

Received August 17, 2020, accepted September 6, 2020, date of publication September 15, 2020, date of current version September 28, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3024122

Joint Apnea and Body Position Analysis for Home Sleep Studies Using a Wireless Audio and Motion Sensor

MARCEL MŁYŃCZAK¹, (Member, IEEE), TULIO A. VALDEZ², AND WOJCIECH KUKWA³

¹Faculty of Mechatronics, Institute of Metrology and Biomedical Engineering, Warsaw University of Technology, 02-525 Warsaw, Poland

²Department of Otolaryngology, Stanford University, Stanford, CA 94305, USA

³Faculty of Dental Medicine, Medical University of Warsaw, 02-091 Warsaw, Poland

Corresponding author: Wojciech Kukwa (wojciechkukwa@gmail.com)

This work was supported in part by The Bekker Programme funded by the Polish National Agency for Academic Exchange (NAWA) through the Short-Term Scholarship Project, and in part by the National Centre for Research and Development, Poland, through the Fast Track Program, under Grant POIR.01.01.01-00-0212/19.

ABSTRACT This study aims to evaluate the accuracy of a wireless tracheal audio sensor including the sleep body position measurement, as a possible tool to screen for obstructive sleep apnea (OSA) and also to distinguish among “positional” and “non-positional” sleep apnea patients. 30 adult subjects were asked to sleep naturally for a single night in the sleep laboratory, having simultaneous polysomnography (PSG) as a reference. The results were compared using, e.g., Pearson’s and Lin’s correlation coefficients, Bland-Altman plots, mean absolute error, intercept of the fitted linear model etc. We found the thresholding approach performed on normalized audio energy data achieved mean absolute error of 5.7, and average difference was -2.0 . Pearson’s and Lin’s R coefficients were 0.92 and 0.70, respectively. The qualitative approach reached 86% accuracy (sensitivity of 96%, and specificity of 76%) when setting a binary border at the level of $AHI = 15$ (combined normal + mild versus combined moderate + severe cases). We also checked the distribution of apneas depending on the body position. Our sensor showed a mean \pm SE difference between supine apnea index and non-supine apnea index of 8.3 ± 3.2 , while PSG reference of 11.2 ± 3.8 . The proposed sensor might be a good complement for home sleep studies, being less disturbing and allowing for longitudinal observations and reliably showing positional OSA. PSG is the gold standard in diagnosing OSA; however, it requires to spend a night in a lab with medical staff, it provides short observation, the cost of the study is very high, and it is less suitable for children. This is why a reliable screening method is needed in sleep medicine.

INDEX TERMS Obstructive sleep apnea, apnea-hypopnea index, home sleep study, tracheal sounds, audio processing, supine sleep, non-supine sleep.

I. INTRODUCTION

Obstructive sleep apnea (OSA) is a highly prevalent sleep disorder that can cause significant daytime sleepiness and result in many cardiovascular comorbidities [1]–[3]. It is characterized by repetitive upper airway obstruction and significant airflow reductions during sleep resulting in intermittent hypoxia and sleep fragmentation [4]. As a moderate-to-severe OSA can be diagnosed in as many as 49% of adult males and 23% of adult females [5], the disorder should be diagnosed and treated as early as possible.

The associate editor coordinating the review of this manuscript and approving it for publication was Agustin Leobardo Herrera-May¹.

Current sleep studies, full polysomnography (PSG) and abbreviated home sleep apnea tests (HSAT), are labor-consuming and costly. They require to spend a night in a lab with medical staff, provide short observation, and are less comfortable and available for children. That is why a simple method is needed to address this problem. One of the most promising methods of simple OSA detection is an audio-based approach. As OSA’s pathophysiology is the upper airway obstruction, the sound of breathing carries a good load of data regarding breathing disorders during sleep [6]. In the last few years, studies on breathing sounds were performed using ambient, together with built-in smartphone microphones, or contact microphones [7]–[12]. The highest accuracy of breathing episodes detection can

be obtained from contact-microphones located on the neck, where the tracheal sounds are recorded.

Interestingly, different sleep positions can alter the breathing sounds, and change the severity of sleep apnea in many patients. In most patients referred to sleep clinics, the severity of sleep apnea increases while sleeping supine. Up to 60% of patients have an apnea-hypopnea index twice as high sleeping supine – this situation is called positional obstructive sleep apnea (pOSA) [13], [14]. That is why the simultaneous recording of sleep position is another important issue in newly developed sleep sensors. It might be even more important in a personal home-based sleep study that can also be used as a sleep position trainer.

In a previous paper, we introduced the wireless acoustic sensor system communicating with a smartphone application that can be used for screening breathing disorders during sleep [15]. We showed high accuracy in the automatic detection of normal breathing and snoring episodes. We currently developed a multi-channel wireless sensor with two acoustic channels and accelerometry unit to measure body position and activity during sleep (see Methods).

A similar system has been recently described by *Saha et al.* in a paper, where a sensor combined of microphone and accelerometer was used [16]; however, the accelerometer was used for respiratory related movements detection (to validate apneic and hypopneic events detected with audio analysis), not to combine audio with body position for positional OSA analysis. Also, *Kalkbrenner et al.* used a tracheal microphone connected with an accelerometer, but here the accelerometer was separately mounted on the chest, and not in a single wireless sensor [17]. Both studies used accelerometry signals and acoustic recordings to detect breathing episodes, apneas, and hypopneas. However, the authors did not address the problems of detecting supine and non-supine sleep apnea. To our knowledge, only a paper by *Levendovsky et al.* was aimed at positional sleep breathing disorders, using a system mounted on the back of the neck. Still, it only analyzed snoring intensity without detecting apneas and hypopneas [18].

This study aimed to show the feasibility of a wireless sensor that records and analyzes tracheal breathing sounds and sleep position and actigraphy to analyze positional sleep apnea. We hypothesized that adding the sleep position data could benefit from getting more information on a patient's sleep. We could measure this way, similar to a regular sleep study, the supine and non-supine sleep breathing disorders.

II. MATERIALS & METHODS

A. PARTICIPANTS

30 adult subjects participated in the study. The characteristics of the group is presented in Table 1.

B. PROTOCOL AND DEVICES

A technician asked all participants to perform a single full polysomnography in the sleep laboratory of Otorhinolaryngology Department at Czerniakowski Hospital, Warsaw,

TABLE 1. The characteristics of the study group; all results presented as median(IQR; min – max).

Sex	Female	Male
N	5	25
AHI [events/h]	10.4 (7.9; 2.2 – 19.4)	29.3 (32.2; 1.1 – 65.4)
Snoring [%]	44.0 (47.4; 0.4 – 59.5)	32.2 (35.0; 0.0 – 67.2)
Age [years]	56 (7; 30 – 59)	47 (18; 26 – 72)
Height [cm]	168 (6; 158 – 172)	176 (13; 168 – 196)
Weight [kg]	77 (9; 48 – 85)	95 (18; 69 – 119)
BMI [kg/m ²]	26.0 (3.7; 17.0 – 31.2)	29.3 (3.1; 23.8 – 38.9)

Poland. The data were recorded using the Nox A1 PSG System, a full and portable polysomnography system, created by Nox Medical, Iceland. During the analysis several exported data and report details were taken into account:

- raw accelerometry data, sampled with 20 Hz,
- estimated body position and activity data, sampled with 200 Hz,
- AHI calculation as single one-night parameter, and
- beginnings and ends of central or obstructive apneas.

During the recordings, the audio and motion sensor was placed in the suprasternal notch on the neck, to record tracheal sounds. It enables recording of two audio signals, from the body and from the field, using digital microphones with 8000 Hz sampling frequency, and 3-axis motion signals - using accelerometer with 52 Hz sampling frequency. An accelerometer enables recording of actigraphy - body position and subject's activity. The device does not have the CE mark yet, and it is not available on the market; however, it does not prevent from results replication, as the sensor comprises generally available audio and motion sensors. The dimensions of the sensor are 33 x 39 x 13 mm, the diameter of the membrane is 21.8 mm, the sensor weight is 18 g. The current battery allows over 14h of operation. The memory capacity is defined internally by 2GB FLASH chip or can be extended by external microSD card. Computations are performed on 32-bit ARM Cortex-M4 micro-controller with DSP dedicated unit, 128 kBRAM, and running at 64 MHz.

The recommended placement is on the neck (suprasternal notch). The sensor can be attached using a medical double-sided patch, which is more comfortable than a band holding it around a neck. Based on the questionnaire reports from the users, it stayed well in place during the entire night, did not disturb natural sleep, and did not introduce any discomfort, even during momentary awakenings at night. Nobody refused to use the sensor in the future. Its photography is presented in Figure 1, and the placement on a neck in Figure 2.

The sensor has an internal storage, enabling to save raw data. All algorithms will be implemented on-chip, allowing real-time analysis during sleep (not affecting the battery life much - it was already tested) and presentation of the results on smartphone screen just after waking up. Therefore, the main setup is that only the results of the analysis is transmitted wirelessly (using Bluetooth protocol) to the mobile app.

The procedure was approved by the Ethics Committee of Medical University of Warsaw (KB/14/2018).



FIGURE 1. The photography of the sensor; outer part on the left, and inner part (directed to the body) on the right.



FIGURE 2. The placement of the sensor on a neck.

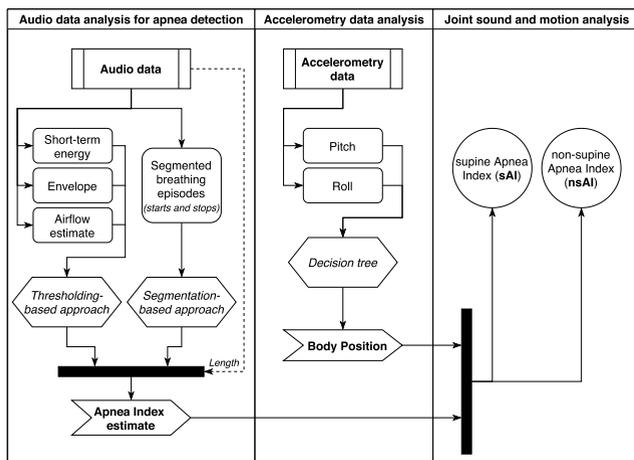


FIGURE 3. The conceptual scheme diagram of the audio- and motion-based analysis.

All participants were informed about the general aims of the measurements, and each had signed a general consent form for the routine medical monitoring, consisting of a statement of acceptance of the use of the results for scientific purposes, before the study.

C. AUDIO DATA ANALYSIS FOR APNEA DETECTION

The conceptual diagram presenting the flow of analysis in the II.C-II.E sections is shown in Figure 3.

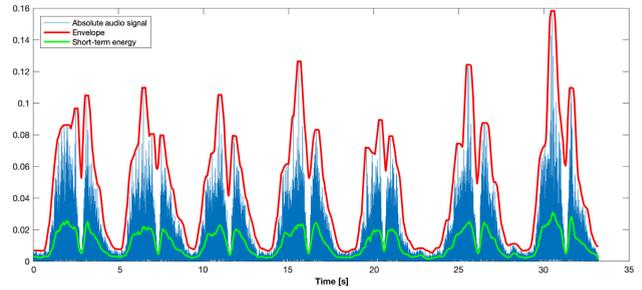


FIGURE 4. The sample audio signal (absolute) with detected envelope and short-term energy signals; Y-axis units are arbitrary.

PSG recording, treated as a reference study was evaluated by a sleep physician using the standard criteria defined by the American Academy of Sleep Medicine after the examination, and based on automatic initial analysis. However, the exact synchronization between PSG and the audio and motion sensor could not be obtained. Therefore, we focused on the indirect approach. Several methods were used to estimate the number of apneas (all central, obstructive and mixed, next divided by total sleep time) from the audio signal. Then, the value was compared with the number of apneas detected in PSG study, also divided by the duration of sleep. In the next step we are going to add hypopnea analysis to get the entire AHI parameter, as the synchronization is reached.

Before the analysis, we performed the pre-processing of the audio data. We took into account only the microphone measuring from the body, as the ambient one is rather used both for snoring and subject’s activity recording, and for a signal to noise ratio reduction. First, we filtered the sounds within 100 – 1800 Hz frequency range. Then, the 2-fold decimation - to 4000 Hz of sampling frequency, was carried out. Next, 3 signals were calculated/detected:

- envelope of the absolute sound;
- short-term energy; and
- airflow estimate.

The first ones were calculated using rolling maximum value within a window of 1/3 second, then decimated to 50 Hz and properly filtered - smoothed; using low-pass FIR filters. The second ones were estimated within a short 16ms windows (64 probes in each) and then decimated in the same way as for the envelope. The last one comes from the short-term energy signal and it is arbitrarily calculated as its square root, which changes the dynamics of the signal - lower values are amplified in comparison to the higher one, which makes the characteristics similar to airflow curve.

The example of an envelope and short-term energy signals calculated for the absolute audio signal is presented in Figure 4.

We decided to implement and evaluate two approaches: thresholding- and segmentation-based.

The first one comes from the consideration presented by Kalkbrenner at el. [17] and based on the AASM guidelines [19]. It assumes, that the apnea is found, when the envelope of the signal is lower than an established threshold

for at least 10-seconds. As it is too arbitrary, we decided to evaluate, in reference to the PSG, all three signals, with or without adaptation of the level of normal breathing, with or without normalization of the data within long segments, and with various threshold levels. We checked the accuracies between 2% and 15% - by 0.5% - of the normal breathing level.

The adaptation is performed in such a way, that the maximum level of normal breathing is established for the last 30 minutes (this level is taken for threshold calculation and is introduced to adjust to the nature of the signal and become more independent of possible artifacts), and not for the entire recording, as the amplitude may vary according to the body position. The normalization of the audio amplitude is based on the same consideration and it is performed in regard to the maximum value within the last 30 minutes.

The second approach is based on the previously published algorithm for segmentation breathing episodes and classifying them as normal breathing or snoring [15]. If there is an interval greater than 10 seconds between subsequent detected episodes (after removing possible “crackles” and noises), we may count in the apnea indexes.

In order to perform comparative analysis, the Shiny web app (with R-language-based calculations in the background), described in [20] is used. Here, we reported classical Pearson’s correlation coefficient (*Pearson’s R*), Lin’s concordance coefficient (*Lin’s R*), intercept of the linear model best fitted to the data points (*Intercept*), mean absolute error (*MAE*), the average difference (*AD*; similar to bias from the Bland-Altman diagram), and limits of agreement from the Bland-Altman diagram). Pearson’s coefficient is reported as a preliminary illustration; the discussion about its possible disadvantages is presented in [20]. Lin’s version is treated as a more suitable index of reliability.

D. ACCELEROMETRY DATA ANALYSIS

Probably all commercial devices operate within four main sleeping body positions - supine, prone, left side, and right side. Therefore, we prepared a simple algorithm of body position calculation based on the pitch and roll estimation from raw 3 axes accelerometry registrations (input information). The equations to calculate the position of the sensor compared to the gravity, are as follows (1-2):

$$\text{Roll}[\text{rad}] = \arctan\left(\frac{Y}{X}\right) \quad (1)$$

$$\text{Pitch}[\text{rad}] = \arctan\left(\frac{-Z}{\sqrt{X^2 + Y^2}}\right) \quad (2)$$

where X-Z are acceleration in all axes, respectively.

Then, the body position (output information) is classified using a simple hierarchical tree. First, when Pitch is less than 75 degrees the position is considered as lying, and the roll is taken to determine one of the four lying positions by dividing the circumference of 360 degrees into 4 parts and small breaks of 2 degrees each to highlight the changes and uncertainty.

The simple activity measure is calculated by estimating the accelerometry “vector”, as in the Equation (3):

$$\text{ACC} = \sqrt{X^2 + Y^2 + Z^2} \quad (3)$$

where ACC is an accelerometry vector, and X-Z have the same meaning as in Equations (1-2).

Then, the constant component is being removed, and the high activity is set, when an arbitrarily set threshold is exceeded. As the PSG device enables to record 3-axis accelerometer signals and stores body position and subject’s activity measures, we used them to evaluate the accuracy of the implemented simple algorithms.

E. JOINT SOUND AND MOTION ANALYSIS

As the used sensor has both sound and accelerometry sensors, we may calculate joint indexes providing new clinical context. Therefore, we use two coefficients:

- sAI - supine Apnea Index, and
- nsAI - Apnea Index during all non-supine body positions (prone, sides - as these are usually less impactful on the number of apneas).

Both are calculated as the sum/number of detected apneas occurring during supine or non-supine sleeping body position respectively, then divided by the total recorded supine or non supine sleeping time. We hope that splitting the analysis into two various conditions (supine- and non-supine-related) enables to estimate the impact of the position on the number of apneas; and then to help to reduce the overall AI metrics by a positional training. Particularly, when sAI is definitely bigger than nsAI, and nsAI is relatively low - then, reducing sleeping time being supine can help to breathe better at night.

F. MACHINE-LEARNING-BASED AUDIO ANALYSIS FOR QUALITATIVE AHI ESTIMATION

In the last step, we implemented non-linear-feature-based approach, that comes from the idea presented in the paper [21]. It relies on estimating parameters, e.g.:

- Sample Entropy;
- Approximate Entropy;
- Q3 of the absolute signal first derivative, treated as evaluation of signal variability; and
- S Transform (simpler equivalent to FFT, for indirect spectral analysis),

within 10-second-lasting windows, with 50% overlapping. And then, quantile distribution of all parameters was taken as the input for classification, as a strictly qualitative approach was considered.

The entire AHI (not only apneas) was assessed in four ranges - normal ($AHI < 5$), mild ($5 \leq AHI < 15$), moderate ($15 \leq AHI < 30$), and severe ($AHI \geq 30$) - based on the medical guidelines for adults [22].

We decided not to implement direct spectral features, e.g. coming from MFCC analysis or just FFT calculations, in order to accelerate the analysis to be performed on the chip during the measurement. We performed 10-fold

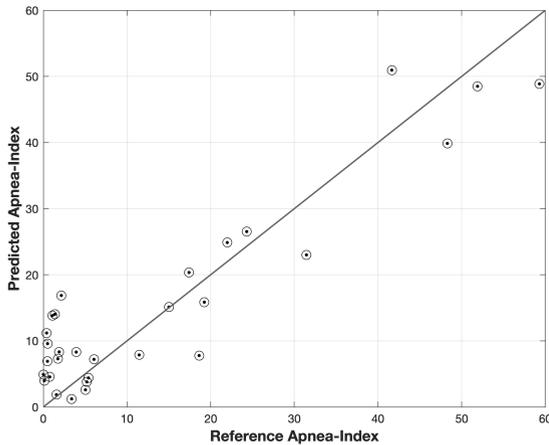


FIGURE 5. The comparison between reference (PSG) and predicted (from our sensor) apnea-index (number of apneas per hour of sleep) for thresholding approach and taking into account short-term energy signals and 5% of threshold level.

hold-out cross-validation, in which for each iteration randomly selected 66% of data (20 subjects) were taken for model training, and the remaining 10 subjects were used for testing. The classification was performed using 3 commonly used machine learning methods.

- Random Forests;
- C5.0 (boosted recursive partitioning); and
- XGBoost.

Entire analyses presented in II.C and II.D sections were carried out using MATLAB. In II.E section, the parameterization was performed in Matlab, and the machine-learning process, comprising data division, cross-validation, modeling and validation, was done using R language. Shiny web app uses R language in the background, as described in [20].

III. RESULTS

A. APNEA RATE ESTIMATION

For the thresholding approach and after exploring all of the combinations, the greatest coherence between audio and motion sensor and PSG was achieved for the short-term energy signal, after normalization within 30-min segments, and the threshold of 5%. Then, resulting in:

- $Pearson'sR = 0.92 **$,
- $Lin'sR = 0.70$,
- $Intercept = 5.0$,
- $MAE = 5.7$,
- $AD = -2.0$, and
- $Limits \pm 1.96SD = 13.3$.

Figure 5 presents the graphical comparison of the apnea-index, not considering the body position, and Figure 6 shows the corresponding Bland-Altman plot.

On the other hand, the segmentation-based method allows searching and counting the moments, where the break between subsequent episodes is greater than 10 second, or when the irregularity of found episodes are above the threshold. Then the results are worse (except the average difference) than for thresholding method:

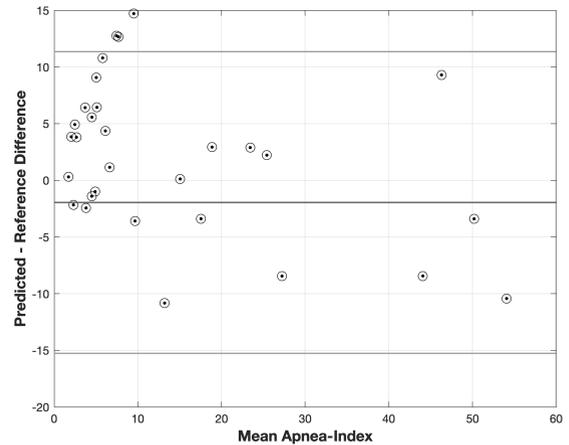


FIGURE 6. The Bland-Altman plot between reference (PSG) and predicted (from our sensor) apnea-index (number of apneas per hour of sleep) for thresholding approach and taking into account short-term energy signals and 5% of threshold level.

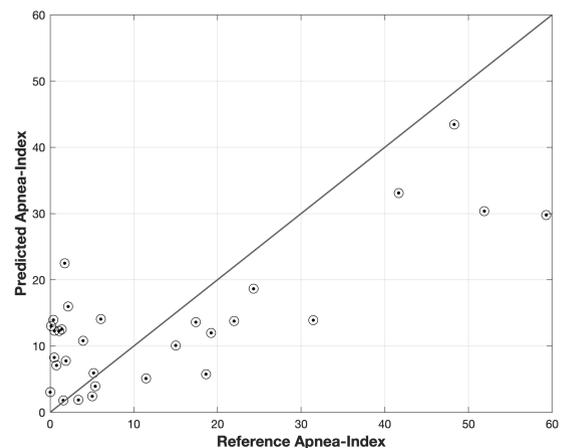


FIGURE 7. The comparison between reference (PSG) and predicted (from our sensor) apnea-index (number of apneas per hour of sleep) for segmentation-based approach.

- $Pearson'sR = 0.77 **$,
- $Lin'sR = 0.60$,
- $Intercept = 7.3$,
- $MAE = 9.0$,
- $AD = 0.1$, and
- $Limits \pm 1.96SD = 22.3$.

Figure 7 presents the graphical comparison of the apnea-index, not considering the body position, for segmentation-based approach, and Figure 8 shows the corresponding Bland-Altman plot.

B. BODY POSITION ESTIMATION

The accuracy of body position classification with audio and motion sensor vs PSG data in a hierarchical system was as follows:

- 99.9% for lying / non-lying distinguishing; and
- 97.3% for lying supine / non-supine.

It was noted that further improvement of the accuracy, mainly by adopting some heuristics related to body positions

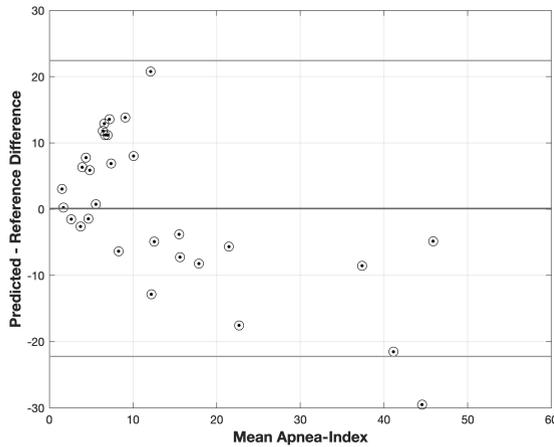


FIGURE 8. The Bland-Altman plot between reference (PSG) and predicted (from our sensor) apnea-index (number of apneas per hour of sleep) for segmentation-based approach.

TABLE 2. The calculation of sAI and nsAI using audio and motion sensor and thresholding method; “-” means that the body position lasted less than 1 hour during the entire night; and arbitrarily, estimating an index would not be meaningful then.

ID	sAI	nsAI	ID	sAI	nsAI	ID	sAI	nsAI
1	3.5	5.9	11	4.4	0.9	21	16.4	8.3
2	9.3	6.8	12	9.7	7.2	22	33.3	7.2
3	6.6	10.1	13	20.2	6.9	23	62.6	33.3
4	29.5	11.6	14	-	7.8	24	30.0	6.5
5	46.9	50.5	15	14.1	8.5	25	3.6	-
6	29.0	45.0	16	7.7	9.6	26	30.9	15.2
7	-	28.2	17	90.6	32.3	27	1.0	3.9
8	-	20.7	18	-	13.8	28	5.1	4.3
9	16.8	14.1	19	8.6	3.9	29	11.1	3.0
10	-	5.8	20	-	25.1	30	10.9	12.5

changes and to parts of the registration when the subject’s activity is the highest, would require fully synchronized registrations.

C. sAI AND nsAI CALCULATION

As the Lin’s concordance coefficient and MAE appears more important than AD, and there are no division for Apnea Index - to perform standardized quantitative analysis, we chose thresholding approach as the most efficient for sAI and nsAI calculation. We gathered sAI and nsAI parameters for all subjects using our sensor. They are stored in Table 2.

According to the results, six out of 30 subjects slept in a supine body position for less than 1 hour during a night; and therefore, we decided not to calculate the sAI then (PSG analysis reported 7 of such cases).

For the remaining subjects we checked the significance of the differences using Wilcoxon signed rank test (with the significance level of 0.05). Mean ± SE of the difference between sAI and nsAI for the audio and motion sensor were 8.3 ± 3.2 (p-value = 0.012 *; 3.4(16.3) median(IQR)), and for PSG reference they were 11.2 ± 3.8 (p-value < 0.001 ***; 5.5(17.3) median(IQR)). Just for reporting, PSG data suggested the sAHIs were greater than nsAHIs with a median difference of 17.1 (p-value < 0.001 ***).

TABLE 3. The confusion matrix presenting the comparison between reference and predicted AHI index based on the non-linear-based approach. As we chose hold-out cross-validation, in each iteration, there are 10 subjects taken for testing. There are also 10 iterations, so 10 times 10 gives 100 - the sum of values in the table.

Prediction \ Reference	Normal	Mild	Moderate	Severe
Normal	0	4	0	0
Mild	4	30	2	0
Moderate	0	6	1	4
Severe	2	4	7	36

TABLE 4. The combined confusion matrix presenting the comparison between reference and predicted AHI index, when normal and mild cases are combined, as well as moderate and severe ones. As we chose hold-out cross-validation, in each iteration, there are 10 subjects taken for testing. There are also 10 iterations, so 10 times 10 gives 100 - the sum of values in the table.

Prediction \ Reference	Normal + Mild	Moderate + Severe
Normal + Mild	38	2
Moderate + Severe	12	48

D. QUALITATIVE AHI ESTIMATION

The non-linear-based approach had the highest performance for random forest classifier. The confusion matrix is presented in Table 3.

The overall accuracy was only 67%, with relatively acceptable Cohen’s Kappa value (0.478). More detailed analysis presents, that Mild and Severe classes are best distinguished (with sensitivity of 68% and the specificity of 89% for the former; and the sensitivity of 90% and the specificity of 78% for the latter).

However, when combining Normal with Mild, and Moderate with Severe we may obtain the overall accuracy of 86%, with the sensitivity (finding more severe cases) of 96%, and the specificity (related with less severe cases) of 76%. It appears it is enough accuracy for screening purposes. The combined confusion matrix is presented in Table 4.

IV. DISCUSSION

The development of new, wireless sensors together with an environment for “big data” analysis is a must nowadays in sleep medicine [23]. Currently used methods for OSA diagnostics like polysomnography or polygraphy have multiple limitations among which the cost of a study, hospital conditions, and inconvenience of cables and sensors, are predominant. We do not try to undervalue PSG in the clinical practice, but rather we would like to emphasize a great need to develop a system for a cheap and reliable OSA screening performed in the multi-night setting, e.g., to monitor treatment effect in an outpatient setting and to select, who would need full PSG. Of course, such screening would not be possible for some groups of patients, like ones with seizures, arrhythmias or coronary artery disease, who still should be scheduled for a regular sleep study, according to the medical recommendations [24], [25].

Currently, questionnaire-based screening is the only possible way to sample big cohorts. Among the many questionnaires, the four most often used are Epworth Sleepiness Scale (ESS), Berlin Questionnaire, STOP, and STOP-BANG questionnaire. As multiple studies show the major problem

with the use of questionnaires is a low specificity of the method. This, in turn, leads to a high false positive result and a failure in the exclusion of low-risk patients [26]. Among newly proposed methods of OSA screening are photoplethysmography [27], ballistocardiography [28], piezoelectric sensors [29], accelerometers [30], audio signals [31], [32], and radar-based systems with non-contact sensing for sleep apnea and sleep body position, respectively: [33], [34]. Most methods which could be used for home detection of OSA were recently reviewed by *Mendonça et al.* [11].

In this study, we used the sensor, which measures the audio data, simultaneously with motion signals from 3-axis accelerometers. Of course, it is not the first study presenting the tracheal audio registration. Other systems were presented and discussed in our previous paper [15] (see Discussion). The sensor appears to fit the wearable sensing paradigm, discussed in [35]. We tested several placements of the sensor, e.g., in different places of the neck, on chest (upper sternum and middle sternum), and based on the signal-to-noise ratio, suprasternal notch appeared the best place to put the device. Moreover, sternum locations amplified heart tones, while muting the respiratory “noises”.

We checked several approaches and found that the thresholding applied at the short-term energy of the audio signal with a normalized level of normal breathing achieved the best performance of the quantitative estimation of apnea index. The reported accuracies are worse than in the study of *Kalkbrenner et al.* [17]. Probably, it is due to signal quality and indirectly due to the lack of synchronization. The discrepancy between predicted Apnea Index (from our sensor) and reference AI (PSG-based) is higher for mild cases. This is a known phenomenon - with lower AHI, the differences between nights are greater even when using one type of reference device. For this reason, overnight tests should be multi-night and “averaged”. This was clearly stated lately by *Roeder and Kohler* [36].

On the other hand, a qualitative technique using a random forest classifier allowed to report 86% of accuracy (with the sensitivity of 96% and the specificity of 76%), when setting a binary border at the level of $AHI = 15$. Interestingly, in the review on detecting sleep apnea using deep learning methods [37], only a single study focused on breathing sounds by *Kim et al.* was found [38]. In that study, they used breathing sounds with the accuracy of 88.3% for four-group AHI classification, and 92.5% using binary classification.

The development of a personal sleep study device would finally enable an era of “personalized sleep medicine”. As there are dozens of different methods of OSA treatment, including weight reduction, lifestyle changes, etc., the use of home-based personal sleep study is crucial to monitor OSA status before and after any intervention. Interestingly, personalized sleep diagnostic sensor can provide not only diagnostic but also therapeutic options. One of the therapies that can be successfully used by the patient himself is positional therapy when “positional OSA” occurs [39], [40]. For this reason, a personal sensor needs not only to be a breathing

disorders sensor but simultaneously has to assess body position during sleep. Our results show close to 100% accuracy in assessing body position relative to the PSG-based results. Although the position of our sensor was different from the PSG accelerometer (located on the thoracic belt), but that did not influence the overall accuracy. This shows that there is probably no need to mount a second sensor on the chest as in the system proposed by *Kalkbrenner et al.* [17], [41]. As most of the actigraphy systems are wrist-worn this should also be discussed. There are several important advantages in using a chest/neck placement of the actigraphy vs. wrist actigraphy. First, the possibility to acquire additional signals from a chest/neck location, such as heart rate variability, needs to be mentioned [42]. Also, a chest/neck actigraphy can add additional data on sleep position unlike wrist actigraphy [43]. This could also be shown in our study where a sleep position detection reached near 100% values comparing to PSG data. The study by *Razjouyan et al.* also showed that the chest actigraphy could identify sleep/wake episodes with an accuracy of on average 6% higher than wrist sensors [43].

The results of our study showed feasibility of distinguishing between positional and non-positional OSA patients. Therefore, positional therapy could also be implemented in pOSA patients. Similarly, a system described by *Levendowski et al.* showed a possible acquiring of positional OSA with a system mounted on the back of the neck [18]. Interestingly, the authors show two cases of under-reporting of supine sleep with their device - one in a situation of trunk supine and neck upright position, and one with trunk supine and head lateral position. We feel our sensor placement would not be vulnerable to such situations and would show a position in agreement with chest actigraphy, but additional studies are needed. *Zhu et al.* for the first time analyzed the role of head position in positional OSA [44]. They found a significant reduction in AHI during supine sleep while the head was rotated to the right or to the left. The drop in AHI was dramatically more significant when a patient was in head-lateral and trunk-lateral position. So, the final location of the sensor position in sleep studies needs to be further assessed, as the chest may not be the ideal place.

The main weakness of this study is a limited number of patients and the lack of complete synchronization between PSG and audio/motion signals. Therefore, we were not able to perform the analysis, in which all references (e.g., beginnings and endings of both apneas and hypopneas) have a temporal relationship with calculated audio features; we could only compare the final predicted number of apneas along with the ground truth value without the possibility to estimate event-based metrics of accuracy. Accordingly, the quantitative prediction of hypopneas was almost impossible in this setup; however, we include them in the qualitative analysis presented in sections 2.6. and 3.4. Those also did not allow us to use more sophisticated methods, like recurrent deep learning techniques, like it was presented by *Nakano et al.* [45]. The device do not allow to distinguish central and obstructive apneas, as there is no EEG and direct respiratory effort

information; however, we will work on it, as in our opinion, audio contact sensor with actigraphy together might be used to analyze respiratory effort indirectly.

Also, there was no assessment of the efficiency of the proposed method in the home uncontrolled environment, where various sounds of the ambiance may interfere with the audio features taken for the analysis.

V. CONCLUSION

The assessment of quantitative and qualitative accuracy of audio sensor in apnea detection showed promising results. The thresholding approach allowed us to achieve around -2 ± 13 (Bland-Altman bias and limits of agreement), mean absolute error of 5.7, and Pearson's and Lin's R at the levels of 0.92 and 0.70, respectively. Random forest classifier reached 86% accuracy (96% of sensitivity, and 76% of specificity) of distinguishing between normal/mild, and moderate/severe cases.

The results might be even better once the full synchronization with PSG is preserved. What is more important, presented audio and motion sensor has also an activity sensor (3-axis accelerometer), enabling positional OSA analysis.

AUTHOR CONTRIBUTION

- Conceptualization, M.M. and W.K.;
- methodology, M.M. and W.K.;
- software, M.M.;
- validation, M.M., T.A.V. and W.K.;
- formal analysis, M.M.;
- investigation, W.K.;
- resources, W.K.;
- data curation, M.M.;
- writing—original draft preparation, M.M. and W.K.;
- writing—review and editing, M.M., T.A.V. and W.K.;
- visualization, M.M.;
- supervision, T.A.V. and W.K.;
- project administration, M.M., T.A.V. and W.K.;
- funding acquisition, M.M. and W.K.

ACKNOWLEDGMENT

The work was partly a result of a Short-Term Scholarship Project (The Bekker Programme) funded by the Polish National Agency for Academic Exchange (NAWA), partly financed as part of the Fast Track program, conducted in Poland by National Centre for Research and Development (POIR.01.01.01-00-0212/19). The authors thank Jonasz Laba for suggesting linguistic adjustments.



CONFLICT OF INTEREST

M.M. and W.K. are shareholders of Clebre, but received no remuneration on performing this study. T.A.V. declares no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

REFERENCES

- [1] A. Okcay, V. K. Somers, and S. M. Caples, "Obstructive sleep apnea and hypertension," *J. Clin. Hypertension*, vol. 10, no. 7, pp. 549–555, 2008.
- [2] S. Mandal and B. D. Kent, "Obstructive sleep apnoea and coronary artery disease," *J. Thoracic Disease*, vol. 10, no. S34, pp. S4212–S4220, Dec. 2018.
- [3] S. Sharma and A. Culebras, "Sleep apnoea and stroke," *Brit. Med. J.*, vol. 1, no. 4, pp. 185–191, Dec. 2016.
- [4] J. A. Dempsey, S. C. Veasey, B. J. Morgan, and C. P. O'Donnell, "Pathophysiology of sleep apnea," *Physiological Rev.*, vol. 90, no. 1, pp. 47–112, 2010.
- [5] R. Heinzer, S. Vat, P. Marques-Vidal, H. Marti-Soler, D. Andries, N. Tobback, V. Mooser, M. Preisig, A. Malhotra, G. Waeber, P. Vollenweider, M. Tafti, and J. Haba-Rubio, "Prevalence of sleep-disordered breathing in the general population: The HypnoLaus study," *Lancet Respiratory Med.*, vol. 3, no. 4, pp. 310–318, Apr. 2015.
- [6] T. Penzel and A. Sabil, "The use of tracheal sounds for the diagnosis of sleep apnoea," *Breathe*, vol. 13, no. 2, pp. e37–e45, Jun. 2017.
- [7] S. Alqassim, M. Ganesh, S. Khoja, M. Zaidi, F. Aloul, and A. Sagahyroun, "Sleep apnea monitoring using mobile phones," in *Proc. IEEE 14th Int. Conf. e-Health Netw., Appl. Services (Healthcom)*, Oct. 2012, pp. 443–446.
- [8] T. Rahman, A. T. Adams, M. Zhang, E. Cherry, B. Zhou, H. Peng, and T. Choudhury, "Bodybeat: A mobile system for sensing non-speech body sounds," in *Proc. MobiSys*, vol. 14, 2014, pp. 2594368–2594386.
- [9] A. A. Ong and M. B. Gillespie, "Overview of smartphone applications for sleep analysis," *World J. Otorhinolaryngol.-Head Neck Surg.*, vol. 2, no. 1, pp. 45–49, Mar. 2016.
- [10] T. Penzel, C. Schöbel, and I. Fietze, "New technology to assess sleep apnea: Wearables, smartphones, and accessories," *FRsearch*, vol. 7, p. 413, Mar. 2018.
- [11] F. Mendonça, S. S. Mostafa, A. G. Ravelo-García, F. Morgado-Dias, and T. Penzel, "Devices for home detection of obstructive sleep apnea: A review," *Sleep Med. Rev.*, vol. 41, pp. 149–160, Oct. 2018.
- [12] E. Fino and M. Mazzetti, "Monitoring healthy and disturbed sleep through smartphone applications: A review of experimental evidence," *Sleep Breathing*, vol. 23, no. 1, pp. 13–24, Mar. 2019.
- [13] A. Oksenberg, E. Arons, H. Radwan, and D. S. Silverberg, "Positional vs nonpositional obstructive sleep apnea patients: Anthropomorphic, nocturnal polysomnographic and multiple sleep latency test data," *Chest*, vol. 112, no. 3, pp. 629–639, 1997.
- [14] W. Richard, D. Kox, C. den Herder, M. Laman, H. van Tinteren, and N. de Vries, "The role of sleep position in obstructive sleep apnea syndrome," *Eur. Arch. Oto-Rhino-Laryngol. Head Neck*, vol. 263, no. 10, pp. 946–950, 2006.
- [15] M. Młyńczak, E. Migacz, M. Migacz, and W. Kukwa, "Detecting breathing and snoring episodes using a wireless tracheal sensor—A feasibility study," *IEEE J. Biomed. Health Informat.*, vol. 21, no. 6, pp. 1504–1510, Nov. 2016.
- [16] S. Saha, M. Kabir, N. M. Ghahjaverestan, M. Hafezi, B. Gavrilovic, K. Zhu, H. Alshaer, and A. Yadollahi, "Portable diagnosis of sleep apnea with the validation of individual event detection," *Sleep Med.*, vol. 69, pp. 51–57, May 2020.
- [17] C. Kalkbrenner, M. Eichenlaub, S. Rüdiger, C. Kropf-Sanchen, R. Brucher, and W. Rottbauer, "Validation of a new system using tracheal body sound and movement data for automated apnea-hypopnea index estimation," *J. Clin. Sleep Med. Off. Publication Amer. Acad. Sleep Med.*, vol. 13, no. 10, p. 1123, 2017.
- [18] D. J. Levendowski, B. Veljkovic, S. Seagraves, and P. R. Westbrook, "Capability of a neck worn device to measure sleep/wake, airway position, and differentiate benign snoring from obstructive sleep apnea," *J. Clin. Monitor. Comput.*, vol. 29, no. 1, pp. 53–64, Feb. 2015.
- [19] C. Iber, S. Ancoli-Israel, A. L. Chesson, and S. F. Quan, "The new sleep scoring manual—the evidence behind the rules," *J. Clin. Sleep Med.*, vol. 3, no. 2, p. 107, 2007.
- [20] M. Młyńczak, T. A. Valdez, and W. Kukwa, "Comparing sleep studies in terms of the apnea-hypopnea index," 2019, *arXiv:1908.00960*. [Online]. Available: <https://arxiv.org/abs/1908.00960>
- [21] M. Halevi, E. Dafna, A. Tarasiuk, and Y. Zigel, "Can we discriminate between apnea and hypopnea using audio signals?" in *Proc. 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2016, pp. 3211–3214.
- [22] W. R. Ruehland, P. D. Rochford, F. J. O'Donoghue, R. J. Pierce, P. Singh, and A. T. Thornton, "The new AASM criteria for scoring hypopneas: Impact on the apnea hypopnea index," *Sleep*, vol. 32, no. 2, pp. 150–157, Feb. 2009.

- [23] J. Pépin, S. Bailly, and R. Tamisier, “Big data in sleep apnoea: Opportunities and challenges,” *Respirology*, vol. 25, no. 5, pp. 486–494, May 2020.
- [24] L. Desteghe, J. M. L. Hendriks, R. D. McEvoy, C. L. Chai-Coetzer, P. Dendale, P. Sanders, H. Heidebuchel, and D. Linz, “The why, when and how to test for obstructive sleep apnea in patients with atrial fibrillation,” *Clin. Res. Cardiol.*, vol. 107, no. 8, pp. 617–631, Aug. 2018.
- [25] J. R. Tietjens, D. Claman, E. J. Kezirian, T. De Marco, A. Mirzayan, B. Sadroonri, A. N. Goldberg, C. Long, E. P. Gerstenfeld, and Y. Yeghiazarians, “Obstructive sleep apnea in cardiovascular disease: A review of the literature and proposed multidisciplinary clinical management strategy,” *J. Amer. Heart Assoc.*, vol. 8, no. 1, Jan. 2019, Art. no. e010440.
- [26] H.-Y. Chiu, P.-Y. Chen, L.-P. Chuang, N.-H. Chen, Y.-K. Tu, Y.-J. Hsieh, Y.-C. Wang, and C. Guilleminault, “Diagnostic accuracy of the berlin questionnaire, STOP-BANG, STOP, and epworth sleepiness scale in detecting obstructive sleep apnea: A bivariate meta-analysis,” *Sleep Med. Rev.*, vol. 36, pp. 57–70, Dec. 2017.
- [27] Y. Li, H. Gao, and Y. Ma, “Evaluation of pulse oximeter derived photoplethysmographic signals for obstructive sleep apnea diagnosis,” *Medicine*, vol. 96, no. 18, p. e6755, May 2017.
- [28] Z. Wang, X. Zhou, W. Zhao, F. Liu, H. Ni, and Z. Yu, “Assessing the severity of sleep apnea syndrome based on ballistocardiogram,” *PLoS ONE*, vol. 12, no. 4, Apr. 2017, Art. no. e0175351.
- [29] A. Rasheed, E. Iranmanesh, W. Li, Y. Xu, Q. Zhou, H. Ou, and K. Wang, “An active self-driven piezoelectric sensor enabling real-time respiration monitoring,” *Sensors*, vol. 19, no. 14, p. 3241, Jul. 2019.
- [30] D. S. Morillo, J. L. R. Ojeda, L. F. C. Foix, and A. L. Jimenez, “An accelerometer-based device for sleep apnea screening,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 2, pp. 491–499, Mar. 2010.
- [31] H. Nakano, M. Hayashi, E. Ohshima, N. Nishikata, and T. Shinohara, “Validation of a new system of tracheal sound analysis for the diagnosis of sleep apnea-hypopnea syndrome,” *Sleep*, vol. 27, no. 5, pp. 951–957, Aug. 2004.
- [32] A. Yadollahi and Z. Moussavi, “Acoustic obstructive sleep apnea detection,” in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Sep. 2009, pp. 7110–7113.
- [33] S. Kang, D.-K. Kim, Y. Lee, Y.-H. Lim, H.-K. Park, S. H. Cho, and S. H. Cho, “Non-contact diagnosis of obstructive sleep apnea using impulse-radio ultra-wideband radar,” *Sci. Rep.*, vol. 10, no. 1, pp. 1–7, Dec. 2020.
- [34] M. Piriyaajakonkij, P. Warin, P. Lakhan, P. Leelaarporn, T. Pianpanit, N. Kumchaiseemak, S. Suwajanakorn, N. Niparnan, S. Chandra Mukhopadhyay, and T. Wilaiprasitporn, “SleepPoseNet: Multi-view learning for sleep postural transition recognition using UWB,” 2020, *arXiv:2005.02176*. [Online]. Available: <http://arxiv.org/abs/2005.02176>
- [35] P. Zalewski, L. Marchegiani, A. Elsts, R. Piechocki, I. Craddock, and X. Fafoutis, “From bits of data to bits of knowledge—An on-board classification framework for wearable sensing systems,” *Sensors*, vol. 20, no. 6, p. 1655, 2020.
- [36] M. Roeder and M. Kohler, “It’s time for multiple sleep night testing in OSA,” *Chest*, vol. 158, no. 1, pp. 33–34, Jul. 2020.
- [37] S. S. Mostafa, F. Mendonça, A. G. Ravelo-García, and F. Morgado-Dias, “A systematic review of detecting sleep apnea using deep learning,” *Sensors*, vol. 19, no. 22, p. 4934, Nov. 2019.
- [38] T. Kim, J.-W. Kim, and K. Lee, “Detection of sleep disordered breathing severity using acoustic biomarker and machine learning techniques,” *Biomed. Eng. OnLine*, vol. 17, no. 1, p. 16, Dec. 2018.
- [39] P. R. Srijiithesh, R. Aghoram, A. Goel, and J. Dhanya, “Positional therapy for obstructive sleep apnoea,” *Cochrane Database Systematic Rev.*, no. 5, pp. 1–51, 2019, Art. no. CD010990, doi: [10.1002/14651858.CD010990.pub2](https://doi.org/10.1002/14651858.CD010990.pub2).
- [40] G. E. de Vries, A. Hoekema, M. H. Doff, H. A. Kerstjens, P. M. Meijer, J. H. van der Hoven, and P. J. Wijkstra, “Usage of positional therapy in adults with obstructive sleep apnea,” *J. Clin. Sleep Med.*, vol. 11, no. 2, pp. 131–137, 2015.
- [41] C. Kalkbrenner, R. Brucher, T. Kesztyüs, M. Eichenlaub, W. Rottbauer, and D. Scharnbeck, “Automated sleep stage classification based on tracheal body sound and actigraphy,” *GMS German Med. Sci.*, vol. 17, Dec. 2019. [Online]. Available: <http://www.ejgms.de/en/journals/gms/2019-17/000268.shtml>, doi: [10.3205/000268](https://doi.org/10.3205/000268).
- [42] M. Aktaruzzaman, M. W. Rivolta, R. Karmacharya, N. Scarabottolo, L. Pugnetti, M. Garegnani, G. Bovi, G. Scalera, M. Ferrarin, and R. Sassi, “Performance comparison between wrist and chest actigraphy in combination with heart rate variability for sleep classification,” *Comput. Biol. Med.*, vol. 89, pp. 212–221, Oct. 2017.
- [43] J. Razjouyan, H. Lee, S. Parthasarathy, J. Mohler, A. Sharafkhaneh, and B. Najafi, “Improving sleep quality assessment using wearable sensors by including information from postural/sleep position changes and body acceleration: A comparison of chest-worn sensors, wrist actigraphy, and polysomnography,” *J. Clin. Sleep Med. Off. Publication Amer. Acad. Sleep Med.*, vol. 13, no. 11, p. 1301, 2017.
- [44] K. Zhu, T. D. Bradley, M. Patel, and H. Alshaer, “Influence of head position on obstructive sleep apnea severity,” *Sleep Breathing*, vol. 21, no. 4, pp. 821–828, Dec. 2017.
- [45] H. Nakano, T. Furukawa, and T. Tanigawa, “Tracheal sound analysis using a deep neural network to detect sleep apnea,” *J. Clin. Sleep Med.*, vol. 15, no. 8, pp. 1125–1133, Aug. 2019.



MARCEL MŁYŃCZAK (Member, IEEE) received the M.Sc. degree in biomedical engineering from the Faculty of Mechatronics, Warsaw University of Technology, Poland, in 2014. He defended a Ph.D. thesis in biocybernetics and biomedical engineering there.

He is currently a Research and a Teaching Assistant Professor with the Warsaw University of Technology. He is engaged in biomedical engineering studies (particularly non-invasive cardiorespiratory physiological measurements and analysis, also in the sleep medicine and sports medicine contexts). His research interests include wearable systems, m-health, and data science for applications in clinical research.

Dr. Młyńczak is a member of the IEEE Engineering in Medicine and Biology Society. He received the NSF Young Professionals Award at the Smart and Connected Health Contest (during IEEE EMBC 2016).



TULIO A. VALDEZ was born in Colombia. He received the master’s degree in clinical and translational research from the University of Connecticut.

He is currently the Surgeon Scientist with a subspecialty interest in pediatric otolaryngology. He attended medical school at Universidad Javeriana in Bogota Colombia before undertaking his residency in Otolaryngology, Head, and Neck Surgery in Boston. He completed his Pediatric

Otolaryngology Fellowship at the Texas Children’s Hospital, Houston, in 2007. Clinically, he has an interest in airway surgery and swallowing disorders. He has a special interest in the management of sinus disease in cystic fibrosis. He has coauthored one textbook and numerous book chapters and scientific manuscripts. He continues his clinical research in these areas, particularly with a focus on aerodigestive disorders. Scientifically, he has developed various imaging methods to diagnose otitis media and cholesteatoma a middle ear condition that can lead to hearing loss. He was a part of the Laser Biomedical Research Center, Massachusetts Institute of Technology. His research interest includes novel imaging modalities to better diagnose ear infections one of the most common pediatric problems. His research has now expanded to include better intraoperative imaging modalities in pediatric patients to improve surgical outcomes without the need for radiation exposure.



WOJCIECH KUKWA is currently the Head of the ENT Department, Czerniakowski Hospital, Warsaw. He is also an Assistant Professor with the Medical University of Warsaw. He is also the Founder of a Healthy Sleep Foundation, which performed a sleep disorders screening program in Polish children. He is engaged in many ongoing sleep medicine projects. He is a member of the European Sleep Research Society, the International Surgical Sleep Society, and the International Sleep Science and Technology Association.