IDENTIFICATION OF PATIENTS AT THE RISK OF LEWY BODY DISEASES BASED ON ACOUSTIC ANALYSIS OF SPEECH

Andrea Fernández Martínez

Bachelor Degree Programme (4th), Faculty of Sciences and Technology, University of Vic – Central University of Catalonia

E-mail: andrea.fernandez@uvic.cat

Supervised by: Jiri Mekyska

E-mail: mekyska@feec.vutbr.cz

Abstract: Lewy body diseases (LBDs) is a group of neurodegenerative diseases that consists of Parkinson's disease and dementia with Lewy bodies, that are generally very crucial to be diagnosed in their prodromal state. In the frame of this study we proposed a multivariate logistic regression model that identifies people in a high risk of LBDs based on their articulatory and prosodic characteristics. More specifically, the model has 80 % specificity and 85 % sensitivity based on quantification of rigidity of tongue/jaw, monoloudness, and inappropriate pausing.

Keywords: Lewy body diseases, Parkinson's disease, dementia with Lewy bodies, speech, voice, acoustic analysis, prodromal diagnosis

1 INTRODUCTION

Lewy body diseases (LBDs) is a group of neurodegenerative diseases that consists of Parkinson's disease (PD) and dementia with Lewy bodies (DLB). LBDs are associated with pathophysiological process of α-synuclein accumulation in specific brain regions leading to the formation of Lewy bodies and Lewy neuritis resulting to cell death [1]. LBDs have a long prodromal interval, i.e. a period during which neurodegenerative symptoms are present, but full clinical disease has not yet developed [2]. Identification of this stage of LBDs is crucial for development of disease-modifying treatment, since the neurodegeneration may be possibly stopped or treated before the pathological cascades start.

Prodromal markers of LBDs are diverse and usually non-specific (except idiopathic REM sleep behaviour disorder) [1]. Nevertheless, a few studies suggest, that speech/voice disorders such as dysfluency, aperiodicity or irregular alternating motion rate can be identified in early stages of PD or DLB [3][4][5]. Based on these findings, the aim of this pilot study is to identify a group of people, who are at the risk of LBDs (i.e. in probable prodromal stage) and identify acoustic features that discriminate them from healthy controls or patients with clinically diagnosed PD.

2 MATERIALS AND METHODS

2.1 DATASET

We enrolled 20 Czech native PD patients (5 females, 15 males), 32 healthy controls HC (22 females, 10 males), and 24 people who were in the risk of LBDs (14 females, 10 males) at the First Department of Neurology, St. Anne's University Hospital in Brno, Czech Republic. The participants in the risk of LBDs were identified based on a screening questionnaire containing several risk factors, e.g. the REM sleep behaviour disorder. None of the PD patients had a disease affecting the central nervous system other than PD. These patients were examined on their regular dopaminergic

medication approximately 1 hour after the L-dopa dose. All participants signed an informed consent form that has been approved by the local ethics committee.

2.2 ACOUSTIC ANALYSIS

Speech/voice of the enrolled participants was recorded by a large capsule cardioid microphone M-AUDIO Nova and sampled at f_s = 16 kHz. More specifically, we acquired the following tasks: TSK1 – a monolog, at least 90 s long without interruption of a clinician; TSK2 – reading a short phonetically balanced paragraph; TSK3 – approximately 3 s (not longer than 5 s) sustained vowel /a/ at a comfortable pitch and loudness; TSK4 – approximately 3 s (not longer than 5 s) sustained vowel /i/ at a comfortable pitch and loudness; TSK5 – approximately 3 s (not longer than 5 s) sustained vowel /u/ at a comfortable pitch and loudness; TSK6 – sustained phonation of /a/ at a comfortable pitch and loudness as constant and long as possible (performed on one breath).

We quantified the following speech/voice disorders: 1) airflow insufficiency (using maximum phonation time MPT in TSK6); 2) irregular pitch fluctuations (relative standard deviation of fundamental frequency relF0SD, TSK3–6); 3) microperturbations in frequency (jitter, TSK3–6); 4) microperturbations in amplitude (shimmer, TSK3–6); 5) increased noise (harmonics-to-noise ratio HNR, TSK3–6); 6) aperiodicity (degree of unvoiced segments DUV, TSK3–6); 7) tremor of jaw (relative standard deviation of first (F1) and second (F2) formant relF1SD, relF2SD, TSK3–6); 8) decreased tongue movement (vowel articulation index VAI, TSK1–5); 9) rigidity of tongue and jaw (relF1SD, relF2SD, TSK1–2); 10) monoloudness (relative standard deviation of SEO relSEOSD, TSK1–2); 11) monopitch (relF0SD, TSK1–2); 12) inappropriate silences (speech index of rhythmicity SPIR, TSK2); 13) higher proportion of silence time (percentual pause ratio PPR, TSK2); 14) longer duration of silences (median duration of silences longer than 50 ms DurMED, TSK2); 15) higher variability of silence duration (median absolute deviation of silences longer than 50 ms, TSK2); 16) unnatural speech rate (articulation rate AR, TSK2).

2.3 STATISTICAL ANALYSIS

In the first step we employed univariate logistic regression to evaluate discrimination power (PD vs. HC) of individual acoustic features. The discrimination power was quantified using the area under curve (AUC). In addition, we used the minimum redundancy maximum relevance (mRMR) filtering feature selection technique to sort the features based on their relevance and non-redundancy.

In the second step, we selected the first five most discriminative and relevant features, that were further fed into a multivariate logistic regression model. Based on a visual inspection of resulting ROC (receiver operating characteristic) curves, we selected optimal combination of the features, that provided at least 80 % specificity and sensitivity. Finally, we used the multivariate model with the optimal threshold to identify people in a high risk of LBDs (we done this group as LBD) and compared the selected acoustic features of these people with corresponding mean values of the PD and HC cohorts.

3 RESULTS

Results of the univariate regression analysis are summarized in Table 1. According to the mRMR filtering technique, the most relevant feature was found to be the relative standard deviation of the 1st formant (TSK1) with the highest AUC = 75.97 %.

Regarding the multivariate regression analysis, the optimal logistic regression model was found combining features relF1SD (TSK1), SPIR (TSK2) and relSEOSD (TSK2): AUC = 87.10 % (see ROC in Figure 1). Using this model (with 80 % specificity and 85 % sensitivity), only seven participants were confirmed to be in the high risk of LBDs.

Finally, Figure 1 displays mean/individual values of the three above-mentioned features for PD, HC, and LBD groups.

Table 1: Results of the univariate regression analysis.

Task	Task category	Speech dimension	Acoustic feature	Mean ± STD (HC)	Mean ± STD (PD)	AUC [%]	mRMR importance
TSK1	monologue	articulation	relF1SD	0.59 ± 0.18	0.69 ± 0.12	75.97	1
TSK2	reading	speech fluency	SPIR	0.04 ± 0.01	0.03 ± 0.01	73.71	10
TSK2	reading	articulation	relF1SD	0.50 ± 0.13	0.61 ± 0.13	73.23	7
GLOBAL	sustained phonation	phonation	VAI	1.20 ± 0.24	1.10 ± 0.17	71.45	18
TSK1	monologue	prosody	relSEOSD	1.70 ± 0.81	1.29 ± 0.47	66.13	3

STD – standard deviation, HC – healthy controls, PD – Parkinson's disease patients, AUC – area under curve, relF1SD – relative std of the 1st formant, SPIR – speech index of rhythmicity, VAI – vowel articulation index, relSEOSD – relative std of squared energy operator

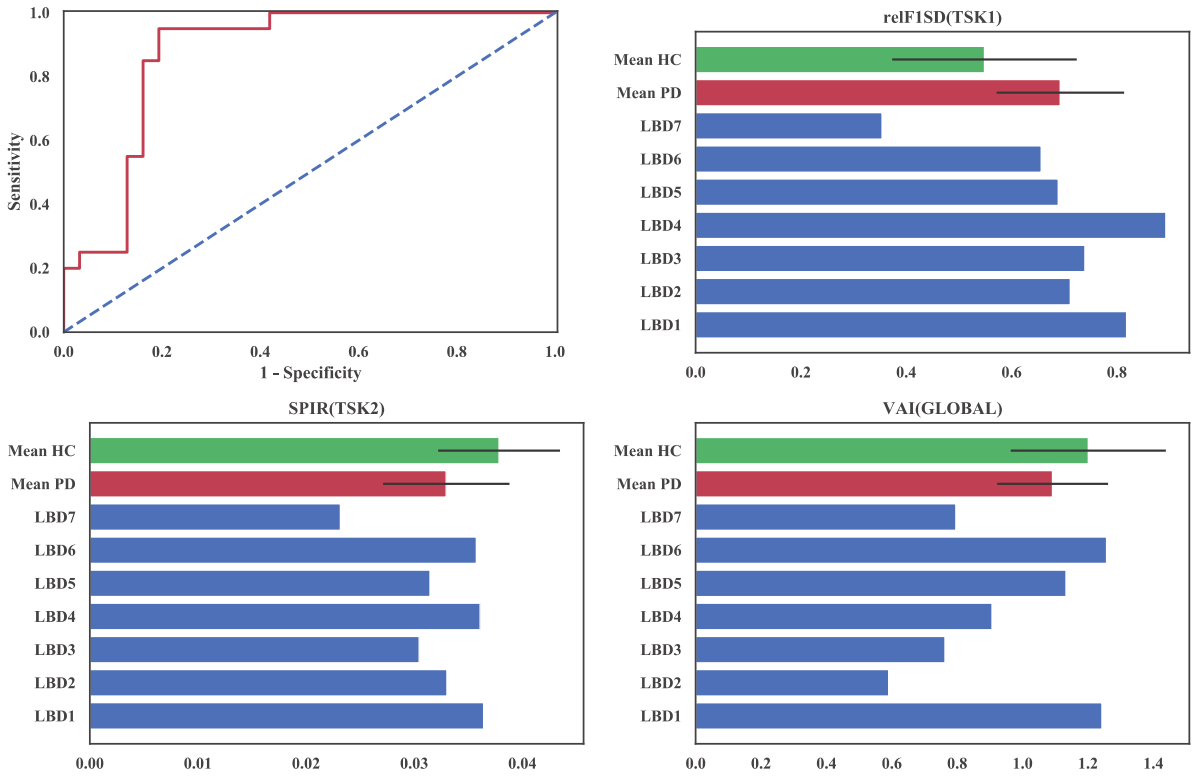

Figure 1: ROC (top-left figure) and values of the selected features (blue colour – people in the high risk of LBDs, green colour – mean value for the HC cohort, red colour – mean value for the PD cohort, relF1SD – relative std of the 1st formant, SPIR – speech index of rhythmicity, VAI – vowel articulation index).

4 DISCUSSION

The most discriminative feature is based on the first formant frequency extracted from the monologue. Generally, formants are related to the resonances of the oro-naso-pharyngeal tract and are modified by position of tongue and jaw. More specifically, the first formant is modified by the vertical position of tongue and jaw. Based on this finding we can conclude that the highest identified difference between PD and HC is in articulation.

The optimal combination of features, that provides 80 % specificity and 85 % sensitivity, contains (beside relF1SD (TSK1)) SPIR (TSK2) and relSEOSD (TSK2). The important role of SPIR in the acoustic analysis of hypokinetic dysathria was identified by Rektorova et al., who used this feature to predict mild cognitive impairment or dementia in PD patients [6]. Therefore we assume, that this

prosodic feature quantifying inappropriate silences is somehow associated with cognitive decline in the PD patients. Finally, relSEOSD assess monoloudness, which is again very typical for PD patients [5].

Based on the trained multivariate model, we finally selected 7 people in a high risk of LBDs. We observed that these people have generally higher values of relF1SD (in monologue and reading task) than HC, lower values of SPIR, and lower values of vowel articulation index, which reflects decreased tongue movement.

5 CONCLUSION

In this study, we identified acoustic features that discriminate PD and HC (with 80 % specificity and 85 % sensitivity), and used these parameters to train a logistic regression model, that identified people with a high risk of LBDs. These subjects, in comparison to HC, are associated mainly with articulatory and prosodic disorders.

This work has several limitations such as the small cohort of participants, gender inequality, and different severity of PD patients. Therefore, we cannot generalize the results, but rather consider them as pilot ones. On the other hand, to the best of our knowledge, it is the first work dealing with the identification of prodromal state of LBDs based on the acoustic analysis of speech/voice, and we believe that our findings will help in further research in this field of science.

ACKNOWLEDGEMENT

This research was funded by the Erasmus+ program and grant LO1401. For the research, infrastructure of the SIX Center was used.

REFERENCES

[1] POSTUMA, Ronald B, Alex IRANZO, Michele HU, et al. Risk and predictors of dementia and parkinsonism in idiopathic REM sleep behaviour disorder: a multicentre study. Brain. 2019, 142(3), 744–759. DOI: 10.1093/brain/awz030. ISSN 0006–8950.

[2] BERG, Daniela, Ronald B. POSTUMA, Charles H. ADLER, et al. MDS research criteria for prodromal Parkinson's disease. Movement Disorders. 2015, 30(12), 1600–1611. DOI: 10.1002/mds.26431. ISSN 08853185.

[3] ASH, Sharon, Corey MCMILLAN, Rachel G. GROSS, et al. Impairments of speech fluency in Lewy body spectrum disorder. Brain and Language. 2012, 120(3), 290–302. DOI: 10.1016/j.bandl.2011.09.004. ISSN 0093934X.

[4] RUSZ, Jan, Jan HLAVNIČKA, Tereza TYKALOVÁ, Jitka BUŠKOVÁ, Olga ULMANOVÁ, Evžen RŮŽIČKA and Karel ŠONKA. Quantitative assessment of motor speech abnormalities in idiopathic rapid eye movement sleep behaviour disorder. Sleep Medicine. 2016, 19, 141–147. DOI: 10.1016/j.sleep.2015.07.030. ISSN 13899457.

[5] BRABENEC, L., J. MEKYSKA, Z. GALAZ and Irena REKTOROVA. Speech disorders in Parkinson’s disease: early diagnostics and effects of medication and brain stimulation. Journal of Neural Transmission. 2017, 124(3), 303–334. DOI: 10.1007/s00702-017-1676-0. ISSN 0300–9564.

[6] REKTOROVA, Irena, Jiri MEKYSKA, Eva JANOUSOVA, et al. Speech prosody impairment predicts cognitive decline in Parkinson’s disease. Parkinsonism & Related Disorders. 2016, 29, 90–95. DOI: 10.1016/j.parkreldis.2016.05.018. ISSN 13538020.