Received: 25 June 2021 | Revised: 20 August 2021 | Accepted: 7 September 2021
DOI: 10.1111/jsr.13490

RESEARCH ARTICLE

Sleep apnea severity based on estimated tidal volume and snoring features from tracheal signals

Nasim Montazeri Ghahjaverestan[1,2] | Shumit Saha[1,2] | Muammar Kabir[1,2] | Bojan Gavrilovic[1,2] | Kaiyin Zhu[1] | Azadeh Yadollahi[1,2]

[1]KITE, Toronto Rehabilitation Institute-University Health Network, Toronto, ON, Canada
[2]Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, ON, Canada

Correspondence
Nasim Montazeri Ghahjaverestan, University Health Network-Toronto Rehabilitation Institute, Room 12-019, 550 University Avenue, Toronto, ON M5G 2A2, Canada.
Email: Nasim.Montazeri@uhn.ca

Funding information
Ontario Centres of Excellence (OCE); NSERC; BresoTEC Inc.

Summary

Sleep apnea can be characterized by reductions in the respiratory tidal volume. Previous studies showed that the tidal volume can be estimated from tracheal sounds and movements called tracheal signals. Additionally, tracheal sounds include the sounds of snoring, a common symptom of obstructive sleep apnea. This study investigates the feasibility of estimating the severity of sleep apnea, as quantified by the apnea/hypopnea index (AHI), using the estimated tidal volume and snoring sounds extracted from tracheal signals. Tracheal signals were recorded simultaneously with polysomnography (PSG). The tidal volume was estimated from tracheal signals. The reductions in the tidal volume were detected as potential respiratory events. Additionally, features related to snoring sounds, which quantified variability, temporal clusters, and dominant frequency of snores, were extracted. A step-wise regression model and a greedy search algorithm were used sequentially to select the optimal set of features to estimate the apnea/hypopnea index and classify participants into healthy individuals and patients with sleep apnea. Sixty-one participants with suspected sleep apnea (age: 51 ± 16, body mass index: 29.5 ± 6.4 kg/m^2, apnea/hypopnea index: 20.2 ± 21.2 event/h) who were referred for a sleep test were recruited. The estimated apnea/hypopnea index was strongly correlated with the polysomnography-based apnea/hypopnea index ($R^2 = 0.76$, $p < 0.001$). The accuracy of detecting sleep apnea for the apnea/hypopnea index cutoff of 15 events/h was 78.69% and 83.61% with and without using snore-related features. These findings suggest that acoustic estimation of airflow and snore-related features can provide a convenient and reliable method for screening of sleep apnea.

KEYWORDS
acoustic analysis, sleep apnea, snoring, tidal volume, tracheal sounds

1 | INTRODUCTION

Sleep apnea is a chronic disorder associated with an increased risk of cardiorespiratory (4 folds; Fletcher et al., 1987) and neurocognitive (1.2 folds) disorders (Davies & Harrington, 2016). Sleep apnea is characterized by intermittent complete (apnea) or partial (hypopnea) reductions in respiratory airflow, each lasting more than 10 s during sleep. The severity of sleep apnea is commonly quantified by the number of apneas and hypopneas per hour of sleep, called the apnea/hypopnea index (AHI). However, to diagnose sleep apnea, individuals need to undergo overnight in-laboratory polysomnography (PSG), which is costly, inconvenient due to the attachment of >20 sensors, and has a long waiting list (2–60 months across the world; Pack, 2004). Therefore, up to 85%

Patent Owner, CleveMed - Ex. 2029, p. 1

of individuals with a high risk of sleep apnea are not diagnosed (Kapur et al., 2002), which is estimated to cost about $150B in the USA (Medicine, 2016) and $25B in Australia (Hillman et al., 2018). To address these challenges, portable home-based devices with fewer sensors have been developed to screen sleep apnea and estimate AHI.

One of the growing modalities for portable sleep apnea screening is based on respiratory sounds/movements recorded non-invasively over the trachea (Hafezi et al., 2020; Kalkbrenner et al., 2018; Kulkas et al., 2009; Nakano et al., 2004, 2019; Penzel & Sabil, 2017, 2018; Yadollahi et al., 2010). Previous studies estimated AHI using tracheal sounds combined with pulse oximetry (Cummiskey et al., 1982; Saha et al., 2019; Yadollahi et al., 2010), nasal pressure (Glos et al., 2019), or actigraphy (Kalkbrenner et al., 2017). Integrating another sensor such as pulse oximetry with the device attached over the trachea can increase screening accuracy; however, the integrated sensor requires synchronization with the tracheal device. Moreover, the synchronization complicates the hardware setting, especially for home-based use. Therefore, a more convenient and cost-effective approach is needed to assess sleep apnea using only the tracheal device.

Tracheal sounds/movements, called tracheal signals in this paper, include rich information about respiration and acoustic characteristics of the pharynx. There is a strong relationship between respiratory airflow and the energy of the tracheal sounds (Gavriely & Cugell, 1996; Sabil et al., 2019; Yadollahi & Moussavi, 2008). Furthermore, respiratory-related movements over the chest and abdomen can be recorded as caudal movements over the trachea (Hafezi et al., 2020; Tong et al., 2019). Using tracheal signals in our previous study on individuals with suspected sleep apnea, we estimated the relative tidal volume overnight with a high correlation with the PSG-based references (Montazeri Ghahjaverestan et al., 2021).

Additionally, tracheal sounds include snoring sounds generated by the vibration of the pharyngeal wall and the turbulence of airflow caused by a narrowing in the pharynx (Azarbarzin & Moussavi, 2010; Yadollahi & Moussavi, 2010). In previous studies, snoring sounds recorded by an ambient microphone have been analyzed for estimating the severity of obstructive sleep apnea (OSA; Abeyratne et al., 2005; Ben-Israel et al., 2012; Ng et al., 2009). However, compared with ambient sounds, tracheal sounds are recorded closer to the source of snores in the pharyngeal airway (Penzel & Sabil, 2017). Thus, they are less noisy (Yadollahi & Moussavi, 2010). Tracheal snoring sound has been used in one study to identify OSA (Azarbarzin & Moussavi, 2013); however, it has never been used for estimating AHI.

This study investigated the possibility of estimating AHI regardless of the type of respiratory events (central or obstructive) using the estimated tidal volume combined with snoring sounds, all extracted from tracheal signals. The estimated AHI was used to classify the population of the study into groups of individuals with (AHI ≥ 15) and without sleep apnea (AHI < 15).

2 | METHOD

2.1 | Study participants

Adults (age ≥ 18) with suspected sleep apnea referred between January 2017 and June 2018 to the sleep laboratory in the Toronto Rehabilitation Institute were recruited for this study. The Research Ethics Board of the University Health Network approved the protocol (IRB #: 15-8967). Candidates with an allergy to medical tape and history of tracheostomy were excluded from the study. All the participants signed written consent.

2.2 | Study procedures

Each participant underwent overnight in-laboratory PSG using Embla® N7000/S4500 (Natus Medical Incorporated). Based on the American Academy of Sleep Medicine (AASM) guidelines (Berry et al., 2012), PSG recordings were analyzed to annotate apneas and hypopneas, from which the PSG-based AHI was calculated for each participant. Apnea was defined as a 90% reduction in airflow lasting for more than 10 s. Hypopnea was characterized as >30% reduction in airflow for more than 10 s, accompanied by a 3% drop in oxygen saturation or cortical arousal (Berry et al., 2012). Simultaneously with PSG, tracheal sounds and movements were recorded with a small wearable device, The Patch (Saha et al., 2019), which included a microphone with a sampling frequency of 15,000 Hz and an accelerometer with a sampling frequency of 60 Hz. The Patch was taped securely over the suprasternal notch. A software program embedded in the microcontroller of The Patch was used to synchronize data recording with the PSG. All the participants slept naturally and reported no issues with the attachment of The Patch.

2.3 | Feature extraction

The analyses were performed using Matlab (2018b, MathWorks) software. To remove the low and high frequency noises while preserving breathing sounds (Beck et al., 2005) and snoring sounds (Lee et al., 2016), the recorded tracheal sounds were bandpass filtered using a zero-phase fifth-order Butterworth filter (70–2000 Hz; Montazeri Ghahjaverestan et al., 2021; Yadollahi & Moussavi, 2006, 2010). Then, as this range of frequencies included breathing sounds (Yadollahi & Moussavi, 2006) while overlapping with snoring sounds, we used an algorithm based on wavelet filtering to remove high energy patterns related to snoring in different levels of spectral decomposition (Montazeri Ghahjaverestan et al., 2020). Next, the raw tracheal movements in cranial (X) and posteroanterior (Z) directions were bandpass filtered with a bandwidth of 0.1–0.35 Hz to extract respiratory-related movements. From the filtered signals, a feature related to the frequency of respiratory event occurrence and a group of snore-related features were extracted.

The thoracoabdominal range of movement (TARM) was estimated from the filtered tracheal sounds and movements. Moreover, the airflow level (AFL) was estimated from tracheal sounds. The algorithms to estimate TARM and AFL signals, as the two estimates of tidal volume, have been validated during overnight studies and shown to be significantly correlated with the PSG-based references (Montazeri Ghahjaverestan et al., 2021). Figure 1 depicts a segment of the tracheal data and the related reference airflow and the thoracoabdominal range of movement of an individual with severe sleep apnea. This figure shows the similar fluctuations between the thoracoabdominal sum movement extracted from the PSG and the movements recorded over the trachea in the X and Z axes. Moreover, it shows the changes in the sound energy reflecting the changes in the level of the reference airflow. Using our previous algorithm (Montazeri Ghahjaverestan et al., 2021), the estimated TARM and AFL were extracted from tracheal signals. In this study, we used the estimated TARM and AFL as the inputs of a mathematical model for detecting respiratory events. The output of the model was a probability indicating the chance of event occurrence at each time instant. In the regions with a high probability of event occurrence, the presence of an event was verified by analyzing the amount of drop in the tidal volume. Figure 2 illustrates a segment of data and the event probability during a hypopnea episode followed by intermittent apneas. More details on the event detection algorithm are presented in the Supporting Information (section 1). Additionally, the sleeping periods were extracted from the tracheal sounds and movements based on a sleep/wakefulness detection algorithm previously presented by Montazeri Ghahjaverestan et al. (2021). By counting the detected events and normalizing them over the sleeping time, a rough estimate of AHI was obtained as the *EventFrequency* feature.

Snoring episodes were detected by analyzing the tracheal sounds and movements (Supporting Information section 2). From the detected snoring episodes, seven features were extracted, including *Melcepstability*, *RunningVariance*, average and standard deviation of *Pitch* and *PitchDensity*.

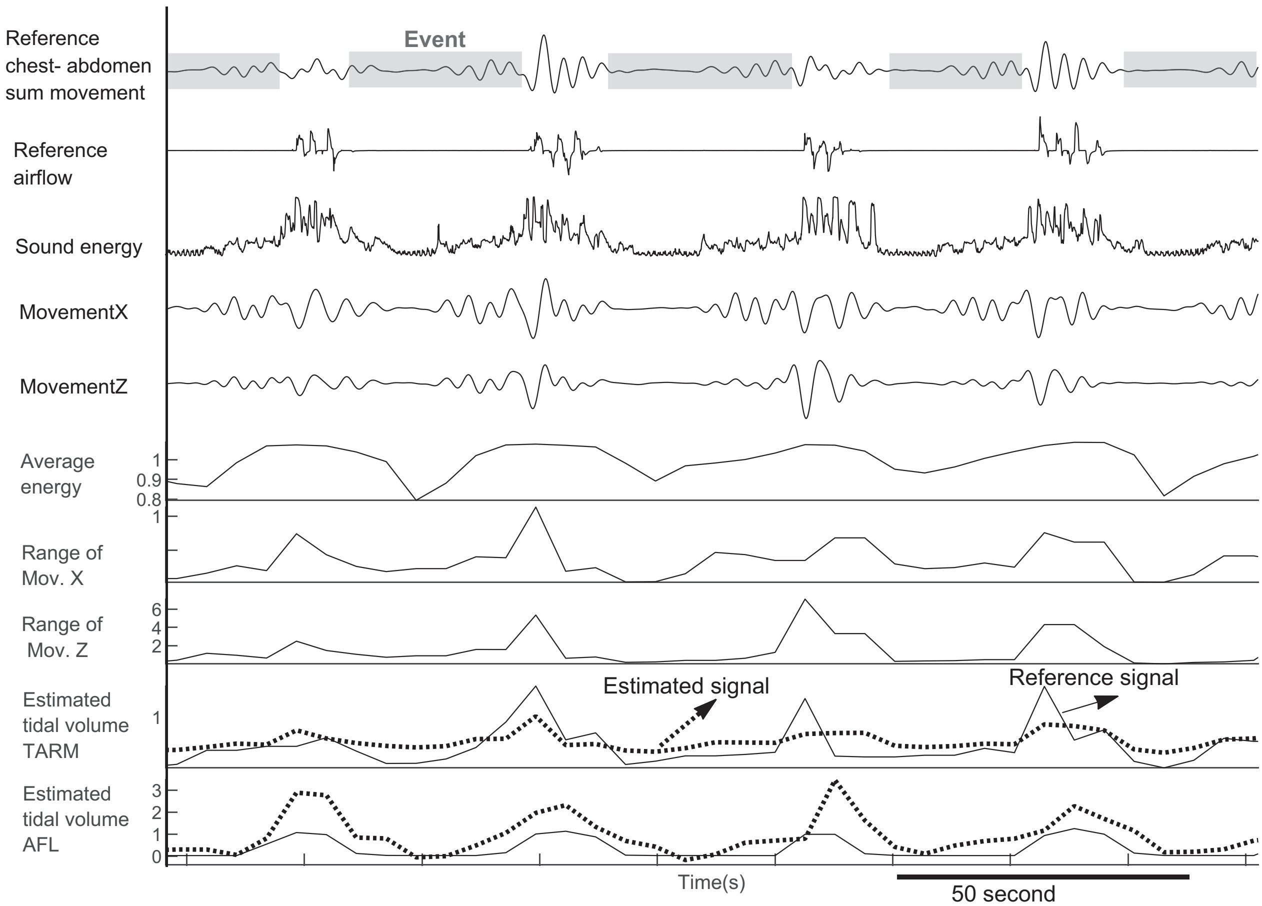

FIGURE 1 An example of the recorded thoracoabdominal sum movement, the related airflow, and synchronized recording of tracheal sound energy and movements in *X* and *Z* directions in a subject with severe sleep apnea with AHI = 83.7. Average energy and range of movements in *X* and *Z* directions were extracted from tracheal signals and used for estimating the tidal volume based on the thoracoabdominal range of movement (TARM). The lower envelope of the sound energy was extracted to estimate the tidal volume based on airflow level (AFL). The last two panels show the reference tidal volume TARM and tidal volume AFL traced with line and the corresponding estimated signals in dotted line

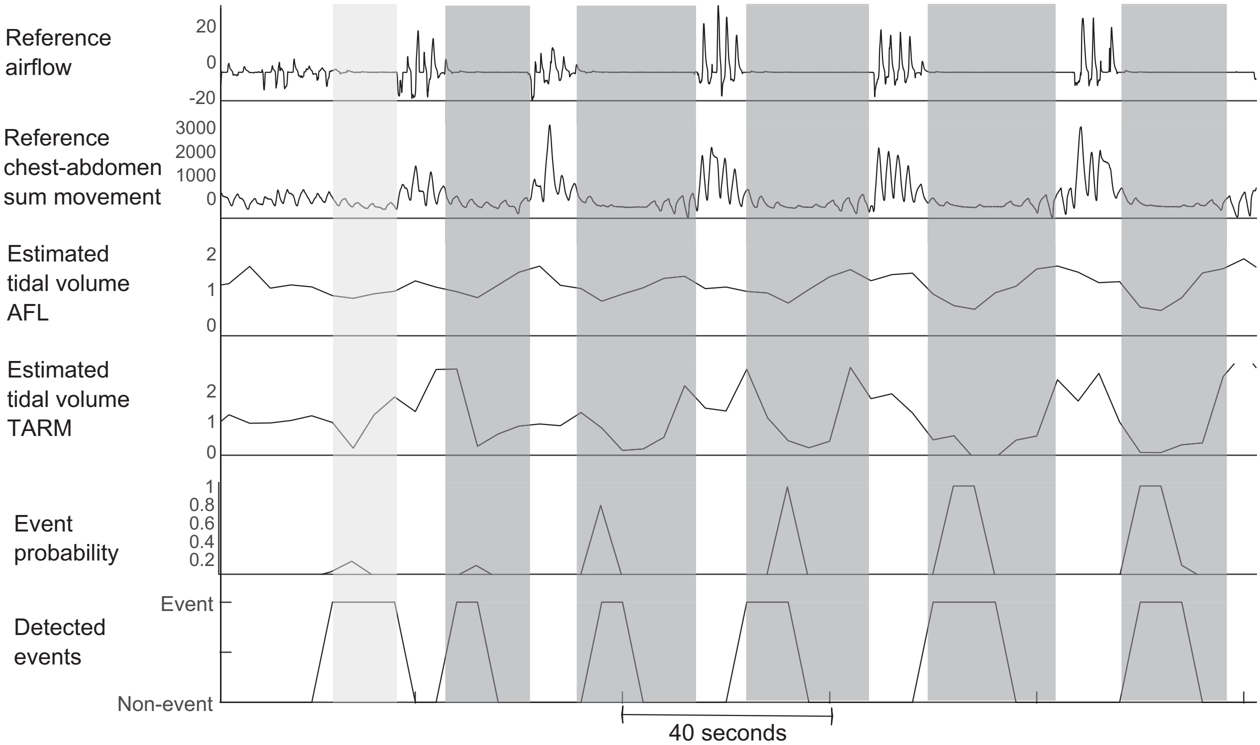

FIGURE 2 An example of event detection performance. The detected events are marked in the regions with higher event probability and drop in the estimated tidal volume signals. Light and dark-shaded regions highlight hypopnea and apnea, respectively

Melcepstability indicates how the snoring patterns are similar in a subject (Ben-Israel et al., 2012). *RunningVariance* was a feature for presenting the variability in the average energy of snores, which were temporally clustered in the overnight data. *Pitch* quantified the dominant frequency of the snores, while *PitchDensity* indicated the portion of snoring where the related Pitch frequency was detectable. A detailed definition of the features is presented in Supporting Information (section 3).

2.4 | Feature selection

The following analysis was performed in R (i386 3.4.1) software. To find the best subset of features to estimate AHI, feature selection was performed in two stages. In the first stage, a step-wise regression (Murphy, 2012) found a subset of features that mostly contributed by predicting AHI variations. In the second stage, different combinations of the selected features were analyzed (greedy analysis) to enhance classifying subjects into healthy and those with sleep apnea based on the AHI cut-off of 15.

2.5 | AHI estimation

After extracting the best combination of the features in the second stage, a multivariate regression model was developed (model$_{\text{Freq-Snr}}$). The performance of the model$_{\text{Freq-Snr}}$ in estimating AHI and classifying participants was compared with another model, which was trained only by *EventFrequency* (model$_{\text{Freq}}$) to quantify the improvement achieved by snoring features.

2.6 | Statistical analysis

The performance of the event-detection algorithm was quantified by reporting the sensitivity, precision, and F1-score for different sleep apnea severity groups. An event was marked as true-positive if there was an overlap between the detected and reference annotations. Undetected reference events and normal segments which were mistakenly detected as events were considered as false-negative and false-positives, respectively.

Using the *EventFrequency* and snore-related features, model$_{\text{Freq-Snr}}$ and model$_{\text{Freq}}$ were developed using a training set randomly selected as 2/3 of the data (40 subjects). Then, the models were validated on the remaining 1/3 of the data (21 subjects) as the validation set. An unpaired *t*-test was used for the statistical comparison of demographics between the training and validation sets. To evaluate the models on a larger training set, another validation process was performed by the leave-one-subject-out (LOSO) process (Wong, 2015). In both validation processes, the estimated AHI was used for classifying each participant as individuals with or without sleep apnea based on different AHI cut-offs of (10, 15, 20, and

30), for which accuracy, sensitivity, and specificity of the classification were reported.

Pearson correlation was used to analyze the agreement of *EventFrequency* and snoring features with PSG-based AHI. For those features presenting a non-linear relationship with AHI, a transform was applied based on the training set. Finally, Pearson correlation and Bland-Altman plots were used to compare the estimated AHI with the PSG-based AHI. For further analysis, the contingency table for classification of the participants into four groups of healthy (AHI < 5), and individuals with mild (5 ≤ AHI < 15), moderate (15 ≤ AHI < 30), and severe (AHI ≥ 30) sleep apnea was reported. Based on the contingency table, the Cohen Kappa score was reported for four-group and two-group classifications (categorized based on AHI cut-off of 15).

3 | RESULTS

Sixty-four participants agreed to participate in this study. Data from three participants were excluded due to the low quality of the recorded tracheal data. Sixty-one participants (30 females, age: 51 ± 16 years, BMI: 29.5 ± 6.4 kg/m^2, AHI: 20.2 ± 21.2 event/h) were included. The participants' characteristics and statistical comparison between the training and validation sets are presented in Table 1. No significant difference was found between the training and validation sets for all the demographics.

The performance of the event-detection algorithm is reported for different sleep apnea severity groups in Table 2. The event detection showed the best performance (precision of 58.53%) for the group with AHI ≥ 30 and lower precision for groups with lower AHI. The example traces in Figure 3 demonstrate the overall agreement between the detected and reference events for a healthy subject and a subject with severe sleep apnea. The *EventFrequency* calculated from the detected events was significantly correlated with PSG-based AHI (Figure 4, $R^2 = 0.67$, $p < 0.001$). Furthermore, there were significant correlations between PSG-based AHI and $Pitch_{std}$, $PitchDensity_{mean}$, $PitchDensity_{std}$, and *Melcepstability* (Figure 5). For *Melcepstability*, a nonlinear relationship was observed. Therefore, using the training set, an exponential transformation with $100 \times \exp(-0.076 \times Melcepstability) - 3.35$ was fitted for *Melcepstability* (Figure 5a, $R^2 = 0.23$, $p < 0.001$).

Table 3 shows the result of step-wise regression analysis. Sequentially, *EventFrequency*, *Melcepstability*, $PitchDensity_{mean}$, *RunningVariance*, and $PitchDensity_{std}$ were chosen as predictors of AHI with an overall $R^2 = 0.81$, $p < 0.001$. The final subset of features set from greedy research included: *EventFrequency*, *Melcepstability*, *RunningVariance*, and $PitchDensity_{std}$ with $R^2 = 0.79$ ($p < 0.001$) and accuracy of 86.50% for classifying participants using the cut-off of 15 in the training set. Table 4 presents additional results for different AHI cut-offs on the training and validation sets obtained from $model_{Freq}$ (Table 4A) and $model_{Freq\text{-}Snr}$ (Table 4B). Using $model_{Freq\text{-}Snr}$, the correlations between estimated and PSG-based AHI for the training set, validation, and LOSO are depicted in Figure 6. For cutoff 15, the validation accuracy of 77.19% ($R^2 = 0.77$, $p < 0.001$) and LOSO accuracy of 83.61% ($R^2 = 0.76$, $p < 0.001$) were obtained. The contingency table related to LOSO cross-validation is shown in Table 5. Based on this table, Cohen Kappa scores of 23.4% and 57.2% were obtained for the four-group and two-group classification, respectively.

Similar analyses were performed for estimating OAHI, for which *EventFrequency*, $PitchDensity_{std}$, *Melcepstability*, $PitchDensity_{mean}$, *RunningVariance*, and $Pitch_{mean}$ were selected with $R^2 = 0.66$ ($p < 0.001$). More results for OAHI estimation have been provided in the Supporting Information (Section 4).

4 | DISCUSSION

The main contributions of this research are to present that: (i) the tidal volume extracted from the overnight tracheal sounds and movements can estimate the frequency of respiratory events that was significantly correlated with PSG-based AHI, and (ii) AHI estimation can be improved by using snoring features extracted from

TABLE 1 Participants' characteristics

Characteristics	Total ($N = 61$)	Training ($N = 40$)	Validation ($N = 21$)	p-Value
Male (female)	31 (30)	19 (21)	12 (9)	0.79
Body mass index (kg/m^2)	29.5 ± 6.4	29.9 ± 6.8	28.7 ± 5.6	0.64
Age (years)	51 ± 16	48 ± 16	57 ± 15	0.64
Epworth sleepiness score	8 ± 4	7 ± 4	8 ± 5	0.62
AHI (events/h)	20.2 ± 21.2	18.2 ± 19.4	24.2 ± 24.1	0.30
OAHI (events/h)	16.9 ± 19.3	15.0 ± 16.2	20.5 ± 24.2	0.56
CAHI (events/h)	2.3 ± 6.4	1.7 ± 3.4	3.6 ± 9.8	0.30
Sleep efficiency (%)	75.9 ± 16.2	77.6 ± 14.2	72.8 ± 19.6	0.47
Total sleeping time (min)	307 ± 88	303 ± 86	315 ± 94	0.52
Detected snoring episodes	2515 ± 1040	2504 ± 944	2535 ± 1228	0.91

Note: Data are presented as mean ± standard deviation.

Abbreviations: AHI, Apnea/Hypopnea Index; CAHI, Central AHI; OAHI, Obstructive AHI; TST, total sleep time.

TABLE 2 Event-detection performance for different apnea-hypopnea index (AHI) groups

AHI groups	*N*	Precision	Sensitivity	F1 score
AHI < 5	19	6.96 ± 8.25	23.97 ± 20.22	9.16 ± 8.60
5 ≤ AHI < 15	12	24.37 ± 11.21	25.71 ± 9.33	23.09 ± 7.60
15 ≤ AHI < 30	16	36.01 ± 18.81	23.7 ± 9.96	27.35 ± 12.65
30 ≤ AHI	14	58.53 ± 20.06	35.56 ± 12.08	42.91 ± 13.5

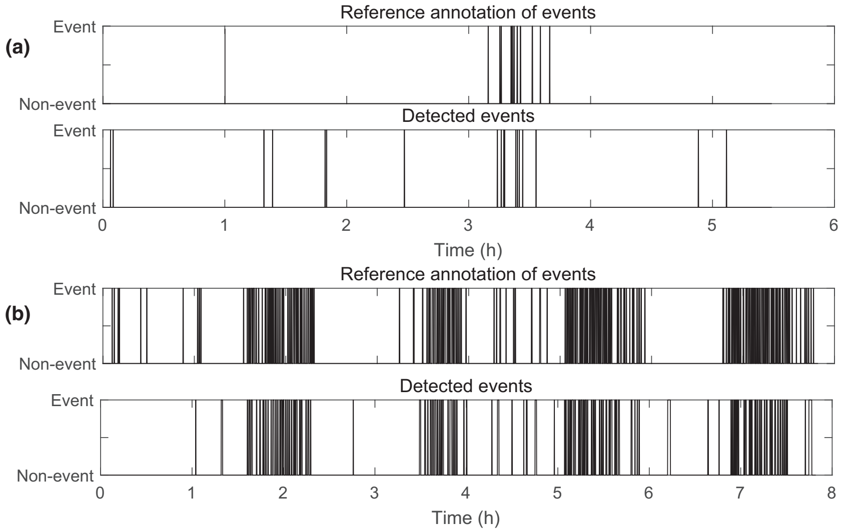

FIGURE 3 Example traces for comparing the detected respiratory events compared with the related reference annotations extracted from PSG for two subjects with A: AHI < 5 and B: AHI ≥ 30

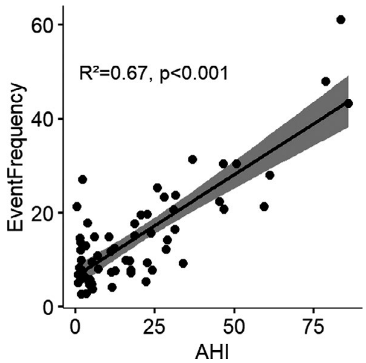

FIGURE 4 The assessment of agreement between *EventFrequency* and PSG-AHI

tracheal sounds. The estimated AHI presented a strong correlation of $R^2 = 0.76$ ($p < 0.001$) with the PSG-AHI for LOSO cross-validation. For the AHI cutoff of 15, the AHI estimation algorithm was able to differentiate individuals with and without sleep apnea with 86.67% sensitivity (no sleep apnea) and 80.65% specificity (sleep apnea).

To extract *EventFrequency*, surrogates of tidal volume were estimated from tracheal signals. Measurement of the absolute tidal volume from PSG requires the airflow and respiratory movements, known as gold standard signals for assessing respiration (Berry et al., 2012), to be calibrated. However, the calibration of gold standard signals is an inconvenient and tedious process during the overnight study. Therefore, we previously proposed two algorithms to extract surrogates of tidal volume: 1, the estimate of TARM from tracheal sounds and movements, and 2, the estimate of AFL from tracheal sounds (Montazeri Ghahjaverestan et al., 2021). In the same study, it was shown that the estimated tidal volumes, similar to the gold standard tidal volume, decreased significantly during apneas and hypopneas compared with normal breathing. This provided a proof of concept that the estimated tidal volume signals can be used for detecting apneas/hypopneas, and ultimately, estimating AHI. Therefore, in this study, the regions with a high chance of apnea/hypopnea occurrence were analyzed to extract the *EventFrequency* feature from estimated tidal volumes. Based on the step-wise regression results, *EventFrequency* had the highest contribution in predicting the variability observed in AHI ($R^2 = 0.67$).

A more accurate estimation of AHI was obtained by including the tracheal snoring sound features. Recent studies have investigated the possibility of assessing OSA using the features extracted from snoring sounds (Abeyratne et al., 2013; Ben-Israel et al., 2012; Herath et al., 2015; Karunajeewa et al., 2010). For OSA analysis using snoring, the features such as *Melcepstability*, *RunningVariability*, and *PitchDensity* were proposed by Ben-Israel et al. (2012). They recorded the snoring sound overnight using an ambient microphone and successfully estimated OAHI with $R^2 = 0.81$ ($p < 0.001$). These features may be less sensitive to snore intensity, which can be affected by the severity of OSA as well as body posture (Nakano et al., 2003). Moreover, the limitations of detecting AHI based on only snoring sounds are that this method cannot estimate AHI in

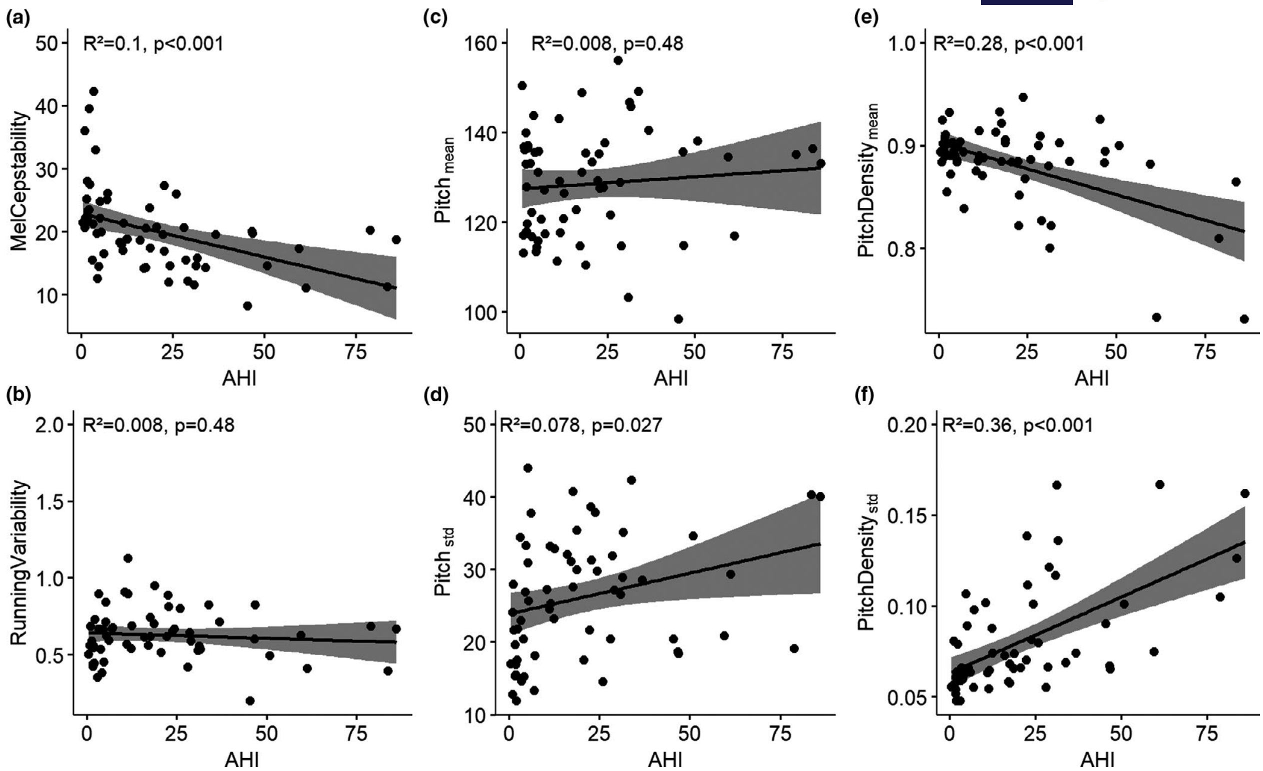

FIGURE 5 The assessment of agreement between A: Melcepstability, B: RunningVariance, C, D: Pitch_{mean} and Pitch_{std}, and E, F: $\text{PitchDensity}_{mean}$ and $\text{PitchDensity}_{std}$ with the PSG-AHI

TABLE 3 Results of step-wise regression analysis to extract the most important features to estimate apnea-hypopnea index (AHI)

Variable entered	Step	Partial R^2	Model R^2	Estimate[a]	*p*-Value[a]
EventFrequency	1	0.67	0.67	1.28	<0.001
Melcapstability	2	0.08	0.74	0.86	<0.001
$\text{PitchDensity}_{mean}$	3	0.04	0.78	−168.28	0.02
RunningVariance	4	0.02	0.80	18.81	0.03
$\text{PitchDensity}_{std}$	5	0.01	0.81	−115.22	0.29

[a]Extracted from the final model.

non-snorer individuals or in those with central sleep apnea caused by the lack of respiratory drive from the brain (White, 2005).

In contrast, in this study, snoring sounds were recorded using a one-directional microphone attached to the trachea, reducing the ambient noises. Moreover, the snore-related features were employed along with *EventFrequency* (based on respiratory patterns) to assess sleep apnea in a population including snorers, non-snorers, healthy subjects, and those with different types of sleep apnea. By including snore-related features in AHI estimation, we improved the agreement with reference AHI by 0.12 ($R^2 = 0.79$) on the training set.

The results on analyzing agreement between the extracted features and the PSG-based AHI demonstrated higher variations in snore-related features for higher AHI, especially in $\textit{PitchDensity}_{std}$ and $\textit{Pitch}_{std}$. Similar results were reported previously where it was found that the variability in snoring sounds was associated with the severity of OSA (Azarbarzin & Moussavi, 2013). These results support the fact that higher AHI is characterized by more instability and collapsibility in the upper airway (Malhotra & White, 2002).

One limitation of this study is related to the potential bias to the snore detection algorithm and the annotator's perception through the procedure of validating the algorithm by manually annotating the snoring episodes. To reduce the bias to snore detection, we trained the algorithm with a large number of snores (11,000 episodes). Another limitation is the small sample size that could be addressed in the future by recruiting more participants. Nonetheless, we evaluated the performance of the classification method based on LOSO cross-validation, which is used in studies with limited sample size. Finally, although accurate and reliable detection and classification of respiratory events as obstructive or central during sleep are crucial concerns for treatment decisions in patients with sleep-disordered breathing, however, this study was not designed to identify the type

TABLE 4 The performance of classifying subjects into healthy and sleep apnea patients, using A: model$_{Freq}$, B: model$_{Freq\text{-}Snr}$

	Hold-out						Leave-one-subject-out		
	Regression training (*N* = 40)			Validation (*N* = 21)			(*N* = 61)		
AHI cut-off	ACC	SEN.	SPC	ACC	SEN	SPC	ACC	SEN	SPC
A									
5	70.0	92.59	23.08	71.43	93.33	16.67	68.85	92.86	15.79
10	70.0	68.18	72.22	66.67	78.57	42.86	67.21	69.44	64.0
15	80.0	61.11	95.45	76.19	75	77.78	78.69	66.67	90.32
20	82.50	57.14	96.15	76.19	60.00	90.91	80.33	58.33	94.59
30	85.00	25.00	100	90.48	66.67	100	86.89	42.86	100
B									
5	77.50	96.30	38.46	76.19	93.33	33.33	77.05	92.86	42.11
10	87.50	100	72.22	80.95	92.86	57.14	80.32	91.67	64.0
15	85.0	88.89	81.81	77.19	91.67	55.56	83.61	86.67	80.65
20	92.5	92.86	92.31	85.71	90.0	81.82	85.25	83.33	86.49
30	92.5	75.0	96.88	85.71	83.33	86.67	86.89	78.57	89.36

Abbreviations: ACC, accuracy of the classification; SEN, sensitivity; SPC, specificity.

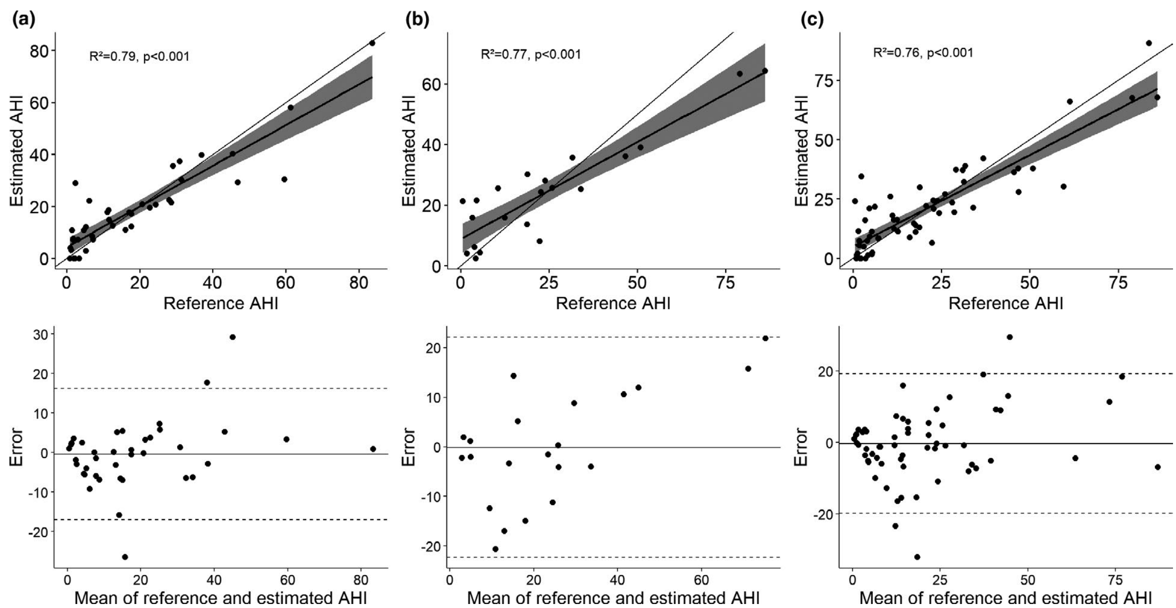

FIGURE 6 The correlation and the related Bland-Altman plots between estimated AHI and PSG-based AHI in, A: Training set, B: Validation set, and C: Leave-One-Subject-Out using model$_{Freq\text{-}Snr}$

TABLE 5 Contingency table for classification of participants into healthy people and those with mild, moderate and severe sleep apnea

	Reference			
	Healthy (AHI < 5)	Mild (5 ≤ AHI < 15)	Moderate (5 ≤ AHI < 15)	Severe (5 ≤ AHI < 15)
Detected				
Healthy (AHI < 5)	3	3	0	0
Mild (5 ≤ AHI < 15)	13	9	9	1
Moderate (15 ≤ AHI < 30)	3	0	7	7
Severe (30 ≤ AHI)	0	0	0	6

of events. The identification of the events will be included in future works.

To conclude, this is the first study of incorporating estimates of tidal volume and snoring sounds extracted from tracheal signals to estimate AHI. This study provides evidence that the acoustic-based wearable device can be used for convenient and robust screening of people with suspected sleep apnea, thus addressing the growing need for developing portable sleep screening devices.

ACKNOWLEDGEMENTS

This study was supported by Ontario Centres of Excellence (OCE), NSERC Discovery grant, and BresoTEC Inc. Toronto, ON, Canada.

CONFLICT OF INTEREST

Dr Yadollahi reports grants from NSERC Discovery, grants from OCE VIPII, grants from NSERC Collaborative Research Development, grants from Ontario Ministry of Research and Innovation, New Investigator Award, during the conduct of the study; In addition, Dr Yadollahi has a patent Acoustic System and Method for Performing Pharyngeal and/or Airway Assessment, Patent Number: US20170119303A1 pending, a patent System and methods for estimating respiratory airflow- Patent number: US10004452B licensed to BresoTec, and a patent Breathing Analysis for Detection of Sleep Apnea/Hypopnea Events – Patent number: US7559903 B2 licensed. Other co-authors and the first author do not have any conflict of interest to disclose.

AUTHOR CONTRIBUTIONS

All authors have reviewed and approved the manuscript. The contributions of the authors are as follows: Nasim Montazeri: algorithm development, statistical analysis, writing the manuscript. Muammar Kabir: algorithm development. Shumit Saha, Kaiyin Zhu: data collection. Bojan Gavrilovic: device development. Azadeh Yadollahi: supervising the research, revising the manuscript.

DATA AVAILABILITY STATEMENT

Research data are not shared.

ORCID

Nasim Montazeri Ghahjaverestan https://orcid.org/0000-0003-0960-9632

REFERENCES

Abeyratne, U., De Silva, S., Hukins, C., & Duce, B. (2013). Obstructive sleep apnea screening by integrating snore feature classes. *Physiological Measurement*, *34*(2), 99. https://doi.org/10.1088/0967-3334/34/2/99

Abeyratne, U. R., Wakwella, A. S., & Hukins, C. (2005). Pitch jump probability measures for the analysis of snoring sounds in apnea. *Physiological Measurement*, *26*(5), 779. https://doi.org/10.1088/0967-3334/26/5/016

Azarbarzin, A., & Moussavi, Z. M. (2010). Automatic and unsupervised snore sound extraction from respiratory sound signals. *IEEE Transactions on Biomedical Engineering*, *58*(5), 1156–1162. https://doi.org/10.1109/TBME.2010.2061846

Azarbarzin, A., & Moussavi, Z. (2013). Snoring sounds variability as a signature of obstructive sleep apnea. *Medical Engineering & Physics*, *35*(4), 479–485. https://doi.org/10.1016/j.medengphy.2012.06.013

Beck, R., Rosenhouse, G., Mahagnah, M., Chow, R. M., Cugell, D. W., & Gavriely, N. (2005). Measurements and theory of normal tracheal breath sounds. *Annals of Biomedical Engineering*, *33*(10), 1344–1351. https://doi.org/10.1007/s10439-005-5564-7

Ben-Israel, N., Tarasiuk, A., & Zigel, Y. (2012). Obstructive apnea hypopnea index estimation by analysis of nocturnal snoring signals in adults. *Sleep*, *35*(9), 1299–1305. https://doi.org/10.5665/sleep.2092

Berry, R. B., Budhiraja, R., Gottlieb, D. J., Gozal, D., Iber, C., Kapur, V. K., Marcus, C. L., Mehra, R., Parthasarathy, S., Quan, S. F., Redline, S., Strohl, K. P., Ward, S. L. D., & Tangredi, M. M. (2012). Rules for scoring respiratory events in sleep: update of the 2007 AASM manual for the scoring of sleep and associated events: deliberations of the sleep apnea definitions task force of the American Academy of Sleep Medicine. *Journal of Clinical Sleep Medicine*, *8*(5), 597–619. https://doi.org/10.5664/jcsm.2172

Cummiskey, J., Williams, T. C., Krumpe, P. E., & Guilleminault, C. (1982). The detection and quantification of sleep apnea by tracheal sound recordings. *American Review of Respiratory Disease*, *126*(2), 221–224.

Davies, C. R., & Harrington, J. J. (2016). Impact of obstructive sleep apnea on neurocognitive function and impact of continuous positive air pressure. *Sleep Medicine Clinics*, *11*(3), 287–298. https://doi.org/10.1016/j.jsmc.2016.04.006

Fletcher, E. C., Schaaf, J. W., Miller, J., & Fletcher, J. G. (1987). Long-term cardiopulmonary sequelae in patients with sleep apnea and chronic lung disease. *American Review of Respiratory Disease*, *135*(3), 525–533.

Gavriely, N., & Cugell, D. W. (1996). Airflow effects on amplitude and spectral content of normal breath sounds. *Journal of Applied Physiology*, *80*(1), 5–13. https://doi.org/10.1152/jappl.1996.80.1.5

Glos, M., Sabil, A., Jelavic, K. S., Baffet, G., Schöbel, C., Fietze, I., & Penzel, T. (2019). Tracheal sound analysis for detection of sleep disordered breathing. *Somnologie*, *23*(2), 80–85. https://doi.org/10.1007/s11818-019-0200-1

Hafezi, M., Montazeri, N., Saha, S., Zhu, K., Gavrilovic, B., Yadollahi, A., & Taati, B. (2020). Sleep apnea severity estimation from tracheal movements using a deep learning model. *IEEE Access*, *8*, 22641–22649. https://doi.org/10.1109/ACCESS.2020.2969227

Herath, D. L., Abeyratne, U. R., & Hukins, C. (2015). Hidden Markov modelling of intra-snore episode behavior of acoustic characteristics of obstructive sleep apnea patients. *Physiological Measurement*, *36*(12), 2379. https://doi.org/10.1088/0967-3334/36/12/2379

Hillman, D., Mitchell, S., Streatfeild, J., Burns, C., Bruck, D., & Pezzullo, L. (2018). The economic cost of inadequate sleep. *Sleep*, *41*(8), zsy083. https://doi.org/10.1093/sleep/zsy083

Kalkbrenner, C., Eichenlaub, M., Rüdiger, S., Kropf-Sanchen, C., Brucher, R., & Rottbauer, W. (2017). Validation of a new system using tracheal body sound and movement data for automated apnea-hypopnea index estimation. *Journal of Clinical Sleep Medicine*, *13*(10), 1123–1130. https://doi.org/10.5664/jcsm.6752

Kalkbrenner, C., Eichenlaub, M., Rüdiger, S., Kropf-Sanchen, C., Rottbauer, W., & Brucher, R. (2018). Apnea and heart rate detection from tracheal body sounds for the diagnosis of sleep-related breathing disorders. *Medical & Biological Engineering & Computing*, *56*(4), 671–681. https://doi.org/10.1007/s11517-017-1706-y

Kapur, V., Strohl, K. P., Redline, S., Iber, C., O'Connor, G., & Nieto, J. (2002). Underdiagnosis of sleep apnea syndrome in US communities. *Sleep and Breathing*, *6*(2), 49–54. https://doi.org/10.1055/s-2002-32318

Karunajeewa, A. S., Abeyratne, U. R., & Hukins, C. (2010). Multi-feature snore sound analysis in obstructive sleep apnea–hypopnea syndrome. *Physiological Measurement*, *32*(1), 83. https://doi.org/10.1088/0967-3334/32/1/006

Kulkas, A., Huupponen, E., Virkkala, J., Tenhunen, M., Saastamoinen, A., Rauhala, E., & Himanen, S.-L. (2009). New tracheal sound feature for apnoea analysis. *Medical & Biological Engineering & Computing*, *47*(4), 405–412. https://doi.org/10.1007/s11517-009-0446-z

Lee, G.-S., Lee, L.-A., Wang, C.-Y., Chen, N.-H., Fang, T.-J., Huang, C.-G., Cheng, W.-N., & Li, H.-Y. (2016). The frequency and energy of snoring sounds are associated with common carotid artery intima-media thickness in obstructive sleep apnea patients. *Scientific Reports*, *6*(1), 1–11. https://doi.org/10.1038/srep30559

Malhotra, A., & White, D. P. (2002). Obstructive sleep apnoea. *The Lancet*, *360*(9328), 237–245.

Medicine, A. A. O. S. (2016). *Hidden health crisis costing America billions. Underdiagnosing and undertreating obstructive sleep apnea draining healthcare system*. Frost & Sullivan. https://j2vjt3dnbra3ps7ll1clb4q2-wpengine.netdna-ssl.com/wp-content/uploads/2017/10/sleep-apnea-economic-crisis.pdf

Montazeri Ghahjaverestan, N., Kabir, M. M., Saha, S., Gavrilovic, B., Zhu, K., Taati, B., & Yadollahi, A. (2021). Relative tidal volume and respiratory airflow estimation using tracheal sound and movement during sleep. *Journal of Sleep Research*, *30*(4), e13279.

Montazeri Ghahjaverestan, N., Saha, S., Gavrilovic, B., & Yadollahi, A. (2020). *Removing of Snoring Segments from Tracheal Breathing Sounds using a Wavelet-based Algorithm*. Paper presented at the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC).

Murphy, K. P. (2012). *Machine learning: A probabilistic perspective*. MIT Press.

Nakano, H., Furukawa, T., & Tanigawa, T. (2019). Tracheal sound analysis using a deep neural network to detect sleep apnea. *Journal of Clinical Sleep Medicine*, *15*(8), 1125–1133. https://doi.org/10.5664/jcsm.7804

Nakano, H., Hayashi, M., Ohshima, E., Nishikata, N., & Shinohara, T. (2004). Validation of a new system of tracheal sound analysis for the diagnosis of sleep apnea-hypopnea syndrome. *Sleep*, *27*(5), 951–957. https://doi.org/10.1093/sleep/27.5.951

Nakano, H., Ikeda, T., Hayashi, M., Ohshima, E., & Onizuka, A. (2003). Effects of body position on snoring in apneic and nonapneic snorers. *Sleep*, *26*(2), 169–172. https://doi.org/10.1093/sleep/26.2.169

Ng, A. K., San Koh, T., Abeyratne, U. R., & Puvanendran, K. (2009). Investigation of obstructive sleep apnea using nonlinear mode interactions in nonstationary snore signals. *Annals of Biomedical Engineering*, *37*(9), 1796–1806. https://doi.org/10.1007/s10439-009-9744-8

Pack, A. I. (2004). Sleep-disordered breathing: access is the issue. *American Journal of Respiratory and Critical Care Medicine*, *169*(6), 666–667. https://doi.org/10.1164/rccm.2401008

Penzel, T., & Sabil, A. (2017). The use of tracheal sounds for the diagnosis of sleep apnoea. *Breathe*, *13*(2), e37–e45. https://doi.org/10.1183/20734735.008817

Penzel, T., & Sabil, A. (2018). Physics and applications for tracheal sound recordings in sleep disorders. In K. Priftis, L. Hadjileontiadis & M. Everard (Eds.), *Breath sounds* (pp. 83–104). Springer. https://doi.org/10.1007/978-3-319-71824-8_6

Sabil, A., Glos, M., Günther, A., Schöbel, C., Veauthier, C., Fietze, I., & Penzel, T. (2019). Comparison of apnea detection using oronasal thermal airflow sensor, nasal pressure transducer, respiratory inductance plethysmography and tracheal sound sensor. *Journal of Clinical Sleep Medicine*, *15*(2), 285–292. https://doi.org/10.5664/jcsm.7634

Saha, S., Kabir, M., Montazeri, N., Gavrilovic, B., Zhu, K., Yadollahi, A. & Alshaer, H. (2019). Apnea-hypopnea index (AHI) estimation using breathing Sounds, accelerometer and pulse oximeter. ERJ Open Research 2019. *5*(suppl 3). European Respiratory Society. https://openres.ersjournals.com/content/5/suppl_3/P63

Tong, J., Jugé, L., Burke, P. G., Knapman, F., Eckert, D. J., Bilston, L. E., & Amatoury, J. (2019). Respiratory-related displacement of the trachea in obstructive sleep apnea. *Journal of Applied Physiology*, *127*(5), 1307–1316. https://doi.org/10.1152/japplphysiol.00660.2018

White, D. P. (2005). Pathogenesis of obstructive and central sleep apnea. *American Journal of Respiratory and Critical Care Medicine*, *172*(11), 1363–1370. https://doi.org/10.1164/rccm.200412-1631SO

Wong, T.-T. (2015). Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognition*, *48*(9), 2839–2846. https://doi.org/10.1016/j.patcog.2015.03.009

Yadollahi, A., Giannouli, E., & Moussavi, Z. (2010). Sleep apnea monitoring and diagnosis based on pulse oximetery and tracheal sound signals. *Medical & Biological Engineering & Computing*, *48*(11), 1087–1097. https://doi.org/10.1007/s11517-010-0674-2

Yadollahi, A., & Moussavi, Z. M. (2006). A robust method for heart sounds localization using lung sounds entropy. *IEEE Transactions on Biomedical Engineering*, *53*(3), 497–502. https://doi.org/10.1109/TBME.2005.869789

Yadollahi, A., & Moussavi, Z. (2008). *Comparison of flow-sound relationship for different features of tracheal sound*. Paper presented at the 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society.

Yadollahi, A., & Moussavi, Z. (2010). Automatic breath and snore sounds classification from tracheal and ambient sounds recordings. *Medical Engineering & Physics*, *32*(9), 985–990. https://doi.org/10.1016/j.medengphy.2010.06.013

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Montazeri Ghahjaverestan, N., Saha, S., Kabir, M., Gavrilovic, B., Zhu, K., & Yadollahi, A. (2022). Sleep apnea severity based on estimated tidal volume and snoring features from tracheal signals. *Journal of Sleep Research*, 31, e13490. https://doi.org/10.1111/jsr.13490