Indirect Assessment of Hyperechogenicity of Substantia Nigra Utilizing Sleep-based Biomarkers

Marek Mikulec
Dept. of Telecommunications
FEEC, Brno University of Technology
Brno, Czech Republic
0000-0002-9405-9796

Jiří Mekyska
Dept. of Telecommunications
FEEC, Brno University of Technology
Brno, Czech Republic
0000-0002-6195-193X

Zoltán Galáž
Dept. of Telecommunications
FEEC, Brno University of Technology
Brno, Czech Republic
0000-0002-8978-351X

Luboš Brabenec
Applied Neuroscience Research Group
CEITEC, Masaryk University
Brno, Czech Republic
0000-0002-8348-5757

Kateřina Bočková
Applied Neuroscience Research Group
CEITEC, Masaryk University
Brno, Czech Republic
0000-0001-9637-4856

Irena Rektorová
International Clinical Research Center
and First Department of Neurology
FM, Masaryk University
St. Anne's University Hospital
Applied Neuroscience Research Group
CEITEC, Masaryk University
Brno, Czech Republic
0000-0002-5455-4573

***Abstract*—Transcranial sonography of the substantia nigra (TCS-SN) may serve as a suitable test for screening groups at a high risk of developing Lewy body diseases (LBDs) such as Parkinson's disease or dementia with Lewy bodies. Although one of the most prominent early markers of these neurodegenerative disorders is the idiopathic rapid eye movement (REM) sleep behavior disorder, the relationship between TCS-SN and sleep alterations has not been fully explored. The aim of this study is to investigate whether sleep-based biomarkers could be used to stratify subjects into three groups with different echogenic areas of the substantia nigra. To achieve this goal, we enrolled 93 participants who underwent TCS-SN and 7-night actigraphy. Additionally, participants completed a sleep diary and the REM sleep behavior disorder screening questionnaire. To assess the severity of pathological echogenicity, we employed a machine learning approach utilizing the XGBoost algorithm. The results show that a multimodal assessment of sleep was able to predict the outcomes of TCS-SN with a balanced accuracy of 96 %. Overall, these findings underscore the potential of a comprehensive approach to model the results of TCS-SN and its implications for the prodromal diagnosis of LBDs.**

***Index Terms*—actigraphy, Lewy body diseases, REM sleep behavior disorder screening questionnaires, sleep diary, substantia nigra, transcranial sonography**

I. Introduction

Lewy body diseases (LBDs) represent a group of neurodegenerative disorders characterized by the accumulation of Lewy bodies within neurons, notably in the cerebral cortex and basal ganglia, including the substantia nigra. These accumulations contribute to the degeneration and eventual death of dopaminergic cells [1]–[3]. LBDs encompass Parkinson's disease (PD), Parkinson's disease dementia (PDD), and dementia with Lewy bodies (DLB) [3]. Unfortunately, as of now, there is no cure for LBDs. However, early diagnosis and tailored treatment can help mitigate disease progression and enhance the patients quality of life [1], [3], [4].

Idiopathic rapid eye movement (REM) sleep behavioral disorder (iRBD) emerges as an early indicator of LBDs. iRBD is characterized by sudden, often vigorous, movements during REM sleep, driven by the content of vivid dreams [5]. This disorder is prevalent in over 80 % of patients with LBDs, with a lower likelihood in PD [2]. E.g., the diagnosis of DLB based on iRBD can achieve a sensitivity of up to 90 % and specificity of 73 % [6]. Notably, in the early prodromal phase of DLB, iRBD manifests in 50 % of cases (evaluated according to NACC-USD version 3) [7], [8]. The mean interval between iRBD onset and conversion to a symptomatic neurodegenerative disorder has been estimated at 14±6 years [9], [10]. This interval offers a window for potential neuroprotective treatments aimed at slowing or halting progressive neuronal loss [9]–[11].

Transcranial sonography (TCS) is a non-invasive medical imaging technique used as an alternative diagnostic method to confirm clinical diagnoses [15]. TCS provides ultrasound images of brain tissues, including the substantia nigra (SN), a region crucial for movement and dopamine production, which may exhibit changes in neurodegenerative diseases in the prodromal stage [15], [16]. For example, increased echogenicity of the substantia nigra is frequently observed in TCS scans of PD patients [15], [17]. TCS of SN (TCS-SN) could potentially be used even in the prodromal state of LBDs. For instance, TCS-SN could predict PD or DLB in iRBD

This work was supported by grant no. NU20-04-00294 (Diagnostics of Lewy body diseases in prodromal stage based on multimodal data analysis) of the Czech Ministry of Health, and by project no. LX22NPO5107 (MEYS): Financed by European Union – Next Generation EU.

patients [18]. However, this prodromal screening has only 40 % sensitivity and 61 % specificity (in the case of PD) [19]. Furthermore, the quality of TCS result images is contingent on the presence of a suitable temporal bone window [17], thus this screening cannot be performed on all individuals.

Polysomnography (PSG), often referred to as the gold standard, is the conventional method for iRBD identification [4], [9]. Nevertheless, its drawback is that it involves specialized equipment found in professional sleep laboratories [12]. In contrast, actigraphy relies on a simpler wearable device known as an actigraph, which records data from a 3-axis accelerometer and can measure temperature, ambient light, and other signals outside of a clinic [9]. This makes actigraphy a more cost-effective prescreening tool for iRBD; however, it is not as accurate as PSG [13]. Another approach of iRBD screening is based on the REM Sleep Behavior Disorder Screening Questionnaire (RBD-SQ), which is an exploratory subjective screening tool consisting of ten questions targeting typical RBD symptomatology [9], [14].

Although some studies have reported a relation between sleep disorders, specifically iRBD, and pathological echogenicity of the SN in patients with LBDs, this field is still relatively young, and some knowledge gaps have not been bridged yet [18]. The aim of this work is to take a step further and explore whether sleep-based biomarkers could be used to stratify subjects into three groups with different echogenic areas of the SN. Such a technique would enable neuroscientists to estimate the results of TCS-SN even in subjects with an inappropriate temporal bone window.

II. Materials and Methods

A. Participants and TCS-SN

For the purpose of this study, we enrolled 93 participants who met the criteria for the temporal bone window. These subjects were stratified into three groups based on the results of TCS-SN:

1) **SN+ group ($N = 25$):** Subjects with pathological echogenicity, i.e., with an echogenic area (EA) $\geq 0.2\,\mathrm{cm}^2$; this cut-off value, established in previous studies [20], [21], provides 95 % sensitivity for diagnosing PD.
2) **nonHC group ($N = 21$):** This group included subjects with EA between 0.16 and $0.19\,\mathrm{cm}^2$; these subjects fell into a gray zone, making the diagnosis uncertain.
3) **HC group ($N = 47$):** Comprising subjects with an EA $\leq 0.15\,\mathrm{cm}^2$; this group is considered healthy based on the literature [15].

Demographic characteristics of this dataset can be found in Table I. All participants signed an informed consent form that was approved by the local ethics committee.

B. Assessment of Sleep

All participants underwent 7-day actigraphy. Subsequently, the actigraphy data were pre-processed in terms of sleep/wake window identification using an algorithm proposed in [22]. This method builds upon the heuristic technique by van Hees et al. [23], which is commonly employed in modern actigraphy tools like `scikit-digital-health` from Pfizer R&D [24]. Enhancements include generating 135 features for each 30-second time window, with one-third being temperature-based. These features serve as input for an XGBoost classifier, producing binary sleep/wake predictions for each 30-second interval [22].

TABLE I
Demographic Characteristics of the Participants

	Total	Men	Women	Age (mean±std) [yrs.]
SN+	25	15	10	70 ± 13
nonHC	21	10	11	65 ± 18
HC	47	14	33	64 ± 16

Next, we extracted nine features from the pre-processed data, as listed in Table II, with four features conforming to the National Sleep Foundation Standards [25]. Similar features were derived from a sleep diary, which was subjectively filled out by the participants. A total of 26 features were available for each night, with 611 successfully recorded nights resulting in a set of 14 053 records.

Finally, the participants completed the REM sleep behavior disorder screening questionnaire (RBD-SQ) [14]. In the subsequent modeling, we utilized its total score.

TABLE II
Sleep Features and Their Abbreviations

Sleep feature	Abbrevation
Time in bed	TIB
Sleep onset latency (norm)	SOL, SOL-N
Wake after sleep onset (norm)	WASO, WASO-N
Wake after sleep offset	WASF
Total sleep time	TST
Wake bouts	WB
Awakening >5 minutes (norm)	WB>5, WB>5-N
Sleep efficiency (norm)	SE, SE-N
Sleep fragmentation	SF

C. Experiments and Machine Learning

Sleep features and RBD-SQ results were used to predict the outcomes of TCS-SN. The dataset was initially divided into 80 % training and 20 % testing data. Various scenarios were considered:

- Stratification of participants:
  - **SN+ vs. HC:** Exclusively considered subjects with SN+ (EA $\geq 0.2\,\mathrm{cm}^2$) and healthy controls (EA $\leq 0.15\,\mathrm{cm}^2$), disregarding subjects in the gray nonHC zone.
  - **SN+ vs. NonHC + HC:** Categorized subjects based on EA $\geq 0.2\,\mathrm{cm}^2$ as the positive class and the rest as healthy.
  - **SN+ vs. NonHC:** Excluded subjects with EA $\leq 0.15\,\mathrm{cm}^2$, treating the NonHC group as the zero

class, focusing on distinguishing subjects with uncertain TCS-SN diagnosis.
  - **NonHC vs. HC:** Ignored SN+ subjects (EA $\geq$ $0.2\,\mathrm{cm}^2$), aiming to ascertain the distinguishability of HC and NonHC groups based on other features.
- Use of actigraphy, sleep diary, and RBD-SQ features:
  - **ACG:** Utilized actigraphy features, including those from Table II, along with features normalized based on [25].
  - **ACG-N:** Subset of ACG, focusing solely on the normalized actigraphy features from [25]: SOL-N, WASO-N, WB>5-N, and SE-N, while ignoring other actigraphy features.
  - **Diary:** Featured the same set of actigraphy-derived features, but sourced from the sleep diaries of subjects.
  - **Diary-N:** Paralleled ACG-N, with only normalized features from sleep diaries considered.
  - **RBD-SQ:** Scenarios relied on raw data from the RBD-SQ, using the final score (0–10) without the cut-off score of 5.

These scenarios were designed to determine which dataset is most suitable for predicting TCS-SN results and to evaluate the potential to differentiate subjects in the gray NonHC zone or distinguish subjects in the HC and SN+ groups based on other parameters.

The subsequent analysis was consistent across scenarios, involving data division into target groups for binary classification, SMOTE upsampling of the minority class, hyperparameter tuning for the XGBoost classifier, and 10-fold cross-validation [26], [27]. We used the following evaluation metrics: balanced accuracy (BACC), Matthew's correlation coefficient (MCC), sensitivity (SEN), specificity (SPE), precision (PRE), and the F1 Score (F1).

Next, the models underwent further analysis using feature importance and SHAP (SHapley Additive exPlanations) values. SHAP values are derived from game theory and quantify the contribution of each feature point to the model's final output [28]. Moreover, the best models underwent a permutation test to assess the significance of cross-validated scores through permutations, yielding p-values. A smaller p-value ($p < 0.05$) indicates a stronger dependency between a feature and the target estimates, while a larger p-value suggests a lack of real dependency and the model's inability to provide accurate predictions [29].

In the final step, the models were tested and evaluated on the remaining 20 % of the dataset. The testing considered the connection between nights recorded by subjects, implementing majority voting for decision-making to maintain the dataset's original imbalance, without introducing synthetic data. Various experiments were conducted, including fine-tuning the threshold for majority voting.

III. Results

The evaluation metrics for all scenarios are presented in Table III, demonstrating a color-coded scale for enhanced readability.

Regarding the SN+ vs. HC scenario, a model was able to discriminate both groups with 96 % sensitivity and specificity when analyzing all kinds of data, i.e., ACG, diary, and RBD-SQ.

The SN+ vs. NonHC scenario, utilizing ACG-N + Diary-N and RBD-SQ features, achieves the highest specificity (97.5 %) and precision (97.4 %). Incorporating all features increases sensitivity to 98.7 %, yet at the expense of specificity and precision.

In the context of the gray zone involving NonHC subjects, high accuracies (BACC = 96 %) were achieved in the NonHC vs. HC scenario, suggesting that both groups should be treated individually; i.e., they cannot be merged and assumed to be one HC group.

Regarding the feature importances, awakening > 5 minutes emerged as the most frequent and important feature (14 scenarios). RBD-SQ score dominated in 5 scenarios, while sleep onset latency and sleep efficiency appeared in 3 and 2 scenarios, respectively. Notably, the difference in feature importance between the most prominent feature and the others was generally minimal.

A. SHAP Values

The SHAP values for the SN+ vs. HC scenario (ACG-N + Diary-N + RBD-SQ) are illustrated in Fig. 1. Notably, WB > 5, derived from sleep diaries, emerged as the most influential feature in this scenario. High values indicate subjects being awake during the night more frequently. According to SHAP values, this fact is likely significant for individuals with SN+. Additionally, the plot suggests a connection between RBD and SN+, as manifested by vivid dreams causing awakening. Intriguingly, the plot shows that high sleep efficiency, as measured by actigraphy, can associate with SN+ subjects, contradicting sleep diary results. RBD-SQ as a third feature aligns with clinical understanding, with low scores indicating healthy controls who experience fewer sleep-related problems.

B. Permutation Test

In all scenarios, the p-values of the permutation test were lower than 0.05. This indicates a strong statistical dependency between features and the model's output.

C. Test With the Remaining 20 % of the Dataset

Classification results using the remaining 20 % of the dataset are detailed in Table IV. Various values of the day threshold in majority voting were tested. The table includes results for the best setting of this threshold for each scenario. The threshold values are listed in column N.

IV. Discussion

Our findings demonstrate the feasibility of distinguishing SN+ individuals from healthy controls with high accuracy using features such as ACG-N, Diary-N, and RBD-SQ. Importantly, we showed that the different echogenic areas of the SN should be divided into three groups, that are distinguishable

TABLE III
CLASSIFICATION RESULTS IN THE CROSS-VALIDATION STEP

Scenario	Selected Features	BACC	MCC	SEN	SPE	PRE	F1
SN+ vs. HC	ACG	0.839	0.679	0.850	0.829	0.832	0.841
SN+ vs. HC	Diary	0.911	0.822	0.931	0.891	0.895	0.912
SN+ vs. HC	ACG + RBD-SQ	0.902	0.803	0.898	0.905	0.904	0.901
SN+ vs. HC	Diary + RBD-SQ	0.938	0.878	0.975	0.902	0.908	0.940
SN+ vs. HC	ACG-N + Diary-N + RBD-SQ	0.960	0.920	0.960	0.960	0.960	0.960
SN+ vs. HC	ACG + Diary + RBD-SQ	0.933	0.867	0.964	0.902	0.907	0.935
SN+ vs. NonHC + HC	ACG	0.879	0.758	0.879	0.879	0.879	0.879
SN+ vs. NonHC + HC	Diary	0.927	0.856	0.960	0.894	0.901	0.929
SN+ vs. NonHC + HC	ACG + RBD-SQ	0.917	0.834	0.899	0.935	0.932	0.915
SN+ vs. NonHC + HC	Diary + RBD-SQ	0.953	0.907	0.955	0.952	0.952	0.954
SN+ vs. NonHC + HC	ACG-N + Diary-N + RBD-SQ	0.957	0.915	0.970	0.945	0.946	0.958
SN+ vs. NonHC + HC	ACG + Diary + RBD-SQ	0.956	0.912	0.960	0.952	0.953	0.956
SN+ vs. NonHC	ACG	0.854	0.709	0.842	0.867	0.864	0.853
SN+ vs. NonHC	Diary	0.940	0.880	0.949	0.930	0.932	0.940
SN+ vs. NonHC	ACG + RBD-SQ	0.946	0.893	0.956	0.937	0.938	0.947
SN+ vs. NonHC	Diary + RBD-SQ	0.937	0.874	0.949	0.924	0.926	0.938
SN+ vs. NonHC	ACG-N + Diary-N + RBD-SQ	0.956	0.912	0.937	0.975	0.974	0.955
SN+ vs. NonHC	ACG + Diary + RBD-SQ	0.956	0.913	0.987	0.924	0.929	0.957
NonHC vs. HC	ACG	0.885	0.770	0.872	0.898	0.895	0.884
NonHC vs. HC	Diary	0.918	0.837	0.949	0.887	0.894	0.920
NonHC vs. HC	ACG + RBD-SQ	0.953	0.905	0.945	0.960	0.959	0.952
NonHC vs. HC	Diary + RBD-SQ	0.960	0.920	0.978	0.942	0.944	0.961
NonHC vs. HC	ACG-N + Diary-N + RBD-SQ	0.945	0.891	0.938	0.953	0.952	0.945
NonHC vs. HC	ACG + Diary + RBD-SQ	0.943	0.887	0.938	0.949	0.948	0.943

[1] BACC – balanced accuracy, MCC – Matthew's correlation coefficient, SEN – sensitivity, SPE – specificity, PRE – precision, F1 – F1 score

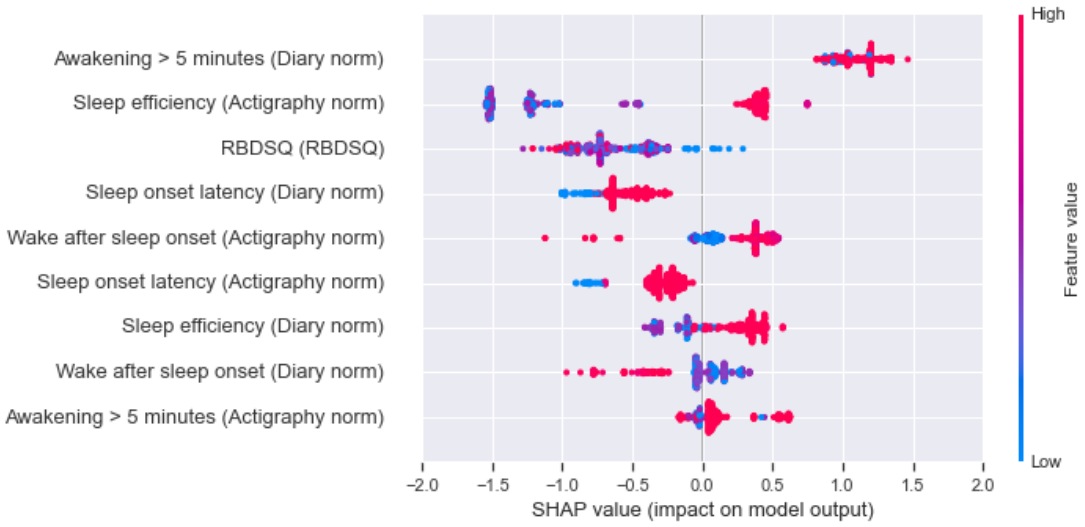

Fig. 1. SHAP values of scenario SN+ vs. HC – ACG-N+Diary-N+RBD-SQ

using sleep-based biomarkers. However, a lower ability to distinguish the NonHC vs. HC group was observed in the test with remaining 20 % of the dataset, where only 76 % specificity and 60 % precision were achieved. The results are consistent with literature [30], where division into three different groups based on echogenic areas of the SN for iRBD patients was suggested (HC, iRBD patients and PD patients).

The highest prediction accuracy was achieved when combining all three feature types. This indicates the complementary nature of actigraphy, sleep diary, and RBD-SQ in

TABLE IV
CLASSIFICATION RESULTS USING THE REMAINING 20 % OF THE DATASET

Scenario	Features	N	BACC	MCC	SEN	SPE	PRE	F1	TN	FP	FN	TP
SN+ vs. HC	ACG	2	0.718	0.375	1.000	0.436	0.323	0.488	34	44	0	21
SN+ vs. HC	Diary	4	0.769	0.514	0.667	0.872	0.583	0.622	68	10	7	14
SN+ vs. HC	ACG + RBD-SQ	6	0.622	0.286	0.333	0.910	0.500	0.400	71	7	14	7
SN+ vs. HC	Diary + RBD-SQ	4	0.769	0.514	0.667	0.872	0.583	0.622	68	10	7	14
SN+ vs. HC	ACG-N + Diary-N + RBD-SQ	3	0.679	0.302	0.667	0.692	0.368	0.475	54	24	7	14
SN+ vs. HC	ACG + Diary + RBD-SQ	6	0.622	0.286	0.333	0.910	0.500	0.400	71	7	14	7
SN+ vs. NonHC + HC	ACG	6	0.667	0.544	0.333	1.000	1.000	0.500	112	0	14	7
SN+ vs. NonHC + HC	Diary	4	0.811	0.648	0.667	0.955	0.737	0.700	107	5	7	14
SN+ vs. NonHC + HC	ACG + RBD-SQ	6	0.635	0.322	0.333	0.938	0.500	0.400	105	7	14	7
SN+ vs. NonHC + HC	Diary + RBD-SQ	6	0.635	0.322	0.333	0.938	0.500	0.400	105	7	14	7
SN+ vs. NonHC + HC	ACG-N + Diary-N + RBD-SQ	6	0.667	0.544	0.333	1.000	1.000	0.500	112	0	14	7
SN+ vs. NonHC + HC	ACG + Diary + RBD-SQ	6	0.635	0.322	0.333	0.938	0.500	0.400	105	7	14	7
SN+ vs. NonHC	ACG	6	0.667	0.486	0.333	1.000	1.000	0.500	34	0	14	7
SN+ vs. NonHC	Diary	4	0.745	0.495	0.667	0.824	0.700	0.683	28	6	7	14
SN+ vs. NonHC	ACG + RBD-SQ	4	0.564	0.142	0.333	0.794	0.500	0.400	27	7	14	7
SN+ vs. NonHC	Diary + RBD-SQ	5	0.730	0.461	0.667	0.794	0.667	0.667	27	7	7	14
SN+ vs. NonHC	ACG-N + Diary-N + RBD-SQ	4	0.730	0.461	0.667	0.794	0.667	0.667	27	7	7	14
SN+ vs. NonHC	ACG + Diary + RBD-SQ	3	0.627	0.248	0.667	0.588	0.500	0.571	20	14	7	14
NonHC vs. HC	ACG	4	0.790	0.540	0.824	0.756	0.596	0.691	59	19	6	28
NonHC vs. HC	Diary	5	0.494	-0.011	0.412	0.577	0.298	0.346	45	33	20	14
NonHC vs. HC	ACG + RBD-SQ	6	0.410	-0.250	0.000	0.821	0.000	0.000	64	14	34	0
NonHC vs. HC	Diary + RBD-SQ	5	0.552	0.101	0.412	0.692	0.368	0.389	54	24	20	14
NonHC vs. HC	ACG-N + Diary-N + RBD-SQ	2	0.552	0.101	0.412	0.692	0.368	0.389	54	24	20	14
NonHC vs. HC	ACG + Diary + RBD-SQ	6	0.455	-0.170	0.000	0.910	0.000	0.000	71	7	34	0

[1] N – a day threshold used in the majority voting, BACC – balanced accuracy, MCC – Matthew's correlation coefficient, SEN – sensitivity, SPE – specificity, PRE – precision, F1 – F1 score, TN – true negative, FP – false positive, FN – false negative, TP – true positive

modelling TCS results. Furthermore, our feature importance analysis suggests that no single direct biomarker is solely linked with the increased echogenicity of SN. Instead, a combination of factors contributes to its diagnosis.

The SHAP values shed light on the importance of specific features in distinguishing SN+ individuals. For instance, frequent awakenings during the night (WB $>$ 5) were strongly associated with SN+. Additionally, the connection between iRBD and SN+ is highlighted, suggesting that vivid dreams causing awakenings may indicate both iRBD and SN+. This underscores the value of a multidimensional approach to diagnosis.

Interestingly, models based solely on sleep diaries, without considering actigraphy features, yielded better results. This finding is supported by the SHAP values analysis, where sleep efficiency from actigraphy data contradicted the trend. On the contrary, actigraphy data proved to be the most relevant for distinguishing the NonHC vs. HC group in the test dataset.

In conclusion, our study provides insights into the potential use of ACG-N, Diary-N, and RBD-SQ features as alternatives to TCS-SN. While further research is needed to validate these findings on a larger and more diverse population, our results offer promising avenues for emulating TCS diagnosis. Additionally, the clinical utility of these predictive models should be assessed in real-world healthcare settings.

Finally, the authors in [31] state that the size of EA remains stable over time and, therefore, cannot be used to monitor the progress of neurodegeneration. In order to confirm or reject this finding and to evaluate the predictive value of TCS-SN and sleep-based features, we plan to repeat the data acquisition several times.

A. Study Limitations

Despite the promising results, our study has several limitations. First, the dataset used for this analysis is relatively small, which may limit the generalizability of our findings. Second, the cross-sectional nature of the data prevents us from establishing causal relationships. Longitudinal studies are needed to investigate the predictive value of these features over time. Finally, our analysis relies on self-reported sleep diary data, which could be susceptible to recall bias. Objective sleep monitoring using wearable devices can provide more accurate sleep-related features.

V. CONCLUSION

Utilizing a machine-learning-based approach, this study explored whether objectively and subjectively measured sleep alterations could predict the outcomes of TCS-SN. When differ-

entiating participants with normal and pathological echogenicity (in a cross-validation experiment), we achieved 96 % sensitivity and specificity while modeling data from actigraphy, sleep diary, and RBD-SQ. The performance measures decreased when evaluating the model on an independent test set; specifically, we obtained 67 % sensitivity and 87 % specificity.

Our findings suggest that combining multiple sources of data, including objective sleep measurements, subjective sleep diary information, and RBD-SQ scores, can support the estimation of echogenic area even in subjects with an inappropriate temporal bone window. In general, the combination of TCS-SN and sleep-based features could present a new alternative for screening LBDs in the prodromal state.

REFERENCES

[1] S. Chokroverty, "Overview of sleep & sleep disorders," Indian J Med Res, vol. 131, no. 2, pp. 126–140, 2010.

[2] A.-M. Tăuţan, B. Ionescu, and E. Santarnecchi, "Artificial intelligence in neurodegenerative diseases: A review of available tools with a focus on machine learning techniques," Artificial Intelligence in Medicine, vol. 117, p. 102081, Jul. 2021, doi: 10.1016/j.artmed.2021.102081.

[3] K. Menšíková et al., "Lewy body disease or diseases with Lewy bodies?," npj Parkinson's Disease, vol. 8, no. 1, p. 3, Jan. 2022, doi: 10.1038/s41531-021-00273-9.

[4] T. F. Outeiro et al., "Dementia with Lewy bodies: an update and outlook," Molecular Neurodegeneration, vol. 14, no. 1, p. 5, Jan. 2019, doi: 10.1186/s13024-019-0306-8.

[5] M. J. Sateia, "International Classification of Sleep Disorders-Third Edition," Chest, vol. 146, no. 5, pp. 1387–1394, Nov. 2014, doi: 10.1378/chest.14-0970.

[6] T. J. Ferman et al., "Inclusion of RBD improves the diagnostic classification of dementia with Lewy bodies," Neurology, vol. 77, no. 9, p. 875, Aug. 2011, doi: 10.1212/WNL.0b013e31822c9148.

[7] K. A. Wyman-Chick et al., "Prodromal Dementia With Lewy Bodies: Evolution of Symptoms and Predictors of Dementia Onset," J Geriatr Psychiatry Neurol, p. 08919887211023586, Jun. 2021, doi: 10.1177/08919887211023586.

[8] L. Besser et al., "Version 3 of the National Alzheimer's Coordinating Center's uniform data set.," Alzheimer disease and associated disorders, 2018.

[9] Š. Kristián et al., "Actigraphic Screening for Rapid Eye Movement Sleep Behavior Disorder," medRxiv, p. 19001867, Jan. 2019, doi: 10.1101/19001867.

[10] C. H. Schenck, B. F. Boeve, and M. W. Mahowald, "Delayed emergence of a parkinsonian disorder or dementia in 81% of older men initially diagnosed with idiopathic rapid eye movement sleep behavior disorder: a 16-year update on a previously reported series," Sleep medicine, vol. 14, no. 8, pp. 744–748, 2013.

[11] A. Iranzo et al., "Neurodegenerative disease status and post-mortem pathology in idiopathic rapid-eye-movement sleep behaviour disorder: an observational cohort study," The Lancet Neurology, vol. 12, no. 5, pp. 443–453, 2013.

[12] K. Crowley, "Sleep and Sleep Disorders in Older Adults," Neuropsychology Review, vol. 21, no. 1, pp. 41–53, Mar. 2011, doi: 10.1007/s11065-010-9154-6.

[13] H. Yuan, E. Hill, S. Kyle, and A. Doherty, "A systematic review of the performance of actigraphy in measuring sleep stages". 2023. doi: 10.1101/2023.08.01.23293507.

[14] J. Bušková et al., "Validation of the REM sleep behavior disorder screening questionnaire in the Czech population," Bmc Neurology, vol. 19, pp. 1–5, 2019.

[15] S. Mehnert, I. Reuter, K. Schepp, P. Maaser, E. Stolz, and M. Kaps, "Transcranial sonography for diagnosis of Parkinson's disease," BMC Neurology, vol. 10, no. 1, p. 9, Jan. 2010, doi: 10.1186/1471-2377-10-9.

[16] P. Damier, E. C. Hirsch, Y. Agid, and A. M. Graybiel, "The substantia nigra of the human brain: II. Patterns of loss of dopamine-containing neurons in Parkinson's disease," Brain, vol. 122, no. 8, pp. 1437–1448, Aug. 1999, doi: 10.1093/brain/122.8.1437.

[17] D. Školoudík et al., "Transcranial Sonography of the Substantia Nigra: Digital Image Analysis," Am. J. Neuroradiol., vol. 35, no. 12, p. 2273, Dec. 2014, doi: 10.3174/ajnr.A4049.

[18] M. Miyamoto and T. Miyamoto, "Relationship of substantia nigra hyperechogenicity to risk of Lewy body disease in idiopathic REM sleep behavior disorder patients: a longitudinal study," Sleep Medicine, vol. 68, pp. 31–34, Apr. 2020, doi: 10.1016/j.sleep.2019.09.008.

[19] Angela E P Bouwmans, Annemarie M M Vlaar, Werner H Mess, Alfons Kessels, and Wim E J Weber, "Specificity and sensitivity of transcranial sonography of the substantia nigra in the diagnosis of Parkinson's disease: prospective cohort study in 196 patients," BMJ Open, vol. 3, no. 4, p. e002613, Jan. 2013, doi: 10.1136/bmjopen-2013-002613.

[20] D. Berg et al., "Vulnerability of the nigrostriatal system as detected by transcranial ultrasound," Neurology, vol. 53, no. 5, p. 1026, Sep. 1999, doi: 10.1212/WNL.53.5.1026.

[21] U. Walter, M. Wittstock, R. Benecke, and D. Dressler, "Substantia nigra echogenicity is normal in non-extrapyramidal cerebral disorders but increased in Parkinson's disease," Journal of Neural Transmission, vol. 109, no. 2, pp. 191–196, Feb. 2002, doi: 10.1007/s007020200015.

[22] M. Mikulec et al., "Prodromal Diagnosis of Lewy Body Diseases Based on Actigraphy," in 2022 45th International Conference on Telecommunications and Signal Processing (TSP), Jul. 2022, pp. 403–406. doi: 10.1109/TSP55681.2022.9851316.

[23] V. T. van Hees et al., "A Novel, Open Access Method to Assess Sleep Duration Using a Wrist-Worn Accelerometer," PLOS ONE, vol. 10, no. 11, p. e0142533, Nov. 2015, doi: 10.1371/journal.pone.0142533.

[24] L. Adamowicz, Y. Christakis, M. Czech, and T. Adamusiak, "SciKit Digital Health: Python Package for Streamlined Wearable Inertial Sensor Data Processing," JMIR mhealth and uhealth, vol. 10, Apr. 2022, doi: 10.2196/36762.

[25] M. Ohayon et al., "National Sleep Foundation's sleep quality recommendations: first report," Sleep Health, vol. 3, no. 1, pp. 6–19, Feb. 2017, doi: 10.1016/j.sleh.2016.11.006.

[26] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," Journal of artificial intelligence research, vol. 16, pp. 321–357, 2002, doi: 10.1613/jair.953.

[27] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," presented at the Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, 2016, pp. 785–794. doi: 10.1145/2939672.2939785.

[28] W. E. Marcílio and D. M. Eler, "From explanations to feature selection: assessing SHAP values as feature selection mechanism," in 2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Nov. 2020, pp. 340–347. doi: 10.1109/SIBGRAPI51738.2020.00053.

[29] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.

[30] H. Y. Shin, E. Y. Joo, S. T. Kim, H.-J. Dhong, and J. W. Cho, "Comparison study of olfactory function and substantia nigra hyperechogenicity in idiopathic REM sleep behavior disorder, Parkinson's disease and normal control," Neurological Sciences, vol. 34, no. 6, pp. 935–940, Jun. 2013, doi: 10.1007/s10072-012-1164-0.

[31] A. Iranzo et al., "Five-year follow-up of substantia nigra echogenicity in idiopathic REM sleep behavior disorder," Movement Disorders, vol. 29, no. 14, pp. 1774–1780, Dec. 2014, doi: 10.1002/mds.26055.