Automated Speech Analysis to Reveal Early Biomarkers in Parkinson's Disease

Abstract

Automated speech analysis examinations have been established as an effective method in detecting early biomarkers in Parkinson's disease, thereby enhancing the existing diagnostic methodology. The objective of this study was to evaluate the speech feature variables used in an automated speech analysis exam to determine whether a reduced model could more effectively predict the presence of neurodegeneration in an individual. To investigate this, a random forest model was developed to assess variable importance. Mean decrease in Gini values were employed to select variables to be incorporated into a reduced model. Through the use of logistic regression analyses and other various metrics, the full and reduced models were compared. Further examination of the culminative results show that an automated speech analysis examination comprised of speech feature variables more strongly associated with the presence of neurodegeneration in an individual may be an improvement to the original examination model. Future studies should account for gender differences and focus on phonation and timing as findings suggest these subsystems of connected speech are the strongest indicators of neurodegeneration.

Background

Parkinson's disease (PD) is a neurodegenerative disease that is progressive in nature and characterized by a decline of cognitive and motor functions. It is the second most-common neurodegenerative disease globally, affecting approximately 10 million individuals worldwide (Ball et al., 2019). In the U.S., the death rate for PD has increased by 63% in the last two decades, while prevalence has increased by 50% since the previous estimate (Parkinson's Foundation, 2022). Earlier detection of PD has been shown to greatly improve quality of life for those diagnosed (Naranjo et al., 2021). However, diagnosis can take years and multiple interventions, a process that may be unaffordable or unattainable for some. Thus, identifying neurodegenerative biomarkers to improve diagnostics and treatments for PD before significant cognitive and motor decline occurs is extremely beneficial (Hlavnička et al., 2017). Prior research conducted has established that early-stage PD biomarkers can be assessed through investigating functional changes in those with rapid eye movement sleep behavior disorder (RBD). RBD is a parasomnia condition where arousals from sleep and abnormal behaviors, movements, or emotions are experienced (Valli et al., 2022). RBD is considered to be a prodromal syndrome for PD, as those diagnosed with the disorder have greater than 80% likelihood of developing PD or another neurodegenerative disorder in their lifetime (St Louis et al., 2017). Given the predictive potential of RBD as a disease indicator for PD, testing diagnostic processes on RBD populations could assess their potential to detect early neurodegeneration.

An area of biomarker research that needs expanding upon is identifying how vocal change patterns manifest in PD. Speech is a motor skill that is very sensitive to deterioration of the basal ganglia (the primary neurological structure associated with PD), which occurs throughout disease progression (Hlavnička et al., 2017). Speech abnormalities are extremely common in PD pathology, with 90% of PD patients experiencing some degree of impairments (Moro-Velazquez et al., 2021). The prevalence of these indicators makes it essential to use them to detect patterns of neurodegeneration early in order to take preventative measures. Using vocal assessments as a biomarker has advantages over other analysis methods such as imaging, given it is inexpensive, non-invasive, and applicable to large groups.

My research is based on a 2017 study (Hlavnička et al., 2017) which sought to determine if a fully automated speech examination system could reliably distinguish patterns of neurodegeneration using acoustic vocal data. In this study, speech recordings of participants were pre-processed and analyzed. These researchers concluded that subliminal speech abnormalities are detectable in both PD and RBD participants, which have clinical applications for preclinical PD detection and thus could revolutionize how neurodegenerative diseases are diagnosed. In my study, my research questions are:

1. Are there specific speech variables that are strongly associated with early biomarkers of PD as identified by the automated speech analysis exam?
2. Is a reduced model comprised of these variables able to classify disease presence more accurately than the full model used in the original study?

Methods

Data was collected at Charles University in Prague, at the Department of Neurology and Centre of Clinical Neuroscience, Prague, Czech Republic, from 2014 to 2016. A total of 130 participants were used: 50 healthy controls, 30 recently diagnosed with PD, and 50 diagnosed with RBD. Healthy controls had no prior history of neurological or communication disorders or conditions. PD subjects were recruited for the study prior to any treatment regimen being initialized and were diagnosed upon the Parkinson's Disease Society Bank Criteria. RBD subjects were diagnosed with idiopathic RBD according to International Classification of Sleep Disorders criteria. Females were underrepresented (n = 27) in the sample, although this gender disparity present is consistent with previously recorded gender-based prevalence rates for both PD and RBD (Appendix 1). PD participants were graded on the Hoen & Yahr scale to determine the stage of Parkinson's disease progression. PD and RBD subjects were additionally measured using the Unified Parkinson Disease Rating Scale (UPDRS) for motor examination (III). The UPDRS, similar to the Hoen & Yahr scale, is utilized by clinicians to determine the severity and

progression of Parkinson's disease. The speech item from the UPDRS III, is used to examine the severity of dysarthria in participants. The mean of the speech item for PD and RBD subjects is <1 (Appendix 1). This is important to note as it demonstrates how the automated speech analysis exam is able to detect extremely subtle abnormalities/dysfunctions in speech patterns.

An automated speech analysis examination was conducted, wherein subjects performed a speaking task meant to demonstrate dysfunction in speech. Subjects were asked to read a phonetic text consisting of 80 words twice (values for these were averaged). These acoustic vocal assessments were recorded in a quiet room with low ambient noise using a head-mounted condenser microphone situated 5 cm from the mouth. Dimensions and parameters were set to evaluate each vocal assessment using 12 acoustic features as variables.

Hlavnička et al. developed a collection of 12 acoustic variables tailored to display speech disturbances in PD for the purpose of assessing the four basic subsystems of connected speech: phonation, articulation, timing, and respiration. Phonatory characteristics were assessed by duration of voiced intervals (ms) and gaping in-between voiced intervals (-/min). Articulation was measured by duration of unvoiced stops (ms) and decay of unvoiced fricatives (%/min). Timing features were evaluated by rate of speech timing (-/min), acceleration of speech timing (-/min2), duration of pause intervals (ms), and entropy of speech timing (-/min). Respiration was assessed by relative loudness of respiration (dB), pause intervals per respiration (-), rate of speech respiration (-/min), and latency of respiratory exchange (ms). For the purpose of these analyses, PD and RBD participants were both coded with a value of 1, corresponding to disease presence. Healthy controls took a value of 0, corresponding to disease absence.

To evaluate my research questions a random forest model was applied to the data to determine which, if any, speech feature variables were more strongly associated with the presence of neurodegeneration in an individual. Optimal tuning parameters were established, including the number of decision trees used in the forest and the number of random variables used in each decision tree. The dataset was then applied to the model wherein the importance of each speech feature variable was examined using mean decrease in Gini (MDG) values. Speech features with high importance (MDG ≥ 4) were used in the reduced model. Logistic regression analyses were conducted, wherein a likelihood ratio test was used to determine if the reduced model performed better than the full model. Deviance, receiver operating characteristic curves (ROC), and Akaike's Information Criteria (AIC) values were used as further metrics in which to compare model accuracy.

Results

The random forest model contains 500 decision trees, with 3 random variables tried in each split (OOB error rate = 25.38%). The model was used to examine variable importance using the calculated values of mean decrease in Gini. Speech feature variables with mean decrease in Gini values above 4 were included in the reduced model. Duration of unvoiced stops (ms) and pause intervals per respiration (-) had mean decrease in Gini values of 3.5512 and 2.6866, respectively (Table 1). Consequentially, these variables were removed from the model. The remaining high importance variables were used to generate the reduced model. The reduced model had a residual deviance of 124.10, while the full model had a residual deviance of 127.54, indicating that the reduced model was indeed more robust (Appendix 2). A likelihood

Speech Feature	Mean decrease in Gini
Duration of pause intervals (ms)	7.435195
Duration of voiced intervals (ms)	7.426821
Relative loudness of respiration (dB)	6.327564
Rate of speech timing (-/min)	5.404944
Acceleration of speech timing (-/min2)	5.301965
Rate of speech respiration (-/min)	4.778706
Decay of unvoiced fricatives (‰/min)	4.656286
Entropy of speech timing (-)	4.56803
Latency of respiratory exchange (ms)	4.542912
Gaping in-between voiced intervals (-/min)	4.416854
Duration of unvoiced stops (ms)	3.551194
Pause intervals per respiration (-)	2.686637

Legend	
Phonation	
Articulation	
Timing	
Respiration	

Table 1. Mean decrease in Gini values for 12 acoustic speech variables.

ratio test was then performed to compare model fits. The likelihood ratio chi-square test was not significant, with a p-value of 0.179 (Appendix 2). When examining the resulting ROC curves for each model, we can observe that the full model (AUC = 0.8275, 82.75%) performed slightly better than the reduced model (AUC = 0.8210, 82.1%) (Appendix 3-4).

Discussion

The findings of this study support the previously conceived notion that an automated speech analysis examination is a viable method for detecting early biomarkers of PD. Based on a thorough analysis of the combined outcomes obtained from the random forest model and the comparative evaluation of logistic regression models, it appears that the reduced model of high importance speech features can serve as a more robust model for detecting neurodegeneration in an individual. A reduced model provides improved efficiency which can be significantly beneficial in resource-constrained environments, such as mobile devices with limited processing capabilities. Considering that the overarching goal of the automated speech analysis examination is to develop a widely accessible and remotely administered exam, a model requiring fewer computational resources is preferred. Additionally, findings revealed that phonation and timing are the subsystems of connected speech that are most disturbed by PD, thus indicating they hold the highest predictive power for disease presence.

It is also important to address potential limitations or shortcomings that can be improved upon in future studies. The underrepresentation of women in the sample of my study should be considered a limitation. Due to sexual dimorphisms between men and women, disease progression and symptoms greatly differ (Cerri et al., 2019). Unfortunately, the model may not provide an accurate representation on gender differences in predicted patterns of neurodegeneration, as the sample size for women was significantly smaller than the sample size of men. It would be beneficial to reproduce the study with a more representative sample, to determine if the examination itself is an effective diagnostic tool for both men and women.

Future research would benefit from conducting the automated speech analysis examinations in a non-laboratory setting. The initial study has high potential for suffering from low-ecological validity, as each subject recorded the reading of the passage in a sound-proofed room with low ambient noise equipped with a professional microphone situated exactly 5 cm from the mouth. Further research should be conducted in non-laboratory settings, wherein participants complete the examination from their homes using a smartphone (or other household devices) to determine if the results of the initial study are applicable and can be generalized in a real-world setting. Additionally, the dataset used for this study was procured from the initial study conducted in Prague, Czech Republic. All participants in the study were native Czech speakers, and the reading passages were spoken in Czech. Prior literature has established that speech characteristics and natural rhythms differ between languages, essentially meaning that the specific speech dimensions used in the initial study would not produce viable results when applied in other languages (Peter et al., 2022). Thus, future research would benefit from a rescaling of these dimensions in order to represent the common speech patterns of healthy speakers to accurately predict early patterns of neurodegeneration in other languages.

Overall, findings suggest that the automated speech analysis examination has the potential to revolutionize the diagnostic process of PD through identifying individuals with early biomarkers far before healthcare professionals are capable to, thereby improving patient outcome and quality of life.

References

Ball, N., Teo, W. P., Chandra, S., & Chapman, J. (2019). Parkinson's Disease and the Environment. *Frontiers in neurology*, *10*, 218. https://doi.org/10.3389/fneur.2019.00218

Cerri, S., Mus, L., & Blandini, F. (2019). Parkinson's Disease in Women and Men: What's the Difference?. *Journal of Parkinson's disease*, *9*(3), 501–515. https://doi.org/10.3233/JPD-191683

Hlavnička, J., Čmejla, R., Tykalová, T. *et al.* Automated analysis of connected speech reveals early biomarkers of Parkinson's disease in patients with rapid eye movement sleep behaviour disorder. *Sci Rep* 7, 12 (2017). https://doi.org/10.1038/s41598-017-00047-5

Moro-Velazquez, L., Gomez-Garcia, J.A., Arias-Londoño, J.D., Dehak, N., Godino-Llorente, J.I. Advances in Parkinson's Disease detection and assessment using voice and speech: A review of the articulatory and phonatory aspects. *Biomedical Signal Processing and Control* 66, 102418(2021):1746-8094, https://doi.org/10.1016/j.bspc.2021.102418.

Naranjo, L., Pérez, J.C, Campos-Roca, Y. Monitoring Parkinson's disease progression based on recorded speech with missing ordinal responses and replicated covariates. *Computers in Biology and Medicine* 134, 104503(2021): 0010-4825, https://doi.org/10.1016/j.compbiomed.2021.104503

Number of People Diagnosed with Parkinson's Increases to Nearly 90,000 Each Year. Parkinson's Foundation. (2022, December 15). https://www.parkinson.org/understanding-parkinsons/statistics

Peter, V., van Ommen, S., Kalashnikova, M. *et al.* Language specificity in cortical tracking of speech rhythm at the mora, syllable, and foot levels. *Sci Rep* 12, 13477 (2022). https://doi.org/10.1038/s41598-022-17401-x

St Louis, E.K., Boeve, A.R. and Boeve, B.F. (2017), REM Sleep Behavior Disorder in Parkinson's Disease and Other Synucleinopathies. Mov Disord., 32: 645-658. https://doi.org/10.1002/mds.27018

Valli M, Uribe C, Mihaescu A, Strafella AP. Neuroimaging of rapid eye movement sleep behavior disorder and its relation to Parkinson's disease. J Neurosci Res. 2022 Oct;100(10):1815-1833. doi: 10.1002/jnr.25099. Epub 2022 Jul 4. PMID: 35790021.

Appendix

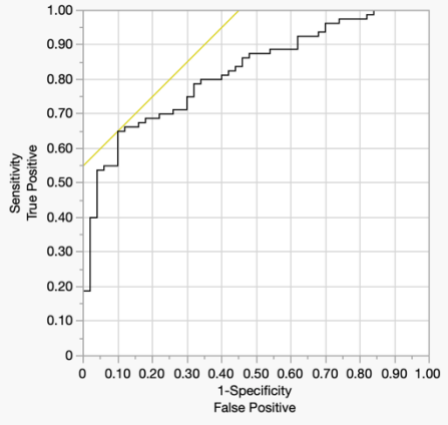
Appendix 1: Descriptive Statistics

	PD (n = 30)	RBD (n = 50)	HC (n = 50)
Men	70% (n = 21)	82% (n = 41)	82% (n = 41)
Women	30% (n = 9)	18% (n = 9)	18% (n = 9)
Mean age	64.9 (SD 10.9, range 34 - 79)	64.9 (SD 9.1, range 40-83)	62.1 (SD 10.8, range 40 - 83)
Mean age of disease onset (years)	63.4 (SD 11.9, range 30–78)	59.2 (SD 9.8, range 33–81)	n/a
Mean Hoehn & Yahr score	2.1 (SD 0.3, range 1.5–2.5)	n/a	n/a
Mean UPDRS III Total	20.2 (SD 12.4, range 6–54)	5.2 (SD 4.1, range 0–21)	n/a
Mean UPDRS III 18 Speech item	0.4 (SD 0.5, range 0–1)	0.06 (SD 0.24, range 0–1)	n/a

Appendix 2: Analysis of Deviance Table

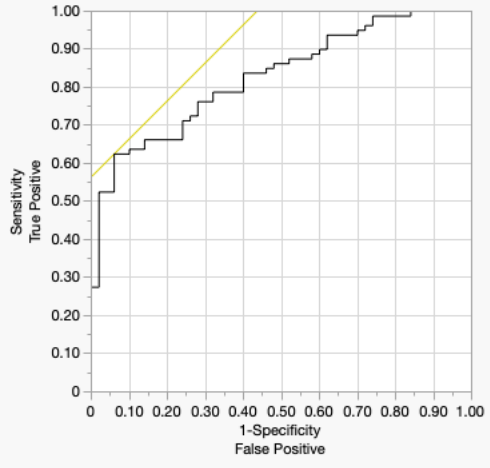
	Residual Df	Residual Deviance	Df	Deviance	Pr(>Chi)
Reduced Model	119	127.54			
Full Model	117	124.1	2	3.441	0.179

Appendix 3: ROC for Reduced Model



AUC = 0.8210

Appendix 4: ROC for Full Model



AUC = 0.8275