkinson's Disease Medication Regimen Using Sensor Technology

Jeremy Watts¹, Anahita Khojandi¹, Rama Vasudevan², Fatta B. Nahab³, and Ritesh Ramdhani^{4,*}

¹University of Tennessee, Department of Industrial and Systems Engineering, Knoxville, 37996, USA

²Oak Ridge National Laboratory, Center for Nanophase Materials Science, Oak Ridge, 37830, USA

³University of California San Diego, Department of Neurosciences, La Jolla, 92093, USA

⁴Donald and Barbara School of Medicine at Hofstra/Northwell, New York, 10021, USA

*rramdhani@northwell.edu

ABSTRACT

Parkinson's disease (PD) medication treatment planning is generally based on subjective data through in-office, physician-patient interactions. The Personal KinetiGraph™ (PKG) has shown promise in enabling objective, continuous remote health monitoring for Parkinson's patients. In this proof-of-concept study, we propose to use objective sensor data from the PKG and apply machine learning to subtype patients based on levodopa regimens and response. We apply k-means clustering to a dataset of within-subject Parkinson's medication changes—clinically assessed by the PKG and Hoehn & Yahr (H&Y) staging. A random forest classification model was then used to predict patients' cluster allocation based on their respective PKG data and demographic information. Clinically relevant clusters were developed based on longitudinal dopaminergic regimens—partitioned by levodopa dose, administration frequency, and total levodopa equivalent daily dose—with the PKG increasing cluster granularity compared to the H&Y staging. A random forest classifier was able to accurately classify subjects of the two most demographically similar clusters with an accuracy of 87.9 ± 1.3 .

Introduction

Levodopa remains the gold standard therapy for treating the cardinal motor symptoms of Parkinson's disease (PD). As Parkinson's progresses, the duration of levodopa's dose efficacy shortens with the emergence of "wearing-off" episodes and dyskinesia. These complications result from a variety of factors, including disease progression and pulsatile stimulation of dopamine receptors due to lack of continuous levodopa administration¹. Typically, a patient's medication regimen is optimized by fragmenting and increasing levodopa dosages while utilizing monoamine oxidase B (MAO-B), dopamine agonists, or catechol-O-methyl transferase (COMT) inhibitors as adjunctive therapies.

The determination of the most effective treatment strategy is based on a clinician's overall impression of motor disability assessed by instruments such as the Unified Parkinson's disease Rating Scale (UPDRS)² and the Hauser paper-based diaries³. The lack of continuous motor assessment coupled with recall bias and limited integration of nonmotor symptomatology into the treatment paradigm present real-world limitations in managing such a heterogeneous condition. Sensor-based technology offers a real-time mechanism to objectively measure motor performance in Parkinson's disease^{4,5}, moving beyond the "snapshot" in-office assessment of clinical impairment.

Motion sensors have demonstrated 70–90% accuracy in measuring fluctuations and dyskinesia^{6–9}. One such inertial sensor is the Personal KinetiGraph™ (PKG) sensor (Global Kinetics Corporation (GKC), Melbourne, Australia). This wrist-worn logger utilizes an accelerometer to collect movement information every two minutes throughout the day resulting in approximately 500 daily assessments and reminds patients to register when they have taken their prescribed dopaminergic medication. The raw data are converted into bradykinesia and dyskinesia scores, as well as time series tracings, curated into a report⁸ using validated algorithms^{8,10–12}. The report shows the continuous changes of bradykinesia and dyskinesia scores, as it relates to levodopa timing as the median, 25th, and 75th percentile, compared to a non-PD control group over six days. Additionally, these scores have been correlated with non-motor symptoms including gastrointestinal tract, sexual function and mood/cognition¹³.

By examining a spectrum of patients with clinical variability and gauging their responsivity to dopaminergic medication with sensor technology, inherent medication similarities may be present within specific clinical subtypes—offering an opportunity to cluster patients in a treatment-related manner. This approach could serve to predict optimal regimens, potentially reducing the lengthy process of optimizing medication for patients. Strategic treatment planning using PD patient subtyping has been shown

effective^{14,15} and stands to offer a data-driven approach to refining clinical management.

Methods

Study Cohort

Characteristics of the patient cohort and selection process are described in the study by Nahab et al.¹⁶, which explored the clinical utility of the PKG in routine care of Parkinson's patients. All patients were selected from the UCSD Movement Disorder Center from June 2016 to March 2017. The study's inclusion criteria included: an age range of 46–83, being on levodopa, and a H&Y range 1–3¹⁶. The participants underwent two clinical assessment visits. Before each visit, the PKG sensor was worn by the patient for a six-day period, which scores patients' key symptoms every two minutes. During the visit, an MDS-UPDRS motor subscales III & IV² and H&Y ranking were conducted. After the first visit, a management plan was developed based on the PKG report¹⁶, the impact of which was evaluated in the second visit by the same clinical metrics, including the PKG.

K-means Clustering

K-mean clustering is an unsupervised ML algorithm that partitions patients into a predetermined number of clusters (k) without a hierarchical structure¹⁷. In this algorithm, clusters are initially formed, and each patient is grouped into their nearest cluster (with respect to Euclidian distance to cluster centroid). The clusters' centroids are then recalculated, seeking to minimize the distance between patients and their assigned centroid. Patients are then reassigned to the nearest clusters. This process is performed iteratively and continues until no patients are reassigned in an update¹⁷.

Daily total levodopa equivalent dose, daily total carbidopa/levodopa IR dose, and administration frequency were used in k-means clustering to group patients into four clusters, per the Within Cluster Sum of Squares (WCSS) measurement, which minimizes the within-cluster variance (e.g., see Fig. S1 in Supplemental information). This technique has been effectively used in healthcare applications for clustering data¹⁸. Consequently, we evaluated demographic information (patient's age at visit 1, age at diagnosis, number of years experiencing PD symptoms, and gender) for each cluster under two comparisons.

Random Forest

A random forest classifier is a supervised ML algorithm that utilizes a large number of decision trees working as an ensemble¹⁹. We opt to use the random forest classifier in this study as it is generally very robust against noisy or high dimensional datasets; it is not susceptible to overfitting²⁰.

Three random forest classifiers were trained to stratify subjects into their designated clusters—identified based on the patients' best medication regimen according to PKG scores—using either demographic information, the H&Y scores from visit 1, and the PKG time series from visit 1. Features were extracted and engineered from the PKG time series via TSFresh²¹, which calculates various time series characteristics frequently used in classification tasks and also included demographic information. Features importance ranking was conducted using the Gini index²². The topmost important features were identified, prior to cross-validation, using preliminary experiments (presented in Table S1 in the Supplementary information) and included in the analysis.

The random forest model's performance was evaluated using 5-fold stratified cross-validation such that in each fold, the ratio of cluster representation remained consistent with that of the original dataset. Due to the nature of the patient data, downsampling was used in each fold to learn from a balanced dataset such that, during training, the model did not favor the more representative cluster. Finally, the process was repeated 100 times, and the mean and confidence interval of the performance metrics were reported. Performance metrics included recall, precision, accuracy, F1 score, and the area under the receiver operating characteristic (AUC). Recall is the proportion of positives that are correctly identified. Precision is the measurement of positive and negative results that are true positives. Accuracy is the measurement of correct predictions out of all predictions. The F1 score is the harmonic average of precision and recall. The AUC is the aggregate comparison of the true positive rate and the false positive rate at different classification thresholds and provides an overall performance metric for the model.

Results

Cohort Characteristics

A total of 26 subjects' (17 male and 9 female) clinical evaluations and PKG readouts were included from the study by Nahab et al.¹⁶. The patient cohort utilized in this study is a subset of that presented in¹⁶, as two participants were excluded from the evaluation: one participant did not have corresponding PKG readouts; the other had dosage inconsistencies in their recorded medication regimen. During visit 2, the overall mean MDS-UPDRS-III was significantly reduced (visit 1: 28.9 ± 14.1 , visit 2: 24.1 ± 13.5 , p -value < 0.028 ¹⁶). Demographic information and clinical characteristics of the participants are provided in Table 1. The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board of the University of Tennessee (UTK-IRB-20-06007-XP).

Characteristics	Mean + SD	Range
Gender (female/male)	9/17	–
Age in years	71.19 ± 9.70	53-87
Age at diagnosis (years)	65.77 ± 10.37	47-80
Visit 1 H&Y stage	1.77 ± 0.71	1-3
Visit 2 H&Y stage	1.85 ± 0.78	1-3
Visit 1 levodopa equivalent dose (mg)	498.94 ± 309.88	200-1350
Visit 2 levodopa equivalent dose (mg)	637.40 ± 322.37	228-1380
Time between clinical visits (days)	65.62 ± 26.46	16-112

Table 1. Patient demographics and clinical characteristics.

Patient Clustering Using Medication Regimen

K-means clustering was utilized to allocate patients to one of four clusters based on their prescribed daily total levodopa equivalent dose, daily total carbidopa/levodopa IR dose, and administration frequency. In the first experiment, subjects were clustered by their visit 2 regimens, in which physicians had adjusted individualized regimens to optimize motor symptoms based on the PKG report. In the second experiment, subjects were clustered according to the regimen associated with best motor function improvement, as defined by the H&Y or PKG scores consistent with the comparisons of Nahab et al.¹⁶.

Figure 1(a-b) presents the clusters based on visit 2's medication regimen. According to Fig. 1(a), when H&Y is used as the comparison metric: four subjects show symptom improvements, while six demonstrated symptom worsening, and sixteen remained unchanged. As shown in Fig. 1(b), when the PKG score is used: seventeen subjects show symptom improvements, eight worsened, and one remains unchanged. The demographic information associated with each cluster is provided in Table 2. For each comparison participants were clustered into four groups denoted as clusters A-D. Cluster D was statistically different from clusters A and B with respect to disease duration and age at diagnosis ($p < 0.05$); no other clusters were statistically different within parameters ($p > 0.05$).

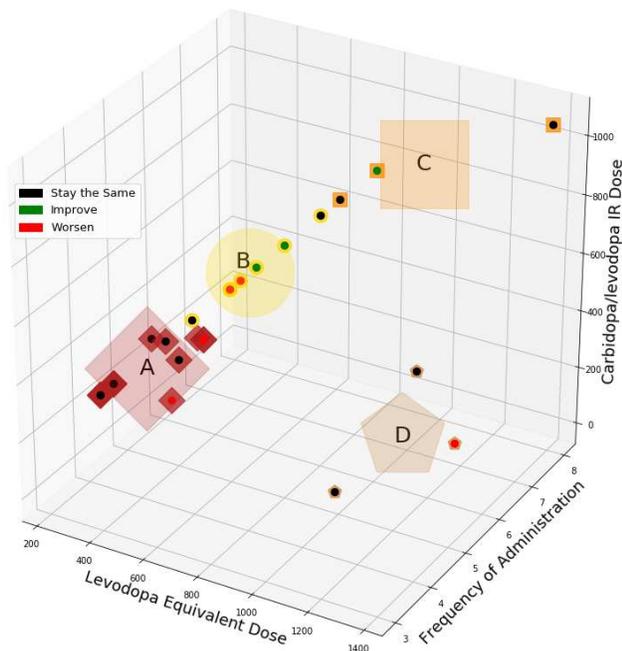
Figure 1(c-d) presents the medication regimens' clusters associated with the most improved motor function between the two study visits. The centroid positions of the clusters are in different locations due to the fact that medication regimens related to improved clinical function differ according to H&Y and PKG measures. The subjects' demographic information for each cluster is provided in Table 2, while the breakdown in PD medication and dosing are provided in Table 3. It should be noted that for H&Y and PKG scores that were unchanged between visit 1 and visit 2, the regimen associated with visit 2 was considered "best" and used in this clustering. No two clusters are statistically different ($p > 0.05$) in terms of patients' age, gender, or age at diagnosis. Cluster D was statistically different from clusters A and B with respect to disease duration ($p < 0.05$) for both the best H&Y and PKG scores.

Random Forest Classification Using PKG Readout

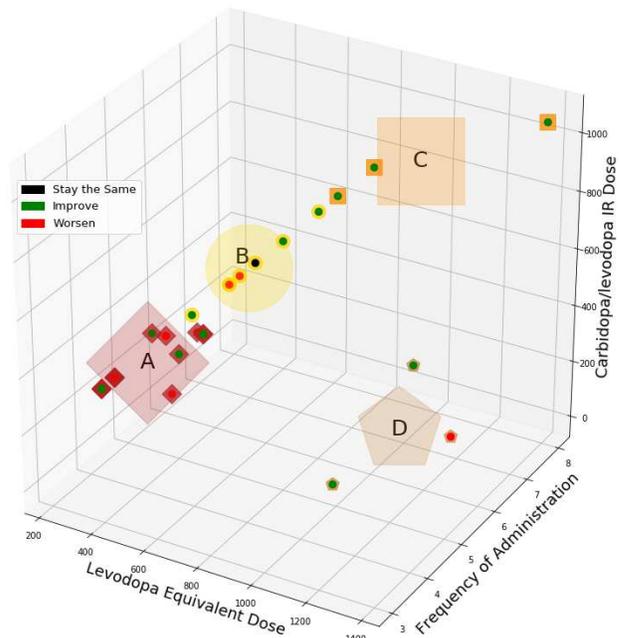
Three random forest classifiers were trained to examine the efficacy of using a combination of demographic information (patient's age at visit 1, age at diagnosis, number of years experiencing PD symptoms, and gender), H&Y scores from visit 1, and PKG time series sensor data from visit 1, to stratify the subjects in clusters A and B, as identified through k-means clustering using the best PKG score (see Fig. 1(d)). As shown in Table 2, "Best PKG Score," clusters A and B contain 17 and 6 participants, respectively. As noted in the Results Section ("Patient Clustering Using Medication Regimen"), these two clusters were the most statistically similar with respect to the demographic information and contained the majority of participants. The performance of each of the classifiers is presented in Table 4.

The random forest classifier using solely demographic information achieved a recall of $31.2 \pm 3.9\%$, a precision of $18.1 \pm 2.1\%$, an accuracy of $49.8 \pm 1.9\%$, an F1 score of $21.0 \pm 2.3\%$, and an AUC of 0.43 ± 0.03 . Whereas the random forest classifier using both demographic information and visit 1 H&Y staging achieved a recall of $49.4 \pm 4.2\%$, a precision of $22.0 \pm 2.2\%$, an accuracy of $45.0 \pm 2.0\%$, an F1 score of $29.6 \pm 2.4\%$, and an AUC of 0.47 ± 0.03 .

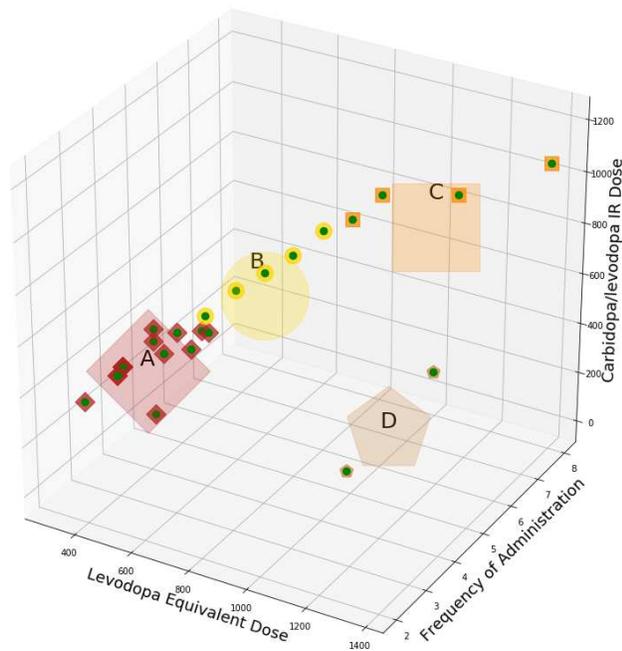
The random forest classifier using PKG time series sensor data to stratify the subjects in clusters A and B—identified through k-means clustering using the best PKG score—had superior performance to the alternative methods. To train this random forest classifier, over 1000 features were extracted from PKG sensors' bradykinesia and dyskinesia time series for each patient, of which the top ten most important features were included in the analysis. These features are provided in Table S1 in Supplemental information. This random forest classifier achieved a recall of $81.0 \pm 3.3\%$, a precision of $69.0 \pm 3.5\%$, an accuracy of $87.9 \pm 1.3\%$, an F1 score of $71.3 \pm 3.3\%$, and an AUC of 0.94 ± 0.01 .



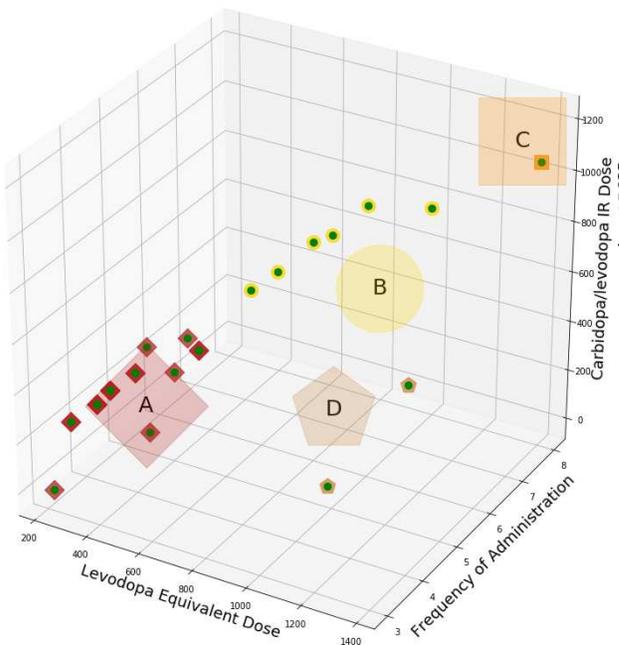
a) Visit 2 HY Change



b) Visit 2 PKG Change



c) Best HY Score Between Visits



d) Best PKG Score Between Visits

Figure 1. Figures 1(a) and 1(b) demonstrates the clusters for the motor function changes between visit 1 to visit 2 based on the H&Y ranking and PKG report, respectively. Figures 1(c) and 1(d) highlight the subjects' best symptom control recorded using H&Y and PKG scores, respectively. The large shapes denote each cluster's centroid, and the exterior marker of each point corresponds to the cluster centroid shape. The capital letters (A-D) are used to refer to each cluster. Each point's interior marker in (a) and (b) represents the state of H&Y and PKG change for each patient, respectively. Note that due to the three-dimensional projection of the plot, the distance between points may appear skewed.

	Cluster A	Cluster B	Cluster C	Cluster D
Visit 2 H&Y & PKG Score				
Study age (years)	74.31 ± 9.99	69.96 ± 9.70	68.15 ± 3.96	68.29 ± 5.68
Age at diagnosis (years)	69.69 ± 9.22	65.41 ± 9.08	60.48 ± 10.32	54.95 ± 4.15 [†]
Years of PD	4.62 ± 3.54	2.55 ± 1.78	7.67 ± 6.65	13.33 ± 2.05 [†]
Gender (female/male)	5/8	2/5	1/2	–
Number of participants	13	7	3	3
Best H&Y Score				
Study age (years)	73.31 ± 9.71	68.41 ± 11.27	66.86 ± 4.09	70.93 ± 5.24
Age at diagnosis (years)	68.77 ± 9.09	66.45 ± 10.05	57.86 ± 10.02	57.43 ± 2.74
Years of PD	4.53 ± 3.38	1.97 ± 1.29	9.00 ± 6.20	13.50 ± 2.50 [†]
Gender (female/male)	6/9	1/4	2/2	0/2
Number of participants	15	5	4	2
Best PKG Score				
Study age (years)	72.92 ± 10.21	67.66 ± 7.54	63.54 ± 0.00	70.93 ± 5.24
Age at diagnosis (years)	68.69 ± 9.47	63.49 ± 8.78	46.54 ± 0.00	57.43 ± 2.74
Years of PD	4.23 ± 3.31	4.17 ± 4.14	17.00 ± 0.00	13.50 ± 2.50 [†]
Gender (female/male)	6/11	3/3	0/1	0/2
Number of participants	17	6	1	2

Table 2. Demographic and clinical information for clusters depicted in Fig. 1(a-d). † Pairwise p-value < 0.05 when compared with both Cluster A and B. Unless denoted by †, each pairwise cluster comparison yields no statistical difference (p-value > 0.05).

Discussion

Utilizing a Parkinson’s dataset of within-subject medication regimen titrations—clinically assessed by the H&Y and PKG scores—k-means clustering was used to group patients in terms of daily total levodopa equivalent dose, daily total carbidopa/levodopa IR dose, and administration frequency. We demonstrate that subjects can be meaningfully clustered based on longitudinal dopaminergic treatment regimens. Further, the sensor-based assessments of the PKG enhance the granularity of this clustering method compared with the H&Y staging.

Figures 1(a) and 1(b) show the difference between H&Y and PKG scores when determining subject improvement. This difference is more pronounced between the two assessment instruments with respect to the regimens yielding the best motor function. This result supports the growing body of literature that the H&Y score lacks the fidelity to capture more nuanced symptom improvement trends^{23–25}. Specifically, at visit 2, the PKG scores indicate clinical improvement in seventeen patients (65%) compared to four patients (15%) based on the H&Y. This comparison shows that determining the optimization of a patient’s medication regimen may be significantly more challenging using only H&Y. However, since the cohort was treated based on PKG changes, further conclusions regarding the robustness of the treatment approaches cannot be drawn. Therefore, our results suggest that efficiently establishing a patient’s best performing regimen would be improved by objective measurements.

Additionally, the cohort’s demographics and clinical characteristics are generally statistically indistinguishable between the identified clusters based on both PKG and H&Y scores. Only subjects with the longest disease duration were grouped into a separate cluster. This group required greater dosages and more frequent administration of dopaminergics for symptom control—a treatment strategy aligned with current clinical practice. Furthermore, using demographic information along with time series sensor data from the PKG yielded a classification model that enabled the random forest classifier to predict the cluster allocation of patients with a high accuracy.

In the past, studies classifying Parkinson’s clinical subtypes have predominantly relied on clinometric scales. Early on, a tremor dominant phenotype was deemed more favorable than the postural instability and gait disorder (PIGD) phenotype²⁶ with younger age at onset associated with slower disease progression compared to a more rapid progression in those who developed symptoms at an older age. However, the myriad of nonmotor symptoms associated with the disease: REM behavioral disorder (RBD), constipation, autonomic dysfunction, and neurobehavioral dysregulation (e.g., impulse control behaviors, executive decline) have expanded the phenotypic spectrum. In several studies, four overarching phenotypic clusters have emerged: 1) mildly affected across motor and nonmotor domains; 2) non-motor dominant; 3) motor dominant; 4) severely affected across all domains^{17,27,28}.

	Cluster A	Cluster B	Cluster C	Cluster D
Best H&Y Score				
LEDD	385 ± 91	700 ± 114	1182 ± 211	1131 ± 206
Carbidopa/levodopa IR	347 ± 114	700 ± 100	1063 ± 95	250 ± 354
Carbidopa/levodopa CR	7 ± 28	–	100 ± 115	500 ± 707
Ropinirole	–	–	1 ± 2	–
Selegline CARBEX	0.4 ± 0.1	–	3 ± 5	1 ± 2
Rasagiline	0.1 ± 0.4	–	–	0.3 ± 0.4
Rytary	54 ± 203	–	–	1170 ± 1655
Best PKG Score				
LEDD	381 ± 105	942 ± 233	1380 ± 0	1131 ± 206
Carbidopa/levodopa IR	335 ± 147	917 ± 183	1050 ± 0	250 ± 354
Carbidopa/levodopa CR	18 ± 73	33 ± 82	200 ± 0	500 ± 707
Ropinirole	–	–	4 ± 0	–
Selegline CARBEX	0.6 ± 2.4	–	10 ± 0	1 ± 2
Rasagiline	0.1 ± 0.3	–	–	0.3 ± 0.4
Rytary	45 ± 184	–	–	1170 ± 1655

Table 3. Dosage and medication types for clusters depicted in Fig. 1(c-d). Medications are aggregate results within each cluster and do not represent a single subject’s regimen.

	Demographics Alone	Demographics & H&Y	Demographics & PKG
Recall	31.2 ± 3.9%	49.4 ± 4.2%	81.0 ± 3.3%
Precision	18.1 ± 2.1%	22.0 ± 2.2%	69.0 ± 3.5%
Accuracy	49.8 ± 1.9%	45.0 ± 2.0%	87.9 ± 1.3%
F1 Score	21.0 ± 2.3%	29.6 ± 2.4%	71.3 ± 3.3%
AUC	0.43 ± 0.03	0.47 ± 0.03	0.94 ± 0.01

Table 4. Random forest classifier performance identifying subjects in clusters A and B, as stratified through k-means clustering under the best PKG score. Each classifier was trained with a combination of features being only demographics, demographics and visit 1 H&Y score, and demographics and visit 1 PKG time series data.

Longitudinal clustering found nonmotor symptoms, especially cognitive impairment, RBD, and orthostatic hypotension were significant determinants of three principle PD subgroups—motor/slow progression, intermediate, diffuse/malignant²⁹. Patients in the motor cluster were least afflicted with NMS and had less worsening than those in the diffuse/malignant cluster, who had RBD, OH, MCI at baseline, and were associated with faster motor progression.

While the prognostication of disease progression is evident in clinical subtyping, the implications on treatment have yet to be established. Similar to the phenotypic variability of PD, the treatment approaches are also heterogeneous. Therefore, a continuous assessment of treatment response not only offers the possibility of more robust medication titrations, but the ability to cluster these sensor-based responses may help potentiate the impact of the emerging clinical phenotypes. This proof-of-concept study establishes that rich information can be extracted from data streams collected from sensors, such as the PKG—that measures both motor function and medication responses—and incorporated into ML algorithms to build predictive models capable of expanding the clinical treatment platform.

The small patient sample potentially introduces bias, such as the underrepresentation of various PD subtypes. Additionally, several subjects’ motor control symptoms within the cohort were never successfully controlled as measured by the PKG’s bradykinesia and dyskinesia scores³⁰. Such subjects when further optimized clinically, may place them in a different cluster—altering the demographic and clinical information associated with that cluster. Future work to examine a more representative patient cohort followed longitudinally stands to enhance the clustering and possibly reveal other inherent treatment subtypes.

Clustering patients in a treatment-related manner has the potential to determine inherent regimen similarities within clinical subtypes. The identification of such regimen similarities could guide strategic treatment planning for a data-driven approach to refine clinical management. This approach can provide clinicians with estimated regimens based on patients’ demographics and limited sensor measurements, potentially expanding the current clinical approach for optimizing regimens.

References

1. Olanow, C., Obeso, J. & Stocchi, F. Continuous dopamine-receptor treatment of Parkinson's disease: scientific rationale and clinical implications. *The Lancet Neurol.* **5**, 677–687 (2006).
2. Goetz, C. *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. Off. J. Mov. Disord. Soc.* **23**, 2129–2170 (2008).
3. Hauser, R. *et al.* Patient evaluation of a home diary to assess duration and severity of dyskinesia in Parkinson disease. *Clin. Neuropharmacol.* **29**, 322–330 (2006).
4. Ramdhani, R., Khojandi, A., Shylo, O. & Kopell, B. Optimizing clinical assessments in Parkinson's disease through the use of wearable sensors and data driven modeling. *Front. Comput. Neurosci.* **12**, 72 (2018).
5. Maetzler, W., Klucken, J. & Horne, M. A clinical view on the development of technology-based tools in managing Parkinson's disease. *Mov. Disord.* **31**, 1263–1271 (2016).
6. Keijsers, N., Horstink, M. & Gielen, S. Ambulatory motor assessment in Parkinson's disease. *Mov. Disord. Off. J. Mov. Disord. Soc.* **21**, 34–44 (2006).
7. Tsipouras, M. *et al.* Automated Levodopa-induced dyskinesia assessment. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, 2411–2414 (IEEE, 2010).
8. Griffiths, R. *et al.* Automated assessment of bradykinesia and dyskinesia in Parkinson's disease. *J. Park. Dis.* **2**, 47–55 (2012).
9. Pulliam, C. *et al.* Continuous assessment of levodopa response in Parkinson's disease using wearable motion sensors. *IEEE Transactions on Biomed. Eng.* **65**, 159–164 (2017).
10. Kotschet, K. *et al.* Daytime sleep in Parkinson's disease measured by episodes of immobility. *Park. & Relat. Disord.* **20**, 578–583 (2014).
11. Horne, M., McGregor, S. & Bergquist, F. An objective fluctuation score for Parkinson's disease. *PloS One* **10**, e0124522 (2015).
12. Braybrook, M. *et al.* An ambulatory tremor score for Parkinson's disease. *J. Park. Dis.* **6**, 723–731 (2016).
13. van Wamelen, D. *et al.* Non-motor correlates of wrist-worn wearable sensor use in parkinson's disease: an exploratory analysis. *NPJ Park. disease* **5**, 1–5 (2019).
14. Marras, C., Chaudhuri, K., Titova, N. & Mestre, T. Therapy of Parkinson's disease subtypes. *Neurotherapeutics* 1–12 (2020).
15. Charvin, D., Medori, R., Hauser, R. & Rascol, O. Therapeutic strategies for Parkinson disease: beyond dopaminergic drugs. *Nat. Rev. Drug Discov.* **17**, 804–822 (2018).
16. Nahab, F., Abu-Hussain, H. & Moreno, L. Evaluation of clinical utility of the Personal KinetiGraph® in the management of Parkinson disease. *Adv. Park. Dis.* **8**, 42 (2019).
17. Van Rooden, S. *et al.* The identification of Parkinson's disease subtypes using cluster analysis: a systematic review. *Mov. Disord.* **25**, 969–978 (2010).
18. Yock, A. & Kim, G. Using k-means clustering to determine the number and position of isocenters in MLC-based multiple target intracranial radiosurgery. *J. Appl. Clin. Med. Phys.* **18**, 351–357 (2017).
19. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
20. Liaw, A. & Wiener, M. Classification and regression by random forest. *R News* **2**, 18–22 (2002).
21. Pedregosa, F. *et al.* Scikit-learn: machine learning in Python. *The J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
22. Tan, P., Steinbach, M. & Kumar, V. *Introduction to data mining* (Pearson Education India, 2016).
23. Goetz, C. *et al.* Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: status and recommendations the Movement Disorder Society Task Force on rating scales for Parkinson's disease. *Mov. Disord.* **19**, 1020–1028 (2004).
24. Goetz, C. *et al.* Levodopa. *Mov. Disord.* **17**, S23–S37 (2002).
25. Diamond, S. & Markham, C. Evaluating the evaluations: or how to weigh the scales of parkinsonian disability. *Neurology* **33**, 1098–1098 (1983).

26. Jankovic, J. Parkinson's disease: clinical features and diagnosis. *J. Neurol. Neurosurg. & Psychiatry* **79**, 368–376 (2008).
27. Mu, J. *et al.* Parkinson's disease subtypes identified from cluster analysis of motor and non-motor symptoms. *Front. Aging Neurosci.* **9**, 301 (2017).
28. Erro, R. *et al.* Clinical clusters and dopaminergic dysfunction in de-novo Parkinson disease. *Park. & Relat. Disord.* **28**, 137–140 (2016).
29. Fereshtehnejad, S. *et al.* New clinical subtypes of Parkinson disease and their longitudinal progression: a prospective cohort comparison with other phenotypes. *JAMA Neurol.* **72**, 863–873 (2015).
30. Pahwa, R. *et al.* Role of the Personal KinetiGraph in the routine clinical assessment of Parkinson's disease: recommendations from an expert panel. *Expert. Rev. Neurother.* **18**, 669–680 (2018).

Acknowledgements

This study was partially supported by the Science Alliance, The University of Tennessee, and by the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory managed by UT-Battelle, LLC, for the U.S. Department of Energy (DOE).

Author contributions statement

J.W., A.K., and R.R. conceived and designed the study. J.W. wrote the first draft and performed the data analysis. F.N. collected patient data. J.W., A.K., R.V., F.N., and R.R. provided critical insights in editing the manuscript. All authors have approved and contributed to the final written manuscript.

Additional information

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board of the University of Tennessee (UTK-IRB-20-06007-XP).

Conflicts of Interest: Dr. Nahab had no competing interests to declare at the time of the data collection, analysis and initial publication. At present, Dr. Nahab is employed by Global Kinetics. He has no other competing interests. The other authors declare no conflicts of interest.

Data Availability Statement: Restrictions apply to the availability of these data. Data was obtained from Nahab, F.¹⁶ and are available from the corresponding author with the permission of Nahab, F. All subjects signed informed consent in Nahab et al.¹⁶

Figures

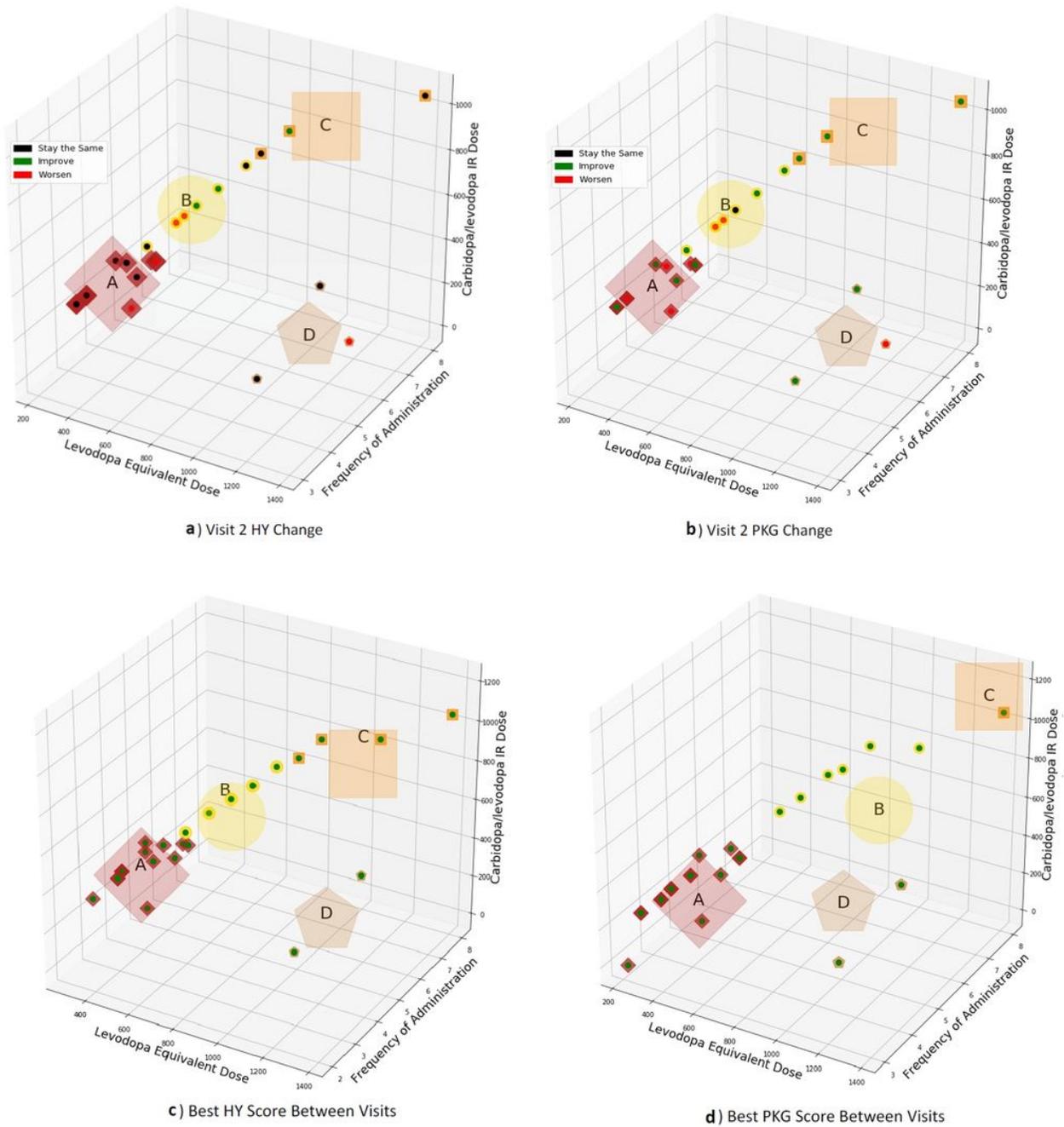


Figure 1

Figures 1(a) and 1(b) demonstrates the clusters for the motor function changes between visit 1 to visit 2 based on the H&Y ranking and PKG report, respectively. Figures 1(c) and 1(d) highlight the subjects' best symptom control recorded using H&Y and PKG scores, respectively. The large shapes denote each cluster's centroid, and the exterior marker of each point corresponds to the cluster centroid shape. The

capital letters (A-D) are used to refer to each cluster. Each point's interior marker in (a) and (b) represents the state of H&Y and PKG change for each patient, respectively. Note that due to the three-dimensional projection of the plot, the distance between points may appear skewed.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementalInformation.pdf](#)