Discriminating Progressive Supranuclear Palsy from Parkinson's Disease using wearable technology and machine learning.

Author names and affiliations

Maarten De Vos[1], John Prince[1], Tim Buchanan PhD[2], James J. FitzGerald PhD[3,4], and Chrystalina A. Antoniades PhD[4*]

[1] Department of Engineering Science, Institute of Biomedical Engineering, University of Oxford, Old Road Campus Research Building, OX3 7DQ, Oxford, UK.

[2] UCB Biopharma SPRL, Brussels, Belgium

[3] Nuffield Department of Surgical Sciences, University of Oxford, Oxford, OX3 9DU, UK.

[4] Nuffield Department of Clinical Neurosciences, NeuroMetrology Lab, University of Oxford, Oxford, OX3 9DU, UK.

Corresponding author

Professor Chrystalina A. Antoniades, Nuffield Department of Clinical Neurosciences, Level 6, West Wing, John Radcliffe Hospital, University of Oxford, United Kingdom

E-mail: chrystalina.antoniades@ndcn.ox.ac.uk

Word count: 2999 (excluding references)

References: 23

RUNNING TITLE: Wearing technology for PSP

KEYWORDS: Progressive Supranuclear Pasly, Parkinson's disease, gait, wearables, inertial sensor array, machine learning.

FINANCIAL DISCLOSURE / CONFLICT OF INTEREST:

MDV, JP,TB, and CAA have no conflicts of interest.

JJF has performed consultancy unrelated to the subject of this study for companies including Abbott, Renishaw, and Herantis, and also has a research grant from Abbott, for a project unrelated to this study.

FUNDING SOURCES FOR THE STUDY:

This was study was funded by a grant from the UCB-Oxford Collaboration. JJF and MDV were supported by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC).

Abstract

Background: Progressive supranuclear palsy (PSP), a neurodegenerative conditions may be difficult to discriminate clinically from idiopathic Parkinson's disease (PD). It is critical that we are able to do this accurately and as early as possible in order that future disease modifying therapies for PSP may be deployed at a stage when they are likely to have maximal benefit. Analysis of gait and related tasks is one possible means of discrimination.

Research Question: Here we investigate a wearable sensor array coupled with machine learning approaches as a means of disease classification.

Methods: 21 participants with PSP, 20 with PD, and 39 healthy control (HC) subjects performed a two minute walk, static sway test, and timed up-and-go task, while wearing an array of six inertial measurement units. The data were analysed to determine what features discriminated PSP from PD and PSP from HC. Two machine learning algorithms were applied, Logistic Regression (LR) and Random Forest (RF).

Results: 17 features were identified in the combined dataset that contained independent information. The RF classifier outperformed the LR classifier, and allowed discrimination of PSP from PD with 86% sensitivity and 90% specificity, and PSP from HC with 90% sensitivity and 97% specificity. Using data from the single lumbar sensor only resulted in only a modest reduction in classification accuracy, which could be restored using 3 sensors (lumbar, right arm and foot). However for maximum specificity the full six sensor array was needed.

Significance: A wearable sensor array coupled with machine learning methods can accurately discriminate PSP from PD. Choice of array complexity depends on context; for diagnostic purposes a high specificity is needed suggesting the more complete array is advantageous, while for subsequent disease tracking a simpler system may suffice.

Introduction

Progressive supranuclear palsy (PSP) (1, 2) is an atypical parkinsonian disorder that can be difficult to discriminate clinically from the much commoner Parkinson's disease (3-5). Well developed PSP is characterized by vertical supranuclear gaze palsy, postural instability, axial rigidity, and mild cognitive impairment (1) (4-6), but early PSP shares many features with PD including abnormalities of gait, speech, and eye movements. The Movement Disorders Society (MDS) has produced diagnostic criteria for PSP, with the aim of improving the detection of the disease in clinical practice and research (7).

Comparative studies using accelerometers (8) have shown many shared gait abnormalities (9) (10-12) including decreased velocity, step length, cadence, and mean acceleration. Differences have also been found, including lower vertical displacement and higher acceleration in PSP. In this paper we use a wearable array of inertial measurement units (IMUs), coupled with machine learning algorithms, to distinguish PSP from PD and from healthy controls.

There is a growing number of wearable technologies (13, 14) for characterizing the motor features of neurodegenerative diseases(15-17). When deciding what measurement technology to use two major issues must be considered. One is the setting for the measurement. Laboratory measurements in controlled and supervised conditions yield high quality standardized datasets, and are well suited to tasks like detailed measurement of individual components of the gait cycle. However there is growing interest in longer term ambulatory measurement because

snapshot laboratory recordings may not give an accurate overall picture, due to variability of symptoms by time of day, inconsistent timing of medication, or the stress of being tested in a hospital environment. Laboratory measurements are also unlikely to capture infrequent but important features such as falls or near-falls, and a 'real world' environment may better elicit features that are of direct clinical relevance to the patient.

The second issue is the device complexity. There is a conflict between maximising data and minimising the effort involved in obtaining it. Complex arrays comprising numerous inertial measurement units attached to all four limbs and the trunk may stream data at tens of megabytes per minute. At the opposite extreme are systems based on single sensors, even those built into consumer smartphones. The complex systems will generate much more complete information; the question is can the simpler systems answer the same diagnostic and measurement questions acceptably well, despite their comparatively meagre dataset?

The two issues are linked. Applying a complex system in a laboratory setting is straightforward and is logical given the intent of obtaining detailed data in a concentrated period of time. Using a complex system at home is another matter: donning and doffing may not be easy, and the multiple sensors may get in the way of everyday activities.

Our aims in this study **are twofold**. Firstly, we will use a complex laboratory system with sensors on each limb and the trunk, to record data during commonly applied gait-related tasks. We will explore how well the data can distinguish PSP from PD and from healthy control (HC) participants, and what the most distinguishing features within the dataset are. Secondly, we will analyse which sensors/body

locations are necessary to acquire this data, to determine how far the sensor array can be simplified while still yielding satisfactory results. It is hoped that this will help maximise compliance with future study protocols while avoiding collecting an inadequate dataset due to oversimplification.

Materials & Methods

Participants

The participants were recruited as part of the Oxford study of Quantification in Parkinsonism (OxQUIP); a large clinical observational study being conducted at the John Radcliffe Hospital, Oxford. We recruited 20 participants with PD, 21 participants with PSP, and 39 healthy control participants. Of the 21 PSP participants, 17 were of the PSP–P (PSP with predominant parkinsonism) subtype and 4 PSP–RS (PSP with Richardson's syndrome). HC participants were spouses of the PD or PSP participants. Participant demographics are summarised in Table 1.

Sensor array and software

Participants wore six synchronized inertial measurement units (IMUs) (Opal^{TM}, APDM, Portland, USA), wirelessly connected to software (Mobility Lab^{TM}, APDM) running on a nearby laptop. IMUs were positioned over the lumbar spine, sternum, left and right wrists, and left and right feet. Each provided tri-axial accelerometer, gyroscope, and magnetometer signals sampled at 100Hz. Participants performed a two minute walk, sway test, and timed up-and-go (TUG). Using data from all sensors, the software automatically extracts a range of clinical features specific to the three tasks (figure 2). The full feature set included 109 parameters from analysis of the gait task, 33 from the sway test, and 14 from the TUG.

Participants taking medication were recorded in the 'ON medication' state.

Tasks

All participants completed a two minute walk on the same straight and level surface to record their gait. Next, to measure sway, participants were asked to stand upright and as still as possible for thirty seconds with their eyes closed on a firm surface. A template was placed on the floor to ensure all participants' feet were the same distance apart. The TUG was then repeated three times. Each repetition entailed the participant starting in a sitting position for three seconds, followed by standing up when instructed, walking forward 3 metres, performing a 180 degree turn, walking back to the chair and sitting back down.

Hypothesis Testing

The first aspect of this analysis aims to identify features that demonstrate statistically significant differences between the disease groups. Specifically, we aim to determine whether the PSP group shows a difference to HC and PD participants in a feature. Each participant contributes one instance in each of the three tasks. For each feature, a one-way-ANalysis Of VAriance (ANOVA) was performed to determine if a statistical difference exists between any of the sub-groups. The null hypothesis of this three-variable, two-tailed ANOVA was that no significant difference exists between the feature distributions of each sub-group at the 0.05 level. The assumption of normality is made here due to the comparatively small sample size of participants in each sub-group. However, as will be demonstrated, a non-parametric assumption is adopted in a later section of the analysis. Where indicated by ANOVA, independent t-tests are used to inspect for statistical significance between the three disease groups using a significance level of 0.05.

Disease Classification

In order to assess to what extent PSP can be automatically discriminated from PD and HC based on the measured features, we performed automated classification with a machine-learning algorithm. For each test, repeated 10-fold cross validation is performed based on the full feature set from each test separately. This means that the full dataset is split in randomly into 10 subsets, where 9 subsets are used for training the machine learning model and the remaining subset is used as an independent validation set. This training and validation is repeated 10 times where each time a different independent validation set is used. Within each repeat of cross validation, the cross sectional baseline subset undergoes minority class balancing prior to being assigned to a fold. (Minority class balancing is a technique commonly employed in the training and validation of machine learning classification models to prevent presenting the model with larger quantities of one class than the other and thus it prevents the introduction of bias. By randomly sampling the class with the largest number of data points (or majority class) such that it possesses an equal number of data points to that present in the smallest class (or minority class), a perfectly class balanced dataset is produced.) Within each fold, the training and validation sets undergo zero-mean unit-variance normalisation with respect to the mean and standard deviation of the training set. Because many of the features within each test are correlated (e.g. Gait Speed Left and Gait Speed Right), feature selection (Least Absolute Shrinkage and Selection Operator, LASSO) is performed on each training set, with the selected features subsequently being extracted from the validation set. Two classifiers were used. A Logistic Regression (LR) classifier was used due to its popularity in the clinical field. A Random Forest (RF) classifier was

also used, without feature reduction as Random Forest classifiers are known to deal well with a high dimensional feature set. We report the accuracy, sensitivity, and specificity for each task separately and all tasks combined.

Reduced sensor set

We took the features selected by LASSO for all tasks combined and assessed from which sensor those features are computed. We then repeated the classification procedure using only those features extracted from the lumbar sensor, because this sensor alone accounted for a majority of the features identified by LASSO. We further repeated the classification using data obtainable from the lumbar sensor plus the right arm and right leg sensors, and reported the same performance metrics.

Results

Participants

Demographics, clinical characteristics, and rating scale scores of the participants for each group are shown in table 1.

Figure 1 shows the distributions of some of the feature values extracted from the three tasks that were found to differ significantly between PSP and the other groups. Considerable redundancy exists amongst the large number of features, and not all measurements that were found to be significant individually emerged from the LASSO as independent predictors. For example, mean coronal sway velocity was individually significant, but did not feature in the LASSO result.

Table 2 shows disease classification accuracy under a number of conditions. The upper three rows of the table show the accuracy, sensitivity, and specificity of discrimination between PSP and PD and between PSP and HC, for each of the three tasks separately as indicated in the left column. Sway is much less effective than the other tasks at discriminating between PSP and PD. The fourth row of table 2 gives the results of analysing the combined feature set from all three tasks; this performs better than any of the tasks individually.

The RF classifier performs better than the LR classifier overall. When distinguishing PSP from PD the sensitivity, specificity, and accuracy of RF was as good as or better than LR in every condition in the table. The greatest difference was 15 percentage points in specificity using the combined tasks and full sensor set (75% with LR versus 90% with RF).

Table 3 lists the features emerging from the LASSO analysis of the combined tasks as the most important discriminators between the groups, with their weightings and which sensor the data are obtained from in each case. The lumbar sensor provides the data for ten of the seventeen features in this list, compared to just one feature that is dependent on the sternal sensor. This prompts one to ask what the results would be if the lumbar sensor were used alone. The answer (table 2 row 5) is that the accuracy of both classifiers is reduced modestly, from 80% to 78% with LR and from 88% to 85% with RF.

We then explored the effect of adding one arm and one leg sensor to the lumbar sensor, making a three sensor array including lumbar, right arm and right foot. Where LASSO identified a left sided parameter we substituted the equivalent parameter from the opposite side, making the assumption that, for example, cadence measured from either leg would be similar. Using this intermediate size network, the accuracy with both classifiers was the same as for the full six sensor network (table 2 row 6). However, the high specificity (90%) observed with the 6 sensor network was reduced to 85% with the lumbar sensor alone, and this was not recovered using 3 sensors.

Discussion

We have shown that a wearable IMU array and machine learning methods can accurately differentiate PSP from PD and from control. To ensure the robustness of these findings we used separate training and validation datasets for the analysis, and to provide validation we used two different classification algorithms. Overall, the Random Forest classifier performed better than the Logistic Regression classifier. LR is probably the most commonly used classifier in biomedical research and therefore the apparent superiority of RF is an important finding.

A recent review found 78 published studies of IMU based gait analysis, but only 16% used more than one sensor (18). Only two previous studies examined gait in PSP. One used a single lumbar IMU (8), and detected only a difference in vertical displacement during gait between PSP and PD. The other used two sensors, one on each foot (19), and identified significant differences in gait speed and cadence. In both cases the information obtained was far less detailed than that obtained here.

There is much redundancy in the measured features within and across the three tasks used. For example, the mean velocity of postural sway in the coronal plane was significantly and substantially different between PSP and the other groups, yet did not occur as an independent predictor in the feature set selected by the LASSO analysis. Only one sway variable was found to contain independent information.

With regard to overall classification accuracy, it may be possible to simplify the sensor array. Of the 17 parameters identified by the LASSO analysis, 10 could be obtained from just the lumbar sensor. In distinguishing PSP from HC, the accuracy, sensitivity, and specificity using just this sensor was virtually identical to that

obtained using the full sensor array. Using just the lumbar sensor did degrade performance in terms of test accuracy when differentiating PSP from PD, but only modestly (for the RF classifier there was a test accuracy of 88% with six sensors versus 85% with the lumbar sensor only). An intermediate 3 sensor array (lumbar, right arm, right foot) gave a test accuracy as good as the full six sensor array.

In some respects however the most important parameter is specificity, because a higher specificity increases positive predictive value (PPV), which is critical if PSP specific treatments are to be initiated based on the result. The specificity of 90% using the 6 sensor array and the RF classifier is remarkably good. Previous biomarker studies using diverse approaches have not yielded specificities this high. MRI morphometry of the brainstem gave a specificity of 85%(20), while a meta-analysis of studies of the most promising CSF biomarker, neurofilament light chain, gave a specificity of 81%(21). Because of the low prevalence of PSP amongst patients presenting with parkinsonism, even the higher 85% figure translated into a PPV of just 57% (20). The improvement in PPV between a test with 85% and 90% specificity is substantial. In moving from the full 6 sensors to just the lumbar sensor, the 90% specificity of our RF classifier fell to 85%, an important reduction in efficacy. Performance was not restored by adding in two more sensors.

The variation in model performance is often subtle: for example, reducing from 6 sensors to just the lumbar sensor reduced the accuracies of the LR and RF models by 2% and 3% respectively. However the likelihood of the variation being due to noise is small. The reduction in performance occurs systematically across all metrics (excluding the sensitivity of the random forest which remains unchanged). As the two classification techniques differ in their underlying mechanisms

(parametric and non-parametric) and their differing use of feature selection, this systematic reduction and agreement between the models adds confidence that the results are genuine. Furthermore, extensive cross-validation was utilized during the training and validation of all models, i.e. the models were trained and validated using the full spectrum of available data in multiple permutations. This results in the models being less prone to over-fitting to the training data and thus more robust in the presence of noise.

The job that the classifiers were asked to do in this study is considerably easier than the applications to which we hope the approach will ultimately extend. Classification accuracies of 90% are unsurprising when discriminating PSP patients from controls. The fact that the classification accuracy between PSP and PD was not far behind (88% with RF) may seem impressive, but we began with groups of clear-cut PSP and PD, who would be readily distinguishable clinically by an experienced movement disorders neurologist. For the technique to be most useful, it must do something that such a person finds challenging. The performance differential between more complete and simpler sensor sets may well widen with more difficult questions and will need to be carefully evaluated in each situation.

All of the PD participants and 7 of the 21 PSP participants were taking dopaminergic medication and all were tested while on their medication. Differences in dosage between the PD and PSP groups are a potential confounding variable and could influence measurements of some variables, for example postural sway, which is known to be increased by levodopa (22). In future work we hope to test participants in the off medication state.

The next step will be to evaluate the effectiveness of this approach in the differential diagnosis of these conditions at a stage when there is still substantial clinical diagnostic uncertainty. This will require a large incident cohort of parkinsonian patients, so that it contains sufficient numbers of cases that ultimately manifest as PSP.

Whereas this research has focused entirely on the role of machine learning to perform disease classification based on a single clinical visit, the focus of future work is the application of machine learning to perform longitudinal monitoring of disease severity, a capability that is very much needed for trials of potential disease modifying drugs. The OxQUIP study monitors participants every three months for up to three years. Future machine learning models will aim to perform longitudinal monitoring of symptoms and to predict disease severity via regression analyses. The use of wearable sensors has proven capable of detecting the deterioration of motor symptoms in remote environments whilst also correlating with a weighted UPDRS score (23).

In conclusion, whether the sensor network can be simplified depends on context and the exact parameters we are interested in. It may be that the more complex sensor array is better suited to the diagnostic setting, where specificity is paramount, but then a simpler arrangement can be used for quantification and tracking once the diagnosis is secure.

ACKNOWLEDGMENTS

We thank the patients and their families for their endless support during the OxQUIP study.

AUTHORS' ROLES

Research project: Conception, organization, execution CAA, JFF

Statistical analysis: Design, execution, review & critique MDV, JP, JJF and CAA

Manuscript: Drafting, review & critique MDV, JP, TB, JJF and CAA

Figure 1: Examples of gait parameters that distinguish PSP from PD and HC. A: Gait cadence; B: Mean postural sway velocity in the coronal plane during the sway test; C: Mean time taken to sit from standing during the timed up-and-go (TUG) task; D: Mean time taken to turn during the gait task; E: Mean time taken to turn during the TUG task; F: Standard deviation of time taken to turn during the gait task. Note that the large number of parameters generated by the three tasks have considerable redundancy so that some parameters that are significantly different between the conditions individually are not in the LASSO output.

Figure 2. A full list of task features.

References

1. Litvan I, Agid Y, Calne D, Campbell G, Dubois B, Duvoisin RC, et al. Clinical research criteria for the diagnosis of progressive supranuclear palsy (Steele-Richardson-Olszewski syndrome): report of the NINDS-SPSP international workshop. Neurology. 1996 Jul;47(1):1-9. PubMed PMID: 8710059. eng.
2. Kuniyoshi S, Riley DE, Zee DS, Reich SG, Whitney C, Leigh RJ. Distinguishing progressive supranuclear palsy from other forms of Parkinson's disease: evaluation of new signs. Annals of the New York Academy of Sciences. 2002 Apr;956:484-6. PubMed PMID: 11960847. eng.
3. Fearnley JM, Lees AJ. Ageing and Parkinson's disease: substantia nigra regional selectivity. Brain. 1991 Oct;114 (Pt 5):2283-301. PubMed PMID: 1933245. eng.
4. Boxer AL, Yu JT, Golbe LI, Litvan I, Lang AE, Hoglinger GU. Advances in progressive supranuclear palsy: new diagnostic criteria, biomarkers, and therapeutic approaches. Lancet neurology. 2017 Jul;16(7):552-63. PubMed PMID: 28653647. Pubmed Central PMCID: 5802400.
5. Lamb R, Rohrer JD, Lees AJ, Morris HR. Progressive Supranuclear Palsy and Corticobasal Degeneration: Pathophysiology and Treatment Options. Current treatment options in neurology. 2016 Sep;18(9):42. PubMed PMID: 27526039. Pubmed Central PMCID: 4985534.
6. Schrag A, Ben-Shlomo Y, Quinn NP. Prevalence of progressive supranuclear palsy and multiple system atrophy: a cross-sectional study. Lancet. 1999 Nov 20;354(9192):1771-5. PubMed PMID: 10577638. eng.
7. Hoglinger GU, Respondek G, Stamelou M, Kurz C, Josephs KA, Lang AE, et al. Clinical diagnosis of progressive supranuclear palsy: The movement disorder society criteria. Mov Disord. 2017 Jun;32(6):853-64. PubMed PMID: 28467028. Pubmed Central PMCID: 5516529.
8. Hatanaka N, Sato K, Hishikawa N, Takemoto M, Ohta Y, Yamashita T, et al. Comparative Gait Analysis in Progressive Supranuclear Palsy and Parkinson's Disease. European neurology. 2016;75(5-6):282-9. PubMed PMID: 27288001.
9. Murray MP, Sepic SB, Gardner GM, Downs WJ. Walking patterns of men with parkinsonism. American journal of physical medicine. 1978 Dec;57(6):278-94. PubMed PMID: 742658.
10. Morris ME, Martin CL, Schenkman ML. Striding out with Parkinson disease: evidence-based physical therapy for gait disorders. Physical therapy. 2010 Feb;90(2):280-8. PubMed PMID: 20022998. Pubmed Central PMCID: 2816030.
11. Hausdorff JM, Cudkowicz ME, Firtion R, Wei JY, Goldberger AL. Gait variability and basal ganglia disorders: stride-to-stride variations of gait cycle timing in Parkinson's disease and Huntington's disease. Mov Disord. 1998 May;13(3):428-37. PubMed PMID: 9613733.
12. Egerton T, Williams DR, Iansek R. Comparison of gait in progressive supranuclear palsy, Parkinson's disease and healthy older adults. BMC neurology. 2012 Oct 2;12:116. PubMed PMID: 23031506. Pubmed Central PMCID: 3517411.
13. Godinho C, Domingos J, Cunha G, Santos AT, Fernandes RM, Abreu D, et al. A systematic review of the characteristics and validity of monitoring technologies to assess Parkinson's disease. Journal of neuroengineering and rehabilitation. 2016 Mar 12;13:24. PubMed PMID: 26969628. Pubmed Central PMCID: 4788909.
14. Brognara L, Palumbo P, Grimm B, Palmerini L. Assessing Gait in Parkinson's Disease Using Wearable Motion Sensors: A Systematic Review. Diseases. 2019 Feb 5;7(1). PubMed PMID: 30764502. Pubmed Central PMCID: 6473911.

15. Baston C, Mancini M, Schoneburg B, Horak F, Rocchi L. Postural strategies assessed with inertial sensors in healthy and parkinsonian subjects. Gait & posture. 2014;40(1):70-5. PubMed PMID: 24656713. Pubmed Central PMCID: 4383136.
16. Mancini M, Carlson-Kuhta P, Zampieri C, Nutt JG, Chiari L, Horak FB. Postural sway as a marker of progression in Parkinson's disease: a pilot longitudinal study. Gait & posture. 2012 Jul;36(3):471-6. PubMed PMID: 22750016. Pubmed Central PMCID: 3894847.
17. Mancini M, Zampieri C, Carlson-Kuhta P, Chiari L, Horak FB. Anticipatory postural adjustments prior to step initiation are hypometric in untreated Parkinson's disease: an accelerometer-based approach. Eur J Neurol. 2009 Sep;16(9):1028-34. PubMed PMID: 19473350. Pubmed Central PMCID: 2840629.
18. Vienne A, Barrois RP, Buffat S, Ricard D, Vidal PP. Inertial Sensors to Assess Gait Quality in Patients with Neurological Disorders: A Systematic Review of Technical and Analytical Challenges. Front Psychol. 2017;8:817. PubMed PMID: 28572784. Pubmed Central PMCID: PMC5435996.
19. Raccagni C, Gassner H, Eschlboeck S, Boesch S, Krismer F, Seppi K, et al. Sensor-based gait analysis in atypical parkinsonian disorders. Brain Behav. 2018 Jun;8(6):e00977. PubMed PMID: 29733529. Pubmed Central PMCID: PMC5991583.
20. Kim YH, Ma HI, Kim YJ. Utility of the Midbrain Tegmentum Diameter in the Differential Diagnosis of Progressive Supranuclear Palsy from Idiopathic Parkinson's Disease. Journal of clinical neurology. 2015 Jul;11(3):268-74. PubMed PMID: 26174787. Pubmed Central PMCID: 4507382.
21. Ge F, Ding J, Liu Y, Lin H, Chang T. Cerebrospinal fluid NFL in the differential diagnosis of parkinsonian disorders: A meta-analysis. Neuroscience letters. 2018 Oct 15;685:35-41. PubMed PMID: 30036569.
22. Rocchi L, Chiari L, Horak FB. The effects of deep brain stimulation and levodopa on postural sway in subjects with Parkinson's disease. J Neurol Neurosurg Psychiatry. 2002 Sep;73(3):240. PubMed PMID: 12185151. Pubmed Central PMCID: 1738008.
23. Prince J, Arora S, de Vos M. Big data in Parkinson's disease: using smartphones to remotely detect longitudinal disease phenotypes. Physiological measurement. 2018 Apr 26;39(4):044005. PubMed PMID: 29516871.

	PD (n=20) Mean (range)	PSP (n=21) Mean (range)	HC (n=39) Mean (range)
Age, yrs	66.4 (50-79)	71 (63-89)	67.1 (51-82)
Gender, Male/Female	11:9	12:9	19:20
Disease duration (yrs)	11.4	2.0	N/A
UPDRS motor score (Part III)	27.9 (9-52)	44.6 (21-72)	3.1 (0-12)
MMSE	26.6 (25-29)	25.8 (20-30)	27.6 (26-30)
MOCA	26.6 (24-30)	22.0 (12-29)	28.5 (27-30)
LED (mg/day)	345.5 (50-600)	598.3 (150-1340)	N/A

Table 1. Demographics, clinical characteristics and cognitive scores in the 3 groups. PD = Parkinson's disease, HC = healthy controls and PSP = Progressive Supranuclear Palsy. UPDRS = Unified Parkinson's Disease Rating Scale, MMSE = Mini Mental state examination. MOCA = Montreal Cognitive Assessment. LED = levodopa equivalent dose.

Task	Sensors used	Comparison	Logistic Regression			Random Forest		
			Acc (%)	Sens (%)	Spec (%)	Acc (%)	Sens (%)	Spec (%)
Gait	Lumbar, sternal, both arms, both feet	PSP vs HC	87	86	87	91	85	94
		PSP vs PD	80	81	75	83	81	85
Sway	Lumbar, sternal, both arms, both feet	PSP vs HC	68	52	77	82	67	90
		PSP vs PD	63	66	60	63	67	60
TUG	Lumbar, sternal, both arms, both feet	PSP vs HC	90	81	97	92	86	95
		PSP vs PD	70	71	70	83	86	80
Combined	Lumbar, sternal, both arms, both feet	PSP vs HC	93	86	97	95	90	97
		PSP vs PD	80	85	75	88	86	90
Combined	Lumbar only	PSP vs HC	93	85	97	95	90	97
		PSP vs PD	78	76	80	85	86	85
Combined	Lumbar, right arm, right foot	PSP vs HC	93	89	96	93	90	97
		PSP vs PD	80	85	75	88	90	85

Table 2. Classification results when automatically classifying different patients based on their gait, sway, and TUG signatures using the entire six-sensor array (first three rows of table). Sway is the least informative test. Classification accuracy is improved when the feature sets from all three tasks are merged (fourth row). The Random Forest (RF) classifier performs better than the logistic regression (LR) classifier. Using data from the lumbar sensor alone (row 5) reduces accuracy modestly, but this can be recovered by adding back in arm and foot sensors from one side of the body (last row of table). The high specificity of PSP versus PD classification obtained with RF (90%) is only seen when using the 6 sensor array.

Activity	Feature	Weighting	Sensor	All sensors	Lumbar only	Lumbar, arm, leg
TUG	'Turns - Duration (s)'	0.11328	Lumbar	X	X	X
TUG	'Gait - Lower Limb - Cadence L (steps/min) STD'	0.09398	Left Leg	X		X
TUG	'Turns - Angle (degrees) STD'	0.06248	Lumbar	X	X	X
TUG	'Sit to Stand - Duration (s)'	0.04518	Lumbar	X	X	X
Gait	'Gait - Trunk - Sagittal Range of Motion (degrees) STD'	-0.03971	Trunk	X		
Gait	'Gait - Lower Limb - Toe Off Angle R (degrees) STD'	0.03182	Right Leg	X		X
TUG	'Turns - Duration (s) STD'	0.03071	Lumbar	X	X	X
TUG	'Stand to Sit - Lean Angle (degrees)'	-0.03009	Lumbar	X	X	X
TUG	'Turns - Angle (degrees)'	-0.02602	Lumbar	X	X	X
TUG	'Duration (s)'	0.01957	Lumbar	X	X	X
Gait	'Gait - Upper Limb - Arm Swing Velocity R (degrees/s)'	-0.01876	Right Arm	X		X
Gait	'Gait - Upper Limb - Arm Range of Motion R (degrees)'	-0.01631	Right Arm	X		X
Gait	'Turns - Steps in Turn (#) STD'	0.01618	Lumbar	X	X	X
Sway	'Postural Sway - Acc - Frequency Dispersion (AD)'	0.01395	Lumbar	X	X	X
Gait	'Gait - Upper Limb - Arm Range of Motion L (degrees)'	-0.00735	Left Arm	X		X
Gait	'Gait - Lower Limb - Toe Out Angle L (degrees)'	0.00101	Left Leg	X		X
Gait	'Turns - Duration (s)'	0.00025	Lumbar	X	X	X

Table 3. Key parameters emerging from LASSO analysis discriminating PSP from PD and HC subjects, together with their relative importance (as reflected by the weights). The 'sensor' column explains which of the sensors in the array is responsible for collecting the data described, and the activity column on the left shows during which test the feature has been collected. The three columns on the right show how much of the feature set is available depending on the extent of sensor array used.