Predicting Daytime Sleepiness from Electrocardiography Based Respiratory Rate Using Deep Learning

Emmi Antikainen[1], Rana Zia Ur Rehman[2], Teemu Ahmaniemi[1], Meenakshi Chatterjee[3]

[1] VTT Techinal Research Centre of Finland Ltd., Tampere, Finland
[2] Newcastle University, Newcastle Upon Tyne, United Kingdom
[3] Janssen Research & Development, Cambridge, MA, USA

Abstract

Daytime sleepiness impairs the activities of daily living, especially in chronic disease patients. Typically, daytime sleepiness is measured with subjective patient reported outcomes (PROs), which could be prone to recall bias. Objective measures of daytime sleepiness, which are sensitive to change, would benefit disease state assessment and novel therapies that impact the quality of life. The presented study aimed to predict daytime sleepiness from two hours of continuously measured respiratory rate using a 1-dimensional convolutional neural network. A wearable biosensor was used to continuously measure electrocardiography (ECG) based respiratory rate, while the participants (N=82) were asked to fill in Karolinska Sleepiness Scale three times a day. Considering the need for a sleepiness measure for chronic diseases, neurodegenerative disease (NDD, N=14) patients, immune-mediated inflammatory disease (IMID, N=42) patients, as well as healthy participants (N=26) were included in the study. The disease-agnostic model achieved an accuracy of 63% between non-sleepy and sleepy states. The result demonstrates the potential of using respiratory rate with deep learning for an objective measure of daytime sleepiness.

1. Introduction

Chronic disease patients commonly experience sleep disturbances and fatigue, which contribute to daytime sleepiness. Daytime sleepiness, in return, interferes with the activities of daily living, ultimately deteriorating the quality of life [1]. It affects cognitive functionalities, increasing the risk of falls resulting in injuries and increased healthcare costs [1, 2]. For instance, among Parkinson's Disease (PD) patients, over 35 % experience excessive daytime sleepiness [3]. Objective measurement of daytime sleepiness is important for both assessing new therapies and evaluating the effect of interventions.

Currently, daytime sleepiness is assessed with subjective patient reported outcomes (PROs), such as the Karolinska Sleepiness Scale (KSS) or the Epworth Sleepiness Scale. However, such subjective measures suffer from recall bias [4]. Objective measures of the physiological signs of sleepiness could provide better accuracy, reliability, and continuous assessment. Electrocardiography (ECG) based wearable sensors can facilitate the continuous monitoring of chronic disease patients in free-living settings and may capture how the disease affects the patient's daily-living.

Previous studies on patients' daytime sleepiness prediction have utilized clinical data or laboratory measurements [5, 6]. To our knowledge, sleepiness prediction with wearable sensors in free-living settings has not been studied extensively: Igasaki et al. predicted sleepiness from respiratory signals with support vector machines during a simulated drive, achieving an 89 % accuracy, whereas Bao et al. used wearable body temperature sensing to assess sleepiness over two days [7, 8]. Both studies only included a small sample (6-7) of healthy adults measured for a short time in simulated or restricted free-living settings.

One intuitive manifestation of daytime sleepiness in respiration is yawning. Previous studies have established that yawning frequency increases with sleepiness [9, 10]. This study uses deep learning (DL) to predict patient reported daytime sleepiness from continuously measured respiratory rate, which is often readily measured by modern wearable sensors and may offer an easily accessible continuous measure for sleepiness. The proposed disease-agnostic model uses a 1-dimensional convolutional neural network (1D CNN) and builds on a longitudinal multi-site data set, covering several days of respiratory rate and KSS responses (three times a day) collected from 82 volunteers, including neurodegenerative disease (NDD) patients (N=14), immune-mediated inflammatory disease (IMID) patients (N=42), and healthy participants (N=26).

2. Material and Methods

The study is based on the data collected in the IDEA-FAST feasibility study [11, 12].

2.1. Study participants

The study data comprised a total of 82 volunteered adults, including NDD patients (N=14), IMID patients (N=42), and healthy volunteers (N=26). The NDD group comprised patients with PD (N=8) and Huntington's Disease (HD, N=6). The IMID group included Inflammatory Bowel Disease (IBD, N=10), Primary Sjögren's Syndrome (PSS, N=13), Rheumatoid Arthritis (RA, N=7), and Systemic Lupus Erythematosus (SLE, N=12).

The participants were recruited at four sites, and the ethical approvals were granted (in June to November 2020) by the research ethics committees of each site: the ethical committee of the Medical Faculty of Kiel University (K) (D491/20), Newcastle upon Tyne Hospitals National Health Service Foundation/Newcastle University (N), Erasmus University Medical Centre in Rotterdam (E), and George-Huntington-Institute in Muenster (G). The study was registered in the German Clinical Trial Registry under DRKS00021693.

2.2. Study protocol

The participants wore a wearable patch sensor, VitalPatch, on their chest [13]. It adheres to the skin, and its battery lasts up to seven days. VitalPatch uses a single lead ECG (and partially a tri-axial accelerometer) to derive respiratory rate readings at 0.25 Hz. It is a class IIa medical device with FDA clearance. The participants wore the biosensor for five consecutive days at a time. The wear period was repeated up to four times during their enrollment and was always followed by at least two rest days.

The PRO for daytime sleepiness was collected using the KSS, which was prompted three times a day (at 13:00, 17:00, and 21:00 local time) via a smartphone application, the VTT Stress Monitor App [14]. The KSS was available for a response for 3 hours in the early and late afternoon and 2.5 hours in the evening. The response was selected from a drop-down menu list of ten options, ranging from "extremely alert" to "extremely sleepy".

2.3. Data pre-processing

The respiratory rate data were pre-processed by sorting the timestamps into monotonically increasing order while removing duplicates and by removing (a) manufacturer-defined invalid values, (b) values beyond the range of 4–60 breaths per minute (bpm), and (c) contextual outliers [15]. For (c), each value was compared to the mean of the surrounding ± 1.5 minutes of data and removed if the inspected value differed from the mean by more than 50%.

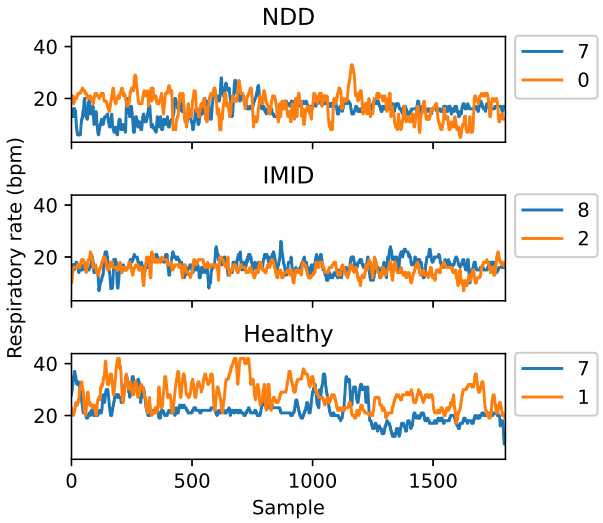

Figure 1. Samples of 2h respiratory rate signals associated with sleepy (blue) and non-sleepy (orange) states. A row shows two samples from an individual randomly chosen from the NDD (top), IMID (middle), and healthy (bottom) groups. The KSS levels (0–9) are shown in the legend.

2.4. Deep learning approach

A 1D CNN was employed to learn respiratory rate patterns and classify the samples into two target classes: (1) non-sleepy and (2) sleepy. The non-sleepy class was defined as KSS levels from "extremely alert" to "rather alert" (indexed 0 to 3), whereas the sleepy class was represented by KSS levels from "neither alert nor sleepy" to "extremely sleepy" (indexed 4 to 9).

The respiratory rate over the 2 hours preceding a KSS response was selected as the prediction input. Thus, the input samples were 1800-value time series. Any missing values were padded with zeros; however, a respiratory rate coverage of at least 90% was required in the 2h window. The coverage was evaluated as the number of observations compared to the observations expected per the sampling frequency. Additionally, for each participant, at least six eligible 2h windows with a corresponding KSS score were required, to capture variation within subject. Figure 1 depicts some eligible samples coupled with the KSS scores.

The 1D CNN model was built from two convolutional layers coupled with max pooling and dense layers. Rectified linear units were used for activation. The training set samples' average respiratory rate and standard deviation were used to standardize the prediction input. The final layers consisted of a dropout layer and a dense layer with softmax activation. The weighted sparse categorical cross-entropy loss function was used together with the Adam optimizer. Model performance was measured via classifica-

tion accuracy, sensitivity, and specificity.

The dataset was grouped by the participant and split randomly into training and test sets, with 20 % of the subjects held out for testing. The training set was further split by 5-fold cross-validation (CV), and the cross-validation was utilized in hyperparameter selection. The final model was trained on the full training data and tested on the held-out 20 % test set. During training, 10% of training data were used to estimate validation metrics and over-fitting.

3. Results

Table 1 shows the participants' demographics. The mean age was 50.9 (±16.1) years, time since diagnosis 11.1 (±9.2) years, and body-mass index 23.8 (±7.2) kg/m^2.

Each participant wore the patch-like sensor for 3–12 days while responding to KSS questionnaires. The total number of 2h respiratory rate samples coupled with a patient reported sleepiness score was 1255, comprising 187 samples for NDD patients, 708 for IMID patients, and 360 for the healthy group. Whilst a minimum number of 6 samples was required, the mean number of obtained samples was 15 (maximum was 30).

The architecture of the 1D CNN is summarized in Figure 2. The hyperparameter optimization yielded a batch size of 30 samples, a learning rate of 0.0001, 50 % dropout rate, and kernel size 5 in the first and 3 in the second convolutional layer. Additionally, the second convolutional layer included an L2 regularization factor of 0.1.

In the cross-validation, the 1D CNN model achieved an average accuracy of 58.3%, sensitivity of 62.5%, and specificity of 52.6%, as detailed in Table 2. Over-fitting was monitored via the training and validation loss curves, and the training was limited to 25 epochs.

The final model achieved 62.6 % accuracy, 57.2 % sensitivity, and 69.2 % specificity in a held-out test set. Based on observations during CV, early stopping was applied after the training and validation loss diverged beyond 0.05 from each other for three consecutive epochs. The test set comprised 10 IMID (5 SLE, 3 RA, 2 PSS) and 2 NDD (both PD) patients and 5 healthy participants, with 7–30 samples per participant. On average, 52% of the participant's samples represented the sleepy class.

Table 1. Participant demographics by participant group.

	NDD	IMID	Healthy
Study sites	K, G	E, K, N	E, G, K, N
Number	14	42	26
Female	7	34	12
Male	7	8	15
Age (mean±SD)	53.6 (±12.6)	53.1 (±16.0)	46.1 (±17.4)
Years since diagnosis (mean±SD)	4.3 (±2.5)	13.5 (±9.4)	–
Body-mass index (mean±SD)	24.0 (±2.4)	22.5 (±9.2)	25.7 (±4.4)

Table 2. Cross-validation results.

	Accuracy	Sensitivity	Specificity
1	0.5226	0.6484	0.4167
2	0.6061	0.6458	0.5686
3	0.5707	0.4884	0.6334
4	0.6396	0.7344	0.4638
5	0.5758	0.6061	0.5455
Average	0.5829	0.6246	0.5257

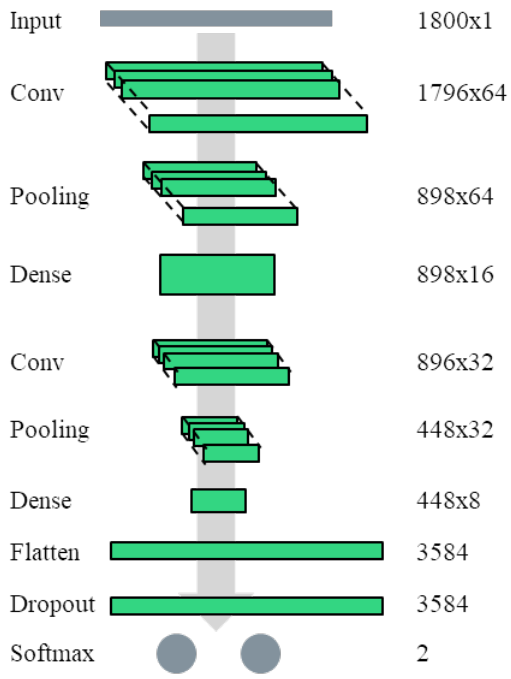

Figure 2. The model architecture comprises 10,426 trainable parameters.

4. Discussion

This study presented a 1D CNN model using ECG-based respiratory rate data to predict daytime sleepiness at the end of a 2-hour monitoring sequence. The final model achieved a 63% accuracy between non-sleepy and sleepy states, together with 57% sensitivity and 69% specificity. The training data included participants from six chronic disease cohorts and healthy participants, capturing several days and several times of the day, from afternoon to evening. We note that notable variations in the performance metrics were revealed in cross-validation within the

training set. However, the average accuracy of 58% was reasonably close to the final test accuracy. Overall, our results suggest that respiratory rate may have potential as a disease-independent predictor of daytime sleepiness.

A well performing objective digital measure can be useful for the clinical assessment of daytime sleepiness in chronic patients. Continuous measures may capture long-term temporal trends more easily than a PRO and enable assessing patient state in the free-living environment, comprehensively describing the effect of sleepiness on the patient's day-to-day quality of life. However, the PROs act as a reference in prediction modelling. This complicates the development of a model that can generalize to new subjects since the scoring may differ significantly from person to person due to individuals' subjective experiences. Thus, personalized models may achieve improved results.

Future studies may explore more specific respiratory patterns from the respiratory signal, and combine with other modalities, e.g, activity measures. Moreover, previous studies have reported that age can affect yawning frequencies [10]. We note that the presented study did not incorporate any demographic information in the model. Personalized prediction models may overcome this and perform better for individual patients. Future studies should eventually focus on developing more complex models that can indicate the level of sleepiness, especially capturing clinically relevant changes in daytime sleepiness.

Acknowledgments

The study was funded via the IDEA-FAST project, which has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No. 853981. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA and associated partners. The authors express their gratitude to the IDEA-FAST project members (https://idea-fast.eu/the-idea-fast-investigators/) for their work facilitating the presented study.

References

[1] Colten HR, Altevogt B (eds.). Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem. National Academies Press (US), 2006.

[2] Ohayon MM, Vecchierini MF. Daytime Sleepiness and Cognitive Impairment in the Elderly Population. Archives of Internal Medicine 2002;162(2):201–208. Doi: 10.1001/archinte.162.2.201.

[3] Feng F, Cai Y, Hou Y, Ou R, Jiang Z, Shang H. Excessive daytime sleepiness in parkinson's disease: A systematic review and meta-analysis. Parkinsonism Related Disorders 2021;85:133–140. Doi: 10.1016/j.parkreldis.2021.02.016.

[4] Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient non-compliance with paper diaries. BMJ May 2002;324:1193–1194. Doi: 10.1136/bmj.324.7347.1193.

[5] Laberge L, Gallais B, Auclair J, Dauvilliers Y, Mathieu J, Gagnon C. Predicting daytime sleepiness and fatigue: a 9-year prospective study in myotonic dystrophy type 1. Journal of Neurology 2020;267:461–468.

[6] Chervin RD, Burns JW, Ruzicka DL. Electroencephalographic changes during respiratory cycles predict sleepiness in sleep apnea. American Journal of Respiratory and Critical Care Medicine 2004;171(6).

[7] Igasaki T, Nagasawa K, Akbar IA, Kubo N. Sleepiness classification by thoracic respiration using support vector machine. In 2016 9th Biomedical Engineering International Conference (BMEiCON). 2016; 1–5. Doi: 10.1109/BMEiCON.2016.7859630.

[8] Bao J, Han J, Kato A, Kunze K. Sleepy watch: Towards predicting daytime sleepiness based on body temperature. In Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers, UbiComp-ISWC '20. 2020; 9–12. Doi: 10.1145/3410530.3414415.

[9] Guggisberg AG, Mathis J, Schnider A, Hess CW. Why do we yawn? Neuroscience Biobehavioral Reviews 2010; 34(8):1267–1276. Doi: 10.1016/j.neubiorev.2010.03.008.

[10] Zilli I, Giganti F, Uga V. Yawning and subjective sleepiness in the elderly. Journal of Sleep Research 2008;17(3):303–308. Doi: 10.1111/j.1365-2869.2008.00666.x.

[11] The IDEA-FAST project consortium. D2.1: First study subject approvals package of the feasibility study (FS) 2020; https://idea-fast.eu/results-and-publications/. Accessed: 2022-08-20.

[12] Chen L, Ma X, Chatterjee M, Kortelainen JM, Ahmaniemi T, Maetzler W, Wang P, Zhang D. Fatigue and sleep assessment using digital sleep trackers: Insights from a multi-device pilot study. In Proceedings of the 2022 44th Annual International Conference of the IEEE Engineering and Medicine Society. 2022; (accepted).

[13] Morgado Areia C, Santos M, Vollam S, Pimentel M, Young L, Roman C, Ede J, Piper P, King E, Gustafson O, Harford M, Shah A, Tarassenko L, Watkinson P. A chest patch for continuous vital sign monitoring: Clinical validation study during movement and controlled hypoxia. J Med Internet Res 2021;23(9). Doi: 10.2196/27547.

[14] Vildjiounaite E, Kallio J, Kyllönen V, Nieminen M, Määttänen I, Lindholm M, Mäntyjärvi J, Gimel'farb G. Unobtrusive stress detection on the basis of smartphone usage data. Personal and Ubiquitous Computing 2018;22:671–688. Doi: 10.1007/s00779-017-1108-z.

[15] Nicolò A, Massaroni C, Schena E, Sacchetti M. The importance of respiratory rate monitoring: From healthcare to sport and exercise. Sensors 2020;20(21). Doi: 10.3390/s20216396.

Address for correspondence:

Emmi Antikainen
Visiokatu 4, P.O. Box 1300, 33101 Tampere, Finland
emmi.antikainen@gmail.com