scientific reports

Check for updates

OPEN

Sleep prediction using data from oximeter, accelerometer and snoring for portable monitor obstructive sleep apnea diagnosis

Diego Munduruca Domingues[1]✉, Paloma Rodrigues Rocha[1], Ana Cláudia M. V. Miachon[1], Sara Quaglia de Campos Giampá[2], Filipe Soares[1], Pedro R. Genta[2] & Geraldo Lorenzi-Filho[2]

The aim of this study was to build and validate an artificial neural network (ANN) algorithm to predict sleep using data from a portable monitor (Biologix system) consisting of a high-resolution oximeter with built-in accelerometer plus smartphone application with snoring recording and cloud analysis. A total of 268 patients with suspected obstructive sleep apnea (OSA) were submitted to standard polysomnography (PSG) with simultaneous Biologix (age: 56 ± 11 years; body mass index: 30.9 ± 4.6 kg/m^2, apnea-hypopnea index [AHI]: 35 ± 30 events/h). Biologix channels were input features for construction an ANN model to predict sleep. A k-fold cross-validation method (k=10) was applied, ensuring that all sleep studies (N=268; 246,265 epochs) were included in both training and testing across all iterations. The final ANN model, evaluated as the mean performance across all folds, resulted in a sensitivity, specificity and accuracy of 91.5%, 71.0% and 86.1%, respectively, for detecting sleep. As compared to the oxygen desaturation index (ODI) from Biologix without sleep prediction, the bias (mean difference) between PSG-AHI and Biologix-ODI with sleep prediction (Biologix-Sleep-ODI) decreased significantly (3.40 vs. 1.02 events/h, p<0.001). We conclude that sleep prediction by an ANN model using data from oximeter, accelerometer, and snoring is accurate and improves Biologix system OSA diagnostic precision.

Keywords Artificial neural network, Sleep prediction, Obstructive sleep apnea

Obstructive sleep apnea (OSA) is characterized by repetitive episodes of upper airway obstruction, resulting in sleep fragmentation and oxygen desaturation[1]. OSA is associated with several health consequences, including poor sleep quality, excessive daytime sleepiness, and increased cardiovascular risk[2]. Polysomnography (PSG) is considered the gold standard method for OSA diagnosis[3]. However, PSG is expensive and inconvenient for patients[3]. Portable monitoring (PM) is a simplified method that has been validated for OSA diagnosis[4]. In contrast to PSG, PM does not detect sleep. The consequence of this limitation is that the number of respiratory events in PM devices are reported by hour of monitoring rather than hours of sleep. Therefore, the absence of sleep monitoring is a potential source of variability between PSG and PM. Biologix system is a new PM device based on a high-resolution wireless oximeter (Oxistar^{TM}, Biologix Sistemas S.A., Brazil) with built-in accelerometer and a smartphone application (app) that is downloaded to the patient's smartphone. The app records snoring, and all information is automatically processed in the cloud. Biologix system has been validated for OSA diagnosis against PSG in the sleep laboratory[5] and against traditional PM at home[6]. However, Biologix system does not monitor sleep and therefore reports oxygen desaturation index (ODI) based on hours of monitoring rather than hours of sleep. Therefore, the objective of this study was to build and validate an artificial neural network (ANN) algorithm using data from oximeter, accelerometer and snoring to detect sleep. We also tested the hypothesis that ANN model improves the Biologix system OSA diagnostic precision.

[1]Biologix Sistemas S.A., São Paulo, SP, Brazil. [2]Laboratório do Sono, LIM 63, Divisão de Pneumologia, Instituto do Coração, InCor, Hospital das Clínicas HCFMUSP, Universidade de São Paulo, Eneas de Carvalho Aguiar 44, 8º andar, São Paulo, SP 05403-900, Brazil. ✉email: diego.domingues@biologix.com.br

Methods

Patients and data collection

The study included patients recruited in the validation study of the Biologix system against PSG. Full details of the protocol have been published elsewhere[5]. The local ethics committee (Comissão de Ética para Análise de Projetos de Pesquisa do HCFMUSP - CAPPesq) approved the protocol (SDC 4515/17/015), and all patients gave their informed consent. The study has been performed in accordance with the Declaration of Helsinki. Briefly, we studied patients with suspected OSA referred for overnight-laboratory PSG at the Sleep Laboratory of the Heart Institute (InCor).

PSG included recording of the electroencephalogram (EEG) central (C) and occipital (O) channels referred to the auricular channel (A) (C3/A2, C4/A1, O1/A2, O2/A1), electrooculogram (EOG), submental electromyogram (EMG), left and right anterior tibialis EMG, electrocardiogram, thoraco-abdominal effort, oronasal airflow (thermistor and nasal pressure based airflow measurement), oxygen saturation (SpO_2) with pulse oximetry, and body position (EMBLA S7000, Embla Systems, USA and Alice 5, Respironics Inc., USA)[5]. Two certified technicians independently analyzed all PSG studies. Hypopnea was defined as a drop in the peak signal excursion of $\geq$ 30% from the pre-event baseline nasal pressure signal lasting for at least 10 seconds. Respiratory events were scored according to the American Academy of Sleep Medicine criteria ($\geq$ 3% reduction in SpO_2 from the pre-event baseline or an event associated with arousal). OSA was classified based on current standards as follows: absence of OSA (AHI < 5 events/hour), mild OSA (5 $\leq$ AHI < 15 events/hour), moderate OSA (15 $\leq$ AHI < 30 events/hour), and severe OSA (AHI $\geq$ 30 events/hour).

Simultaneously, the patients also wore a high-resolution oximeter ($Oxistar^{TM}$, Biologix Sistemas S.A., Brazil) with built-in accelerometer linked by Bluetooth to a smartphone app that recorded snoring. The $Oxistar^{TM}$ firmware captures data at a rate of 100 samples per second, providing beat-to-beat raw SpO_2 measurements with a precision of 0.1%. To smooth the data, a moving average over 4 heartbeats was applied. Oxygen desaturations are calculated providing the ODI. The ODI was calculated as the number of desaturations ($\geq 3\%$ reduction in SpO_2) per hour, using either total recording time or total sleep time. The oximeter information was sent to the cloud, and automatically analyzed (Fig. 1). The PSG and Biologix data were time-synchronized.

ANN algorithm

ANN are algorithms based on the biological structure of the human brain, in which several neurons are connected[7,8]. These neurons are divided into at least three layers: inputs, a variable number of hidden layers, and outputs. Each of these layers is connected to the next layer by an activation function, a weight associated with its signal, and a bias[8]. To build and validate our ANN algorithm, we used data derived from the Biologix system including oximeter (SpO_2, heart rate [HR]), with built- in accelerometer (movement), and smartphone app (snoring). Snoring was obtained by recording the audio of the environment performed by the smartphone app and processed by another neural network. This algorithm provides a binary output indicating whether the patient is snoring or not during the audio recording stretches, similar to other approaches found in the literature[9]. The k-fold cross-validation method (k=10) was used to build and validate the ANN algorithm[10,11].

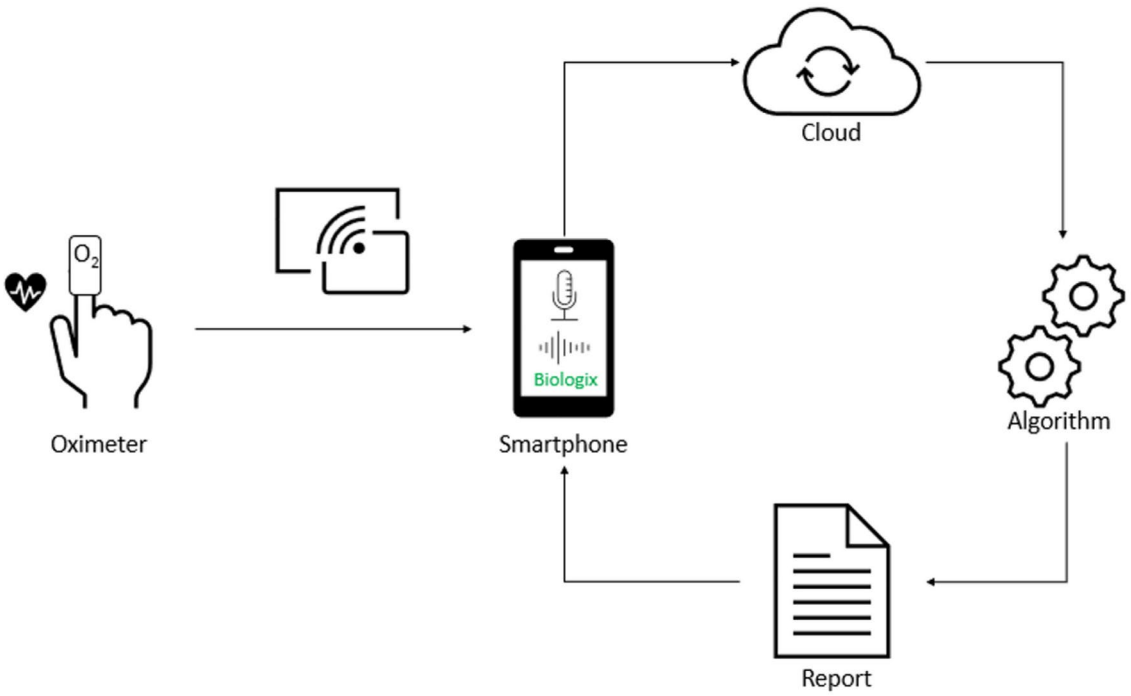


Fig. 1. Biologix system. The wireless oximeter connects via Bluetooth to the smartphone's Biologix application. The data is sent to the cloud and automatically analyzed by the algorithm.

The sleep studies were randomly divided into 10 folds, with each fold used for cross-validation to ensure that all studies were both trained and tested across multiple iterations. In each fold, the training and test datasets were employed to optimize the weights and biases, reducing the error between the predicted value by the neural network results. The gold standard for sleep classification was a binary variable (sleep or awake) determined by PSG epochs of 30 seconds. The process starts with a forward pass for initial values of weight and bias and for pre-defined activation functions. Outputs are then calculated, and errors are determined. In the next step, the values of weights and biases are redefined through a process called backpropagation[12]. Using the newly calculated values, the process is redone. This occurs recursively until a maximum number of iterations[8,13]. Our model consists of an input layer, one hidden layer and an output layer. The input layer has 97 neurons, the hidden layer has 128 neurons and a ReLU (rectified linear unit)[14] activation function. Finally, the output layer has 1 neuron and a sigmoid activation function (Fig. 2). In order to test the accuracy of the accelerometer alone in predicting sleep, another ANN model was built using only the accelerometer channel, with 23 neurons in the input layer, while the other layers remained the same.

Features

The first step was the treatment of missing values (less than 1% of the data was missing), which consisted of replacing these values by zero, in the case of the accelerometer, and by the maximum values for the cases of SpO_2 and HR. Subsequently, the features were calculated using epochs of 30-seconds synchronize to PSG epochs. The features used in the model were calculated based on the signals obtained by the Biologix system and included SpO_2, HR, movement detected by the accelerometer, and snoring detected by the smartphone app. The SpO_2 features were: (1) presence or absence of oxygen desaturation, expressed as a binary variable; (2) desaturation range; (3) SpO_2 quartile (75th) as a measure of the tendency of the patient's SpO_2 values during sleep. The HR signal provided several features: (1) average pulse interval; (2) standard deviation of HR; (3) HR variability (HRV) time domain features (SDNN, RMSSD, PNN50, SD1, SD2); (4) HRV frequency domain features (LF power, HF power, LF/HF ratio). The accelerometer data generated multiple features: (1) variance; (2) root mean square (RMS); (3) skewness; (4) kurtosis. Other variables of interest associated with snoring, obtained by the

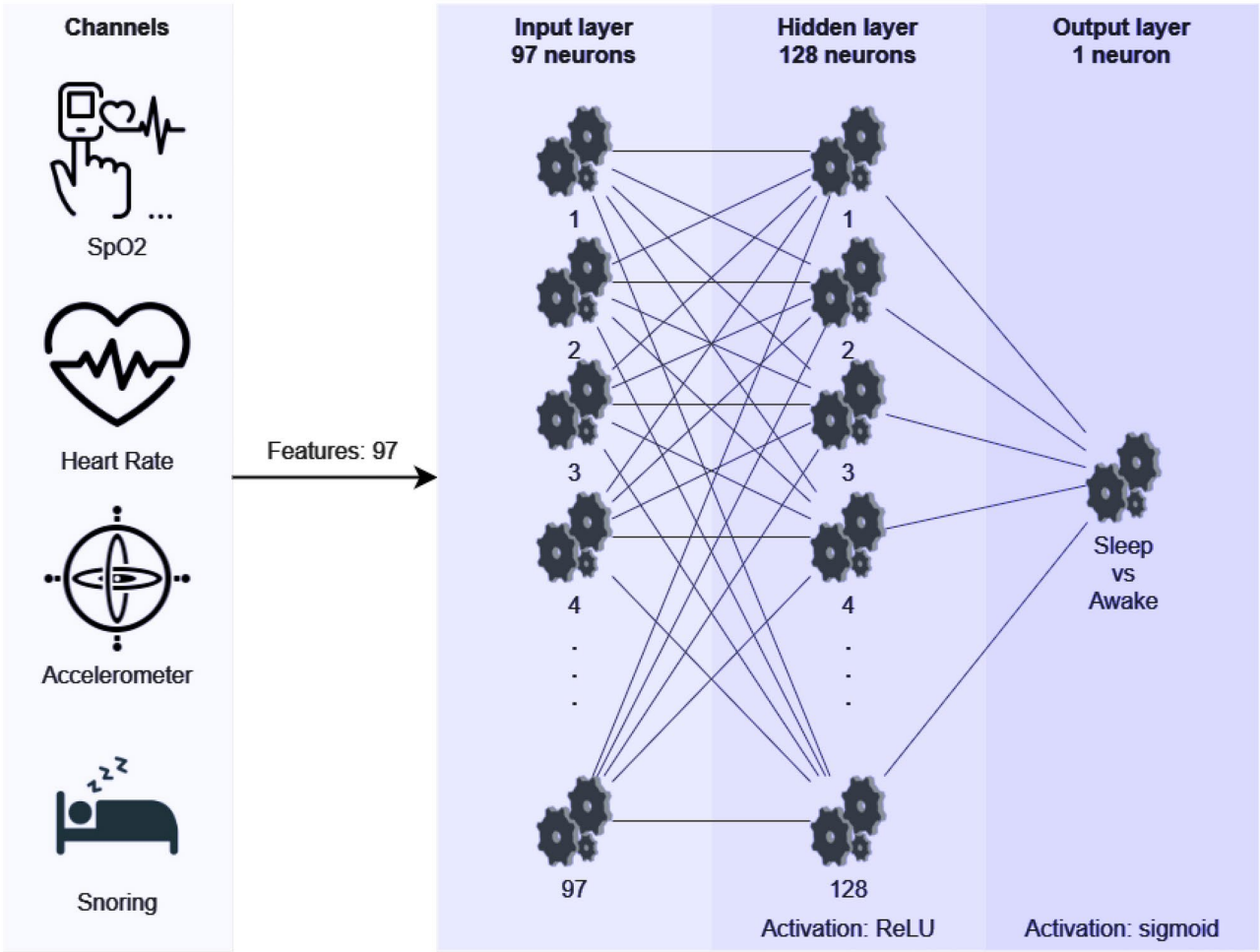


Fig. 2. Artificial neural network algorithm diagram. The 97 input features were extracted from the channels, processed by the hidden layer with 128 neurons, resulting in a single neuron output that predicted whether the patient was sleep or awake. SpO_2 oxygen saturation, *ReLU* rectified linear unit.

Biologix app, were also used, which improved the performance of the machine learning model. In addition, some of these features were considered shifted in relation to the current time step for the better composition of the predictive model, totaling 97 inputs for the neural network model summarized in the Table 1. Finally, the data was standardized by removing the mean and scaling to unit variance[15,16].

$$z = \frac{(x - u)}{s}$$

Where u is the mean value, s is the standard deviation, x is the samples and z is the new samples[15,16].

Statistical analysis

Accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), Cohen's kappa coefficient (κ), F1-score (weighted average between the precision score [PPV] and recall score [sensitivity]), and area under the curve (AUC) were calculated for evaluation of the ANN model. Because several previous studies used only accelerometer data to predict sleep, we used the McNemar's test[17] to compare the ANN performance to predict sleep using only accelerometer data with ANN performance using all Biologix channels (SpO_2, HR, accelerometer, and snoring). Overall summary statistics were calculated in terms of means and standard deviations for continuous variables and counts and percentages for categorical variables. Shapiro-Wilk test was used for checking the data normality of the PSG and the Biologix system. Since the data distribution was not normal and they were not independent, the differences were analyzed by the Wilcoxon signed-rank test. In addition, we calculated the sensitivity, specificity, accuracy, PPV, and NPV of the Biologix system, without and with sleep prediction, versus PSG, in the detection of OSA severity. Mc Nemar's test[17] compared the Biologix system performance without and with sleep prediction. Finally, to assess the amount of agreement on OSA diagnosis between PSG-AHI and Biologix-ODI, without and with sleep prediction, Bland-Altman plots were performed. RStudio 2023.06.1 software (R Foundation for Statistical Computing) was used for all statistical analysis. Significance was assessed with a p-value < 0.05.

Results

Out of 304 consecutive patients previously evaluated to validate the Biologix system against PSG[5], 268 had snoring recordings and were used for this study. The patients had typical characteristics of patients referred for OSA diagnosis, and were predominantly obese middle-aged adults, with comorbidities and with a high proportion of moderate to severe OSA (Table 2). As described in the method section, the sleep studies underwent 10-fold cross-validation. Each fold consisted of a training set of 90% of the patients (approximately 241 patients, corresponding to 221,639 epochs) and a test set of 10% of the patients (approximately 27 patients, corresponding to 24,627 epochs).

The total sleep time determined by PSG and by the Biologix ANN model using all channels was similar (353 ± 65 min vs. 359 ± 56 min, respectively, p=0.15), as well as sleep efficiency (0.76 ± 0.12 % vs. 0.77 ± 0.10 %, respectively, p=0.15). Table 3 shows the performance metrics of the ANN model using only the accelerometer and the ANN model using all Biologix channels to predict sleep, as assessed by k-fold cross-validation. The performance of the ANN when all Biologix channels were used was significantly higher than when only accelerometer data was used ($p < 0.001$), as revealed by McNemar's test. The ANN model using data from all Biologix channels achieved higher AUC values of the receiver operating characteristic (ROC) curve, indicating superior performance in predicting sleep (Fig. 3).

Data	Features	No. of features
SpO_2	Desaturation events	18
	Desaturation range	
	SpO_2 quartile (75th)	
HR	Average pulse interval	40
	Standard deviation	
	Heart rate variability	
Accelerometer	Variance	22
	Root mean square	
	Skewness	
	Kurtosis	
Snore	Snore events	16
	Snoring proportion	
Epoch position	Normalized time of the epoch	1

Table 1. Features inputted into the ANN algorithm with all Biologix channels. Epoch position is the normalized time of the epoch position in relation to the duration of the entire sleep study (expressed as a percentage of the sleep study, being 0% and 100% the beginning and end of the study, respectively). SpO_2 oxygen saturation, *HR* heart rate.

	Total
	(N = 268)
Male (%)	146 (54.5)
Age (years)	56 ± 11
Body mass index (kg/m^2)	30.9 ± 4.6
Epworth sleepiness scale score	11 ± 5
Arterial hypertension (%)	132 (49.3)
Dyslipidemia (%)	76 (28.4)
Diabetes mellitus (%)	51 (19.0)
Depression (%)	24 (9.0)
Coronary artery disease (%)	18 (6.7)
Asthma/chronic obstructive pulmonary disease (%)	13 (4.9)
Polysomnography	
Total recording time (min)	465 ± 42
Total sleep time (min)	353 ± 65
Sleep latency (min)	43 ± 36
Wakes after sleep onset (min)	66 ± 46
Sleep efficiency (%)	76 ± 12
Apnea-hypopnea index (events/h)	35 ± 30
No OSA (%)	29 (10.8)
Moderate to severe OSA (%)	176 (65.7)
Oxygen desaturation index (events/h)	32 ± 29
Biologix	
Total recording time (min)	465 ± 42
Oxygen desaturation index (events/h)	32 ± 26

Table 2. Demographic and sleep data of the population studied. *OSA* obstructive sleep apnea. Apnea-hypopnea index, using 3% desaturation criterion for hypopnea definition

	Algorithm	
	ANN with only accelerometer	ANN with accelerometer, SpO_2, HR and snoring
Accuracy	84.5 (76.3-92.6)	86.1 (78.9-93.2)
Sensitivity	91.7 (84.0-99.5)	91.5 (84.4-98.6)
Specificity	65.4 (47.3-83.5)	71.0 (54.8-87.3)
PPV	88.0 (77.7-98.3)	89.8 (80.6-99.1)
NPV	71.2 (52.9-89.5)	72.3 (54.6-89.9)
κ	0.55 (0.39-0.71)	0.60 (0.45-0.75)
F1-score	0.89 (0.82-0.96)	0.90 (0.84-0.97)
AUC	0.88 (0.80-0.96)	0.90 (0.83-0.98)

Table 3. Performance assessment of the ANN models to predict sleep. Data in parentheses are 95% confidence intervals. *ANN* artificial neural network, SpO_2 oxygen saturation, *HR* heart rate, *PPV* positive predictive value, *NPV* negative predictive value, κ Cohen's kappa coefficient, *AUC* area under the curve.

The ability to predict sleep was evaluated for its impact on OSA severity determination. AHI was classified using cut-offs of 5, 15, and 30 events/h. We calculated the specificity, sensitivity, accuracy, PPV, and NPV of the Biologix system, without and with sleep prediction (Table 4). Compared to Biologix-ODI, the specificity to predict OSA as defined as AHI $\geq$ 5 or 15 events/h increased with sleep prediction, however, the improvement did not reach statistical significance.

Bland-Altman analysis between PSG-AHI versus Biologix-ODI showed that the limits of agreement were similar without and with sleep prediction ([-20,27] vs. [-20,22], respectively), whereas the bias (mean difference) decreased significantly with sleep prediction (3.40 vs. 1.02, $p < 0.001$) (Fig. 4a and b).

Discussion

In the present study, we built and validated a new algorithm to predict sleep based on the ANN model using data derived from the Biologix system, including a high-resolution oximeter with built-in accelerometer, and snoring detected by the smartphone with the Biologix app.[5,6]. Firstly, the algorithm exhibited good performance in distinguishing sleep from awake, with a sensitivity of 91.5%, specificity of 71.0%, accuracy of 86.1%, Cohen's

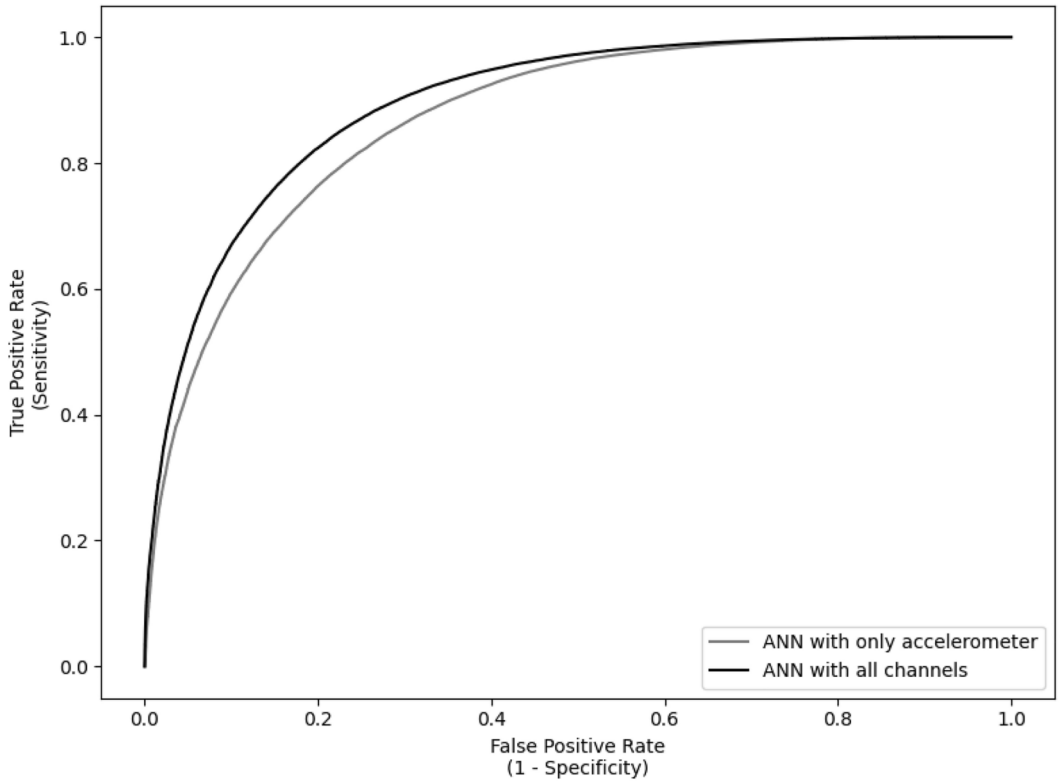


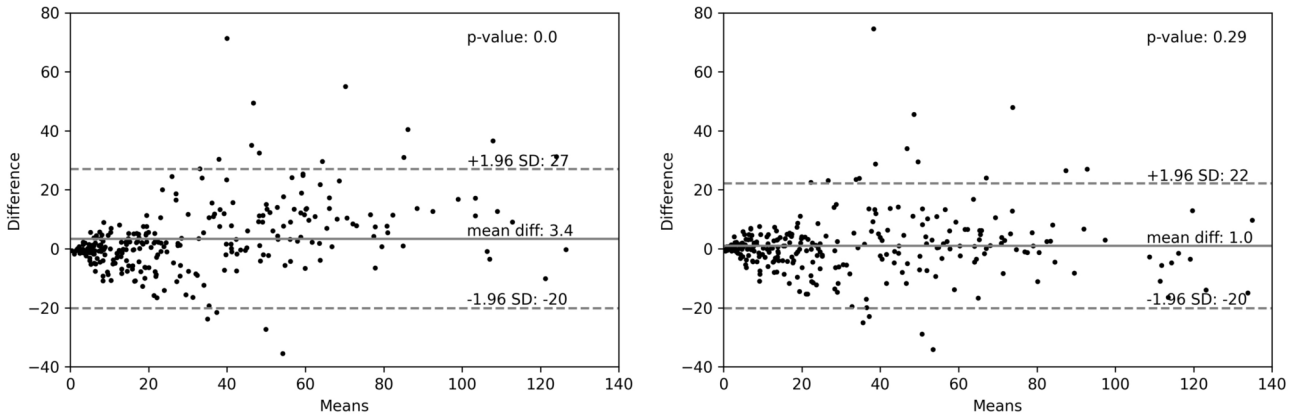
Fig. 3. Receiver operating characteristic (ROC) curve. The ROC curve for the ANN using only the accelerometer and the ANN using all Biologix channels (SpO_2, HR, accelerometer, and snoring).

PSG, AHI	OSA ≥ 5 events/h		Mod-Sev OSA ≥ 15 events/h		Mild 5-14.9 events/h		Moderate 15-29.9 events/h		Severe ≥ 30 events/h	
Biologix	**ODI**	**Sleep ODI**	**ODI**	**Sleep ODI**	**ODI**	**Sleep ODI**	**ODI**	**Sleep ODI**	**ODI**	**Sleep ODI**
Specificity (%)	58.6	75.9	83.8	87.7	87.7	85.4	77.7	88.4	90.9	91.6
Sensitivity (%)	96.2	93.7	92.0	91.4	90.1	91.4	91.4	91.4	87.7	90.4
Accuracy (%)	92.2	91.8	89.2	90.0	84.4	86.8	89.5	90.7	81.7	82.5
PPV (%)	95.0	97.0	91.5	93.0	85.7	88.7	94.4	92.7	59.7	64.4
NPV (%)	65.4	59.5	84.8	87.4	73.4	84.6	84.7	91.2	88.3	87.7

Table 4. Diagnostic performance of Biologix system without and with sleep prediction in OSA severity determination. *PSG* polysomnography, *AHI* apnea-hypopnea index, *OSA* obstructive sleep apnea, *ODI* oxygen desaturation index considering total recording time, *Sleep ODI* oxygen desaturation index considering total sleep time, *PPV* positive predictive value, *NPV* negative predictive value.

kappa coefficient of 0.60, F1-score of 0.90, and AUC of 0.90. Secondly, the algorithm performance to predict sleep was better using all Biologix channels compared to using only the accelerometer channel. Thirdly, Biologix system with sleep prediction decreased the bias between PSG-AHI and Biologix-ODI.

PM has been validated to detect OSA[18,19]. However, in contrast to PSG, PM does not discriminate wakefulness from sleep. In this context, several studies have attempted to predict sleep using a binary classification (sleep versus wakefulness) mainly using actigraphy. Overall, actigraphy is recognized as an accurate and sensitive method to detect sleep periods, but with poor specificity to identify wakefulness (ranging from 32 and 61%)[20–22]. A study carried out with 8 commercial sleep tracking devices, showed that the sensitivity to detect sleep, as compared to PSG, was very high (all greater than or equal to 93%). However, the specificity for predicting sleep was variable and generally low, ranging from 18 to 54%[23]. Alternatively, algorithm models based on data from several channels, either isolated or in combination, including accelerometer, respiratory signals, breathing sounds, and HRV, have been used in an attempt to improve sleep/wakefulness detection by PM[22,24–31]. For instance, Dafna et al.[26] developed and validated an algorithm for detecting sleep periods in patients with OSA based on the analysis of respiratory sounds. Despite the high sensitivity of 92.2%, the specificity to detect wakefulness was low (56.6%). In turn, Montazeri et al.[32] reported promising results with a sensitivity of 87.8%, specificity of 71.4%, and accuracy of 82.3% to detect sleep using an algorithm model based on tracheal sound and movement



(a) Bland-Altman plot for PSG-AHI and Biologix-ODI. **(b)** Bland-Altman plot for PSG-AHI and Biologix-Sleep-ODI.

Fig. 4. Bland-Altman plots for PSG-AHI and Biologix-ODI. (**a**) Bland-Altman plot for PSG-AHI and Biologix-ODI without sleep prediction. (**b**) Bland-Altman plot for PSG-AHI and Biologix-Sleep-ODI with sleep prediction. *AHI* apnea-hypopnea index, *PSG* polysomnography, *ODI* oxygen desaturation index considering total recording time, *Sleep ODI* oxygen desaturation index considering total sleep time.

data recorded with a small wearable device attached over the trachea. The reported device is also able to predict the AHI based on tracheal sounds, however, it does not measure SpO_2. In general, the main challenge of all systems that do not measure EEG is to achieve high sensitivity to detect sleep while also maintaining reasonable specificity. To address this challenge, our algorithm combines built-in accelerometer data with other variables, including SpO_2, HR, and snoring. As compared to the ANN algorithm using only the accelerometer, the ANN algorithm using all channels had a better performance in detecting sleep. Specifically, the specificity increased from 65.4% to 71.0%, without compromising sensitivity (91.7% vs. 91.5%, respectively), resulting in an increase of the accuracy (84.5% vs. 86.1%). In addition, the AUC increased from 0.88 to 0.90, indicating a greater ability to discriminate between sleep and wakefulness when all Biologix channels were used.

Compared to PSG-AHI, Biologix-ODI with sleep prediction (Biologix-Sleep-ODI) improved the performance of Biologix-ODI.The Bland-Altman plots (Fig. 4) showed a significant decrease in the bias between PSG-AHI and Biologix-ODI when sleep prediction was taken into account (Biologix-Sleep-ODI). As compared to Biologix-ODI, the specificity to predict OSA as defined as AHI $\geq$ 5 or 15 events/h increased with sleep prediction (Biologix-Sleep-ODI), however, the improvement did not reach statistical significance (Table 4).

Despite the strengths, our study has limitations. Firstly, the number of patients was relatively small. On the other hand, because sleep prediction was based on 30-second epoch, the study used a large data set to build an algorithm to predict sleep. Secondly, the study was carried out using data from patients with suspected OSA, so the accuracy of our ANN model to predict sleep in healthy subjects may be different. Another caveat is that in contrast to wrist actigraphy, Biologix built-in actigraphy is placed on the finger. Although we have no reason to believe that there is a substantial difference between wrist and finger movements, we acknowledge that this comparison was not performed.

Conclusion

In conclusion, we showed that an algorithm based on ANN using all Biologix channels, including SpO_2, HR, accelerometer, and snoring is able to detect sleep with a good accuracy. Sleep prediction resulted in a reduction in the bias between PSG-AHI and Biologix-Sleep-ODI.

Data availability

The data that support the findings of this study are available from Biologix Sistemas S.A., but restrictions apply to the availability of these data, which were used under license for the current study and are not publicly available. Data are, however, available from the authors upon reasonable request and with the permission of Biologix Sistemas S.A.

Received: 19 February 2024; Accepted: 9 October 2024
Published online: 19 October 2024

References

1. Sleep-related breathing disorders in adults. Recommendations for syndrome definition and measurement techniques in clinical research. *Sleep* **22**, 667–689 (1999).
2. Colten, H. R. & Altevogt, B. M. *Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem* (Institute of Medicine, 2006).
3. Rundo, J. V. & Downey, R. Chapter 25 - polysomnography. In Levin, K. H. & Chauvel, P. (eds.) *Clinical Neurophysiology: Basis and Technical Aspects*. In *Handbook of Clinical Neurology*, vol. 160, 381–392 (Elsevier, 2019).

4. Practice Committee of the American Sleep Disorders Association. Practice parameters for the use of portable recording in the assessment of obstructive sleep apnea. *Sleep* **17**, 372–377 (1994).
5. Do Lago Pinheiro, G. et al. Validation of an overnight wireless high-resolution oximeter plus cloud-based algorithm for the diagnosis of obstructive sleep apnea. *Clinics* **75**, e2414 (2020).
6. Hasan, R. et al. Validation of an overnight wireless high-resolution oximeter for the diagnosis of obstructive sleep apnea at home. *Sci. Rep.* **12**, 15136 (2022).
7. Johnson, K. *Supervised Learning for Sequence Prediction Using Keras Sequential Models*. Master's thesis, California State University, Northridge (2018).
8. Poole, D. L. & Mackworth, A. K. *Artificial Intelligence: Foundations of Computational Agents* (Cambridge University Press, 2010).
9. Camacho, M. et al. Smartphone apps for snoring. *J. Laryngol. Otol.* **129**, 974–979 (2015).
10. Poldrack, R. A., Huckins, G. & Varoquaux, G. Establishment of best practices for evidence for prediction: a review. *JAMA Psychiat.* **77**, 534–540 (2020).
11. Mencar, C. et al. Application of machine learning to predict obstructive sleep apnea syndrome severity. *Health Inform. J.* **26**, 298–317 (2020).
12. Leung, H. & Haykin, S. The complex backpropagation algorithm. *IEEE Trans. Signal Process.* **39**, 2101–2104 (1991).
13. Prechelt, L. *Early stopping-but when? In Neural Networks: Tricks of the trade* 55–69 (Springer, 2002).
14. Rasamoelina, A. D., Adjailia, F. & Sinčák, P. A review of activation function for artificial neural network. In *2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI)*, 281–286 (IEEE, 2020).
15. Pedregosa, F. et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
16. Buitinck, L. *et al.* API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, 108–122 (2013).
17. Kim, S. & Lee, W. Does Mcnemar's test compare the sensitivities and specificities of two diagnostic tests?. *Stat. Methods Med. Res.* **26**, 142–154 (2017).
18. Khor, Y. H. *et al.* Portable evaluation of obstructive sleep apnea in adults: A systematic review. *Sleep Med. Rev.* 101743 (2023).
19. Collop, N. A. et al. Obstructive sleep apnea devices for out-of-center (ooc) testing: technology evaluation. *J. Clin. Sleep Med.* **7**, 531–548 (2011).
20. Marino, M. et al. Measuring sleep: accuracy, sensitivity, and specificity of wrist actigraphy compared to polysomnography. *Sleep* **36**, 1747–1755 (2013).
21. Montgomery-Downs, H. E., Insana, S. P. & Bond, J. A. Movement toward a novel activity monitoring device. *Sleep Breath.* **16**, 913–917 (2012).
22. Banfi, T. et al. Efficient embedded sleep wake classification for open-source actigraphy. *Sci. Rep.* **11**, 345 (2021).
23. Chinoy, E. D. *et al.* Performance of seven consumer sleep-tracking devices compared with polysomnography. *Sleep* **44** (2020).
24. Redmond, S. J. *et al.* Sleep staging using cardiorespiratory signals. *Somnologie* **11** (2007).
25. Kalkbrenner, C. *et al.* Automated sleep stage classification based on tracheal body sound and actigraphy. *GMS German Med. Sci.* **17** (2019).
26. Dafna, E., Tarasiuk, A. & Zigel, Y. Sleep-wake evaluation from whole-night non-contact audio recordings of breathing sounds. *PLoS One* **10**, e0117382 (2015).
27. Schade, M. M. et al. Sleep validity of a non-contact bedside movement and respiration-sensing device. *J. Clin. Sleep Med.* **15**, 1051–1061 (2019).
28. Aktaruzzaman, M. et al. Performance comparison between wrist and chest actigraphy in combination with heart rate variability for sleep classification. *Comput. Biol. Med.* **89**, 212–221 (2017).
29. García-Díaz, E. et al. Respiratory polygraphy with actigraphy in the diagnosis of sleep apnea-hypopnea syndrome. *Chest* **131**, 725–732 (2007).
30. Fonseca, P. et al. Validation of photoplethysmography-based sleep staging compared with polysomnography in healthy middle-aged adults. *Sleep* **40**, zsx097 (2017).
31. Devot, S., Dratwa, R. & Naujokat, E. Sleep/wake detection based on cardiorespiratory signals and actigraphy. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, 5089–5092 (IEEE, 2010).
32. Montazeri Ghahjaverestan, N. *et al.* Sleep/wakefulness detection using tracheal sounds and movements. *Nat. Sci. Sleep* 1009–1021 (2020).

Acknowledgements

Sleep laboratory of the Heart Institute (where PSG were performed), study participants and each coauthor are acknowledged.

Author contributions

D.M.D., A.C.M.V.M., and F.S. led the development and validation of the artificial neural network algorithm. D.M.D and P.R.R. analyzed the results. D.M.D, P.R.R, A.C.M.V.M., S.Q.C.G and G.L.F wrote the initial manuscript draft with input from all authors. P.R.G contributed substantially to the data analysis, interpretation of the data, or a combination thereof. All authors reviewed the manuscript.

Declarations

Competing interests

Authors Diego Munduruca Domingues, Paloma Rodrigues Rocha, Ana Cláudia M. V. Miachon and Filipe Soares are employees of Biologix. Geraldo Lorenzi-Filho is co-founder of Biologix. This research was funded by Biologix Sistemas S.A., Brazil.

Additional information

Correspondence and requests for materials should be addressed to D.M.D.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.