





Article

Irregularity and Variability Analysis of Airflow Recordings to Facilitate the Diagnosis of Paediatric Sleep Apnoea-Hypopnoea Syndrome

Verónica Barroso-García ^{1,*}, Gonzalo César Gutiérrez-Tobal ¹ , Leila Kheirandish-Gozal ², Daniel Álvarez ^{1,3} , Fernando Vaquerizo-Villar ¹, Andrea Crespo ^{1,3}, Félix del Campo ^{1,3} , David Gozal ² and Roberto Hornero ¹ 

¹ Biomedical Engineering Group, E.T.S.I. Telecomunicación, Universidad de Valladolid, Valladolid 47011 Spain; gonzalo.gutierrez@gib.tel.uva.es (G.C.G.-T.); dalvgon@ribera.tel.uva.es (D.Á.); fernando.vaquerizo@gib.tel.uva.es (F.V.-V.); andreacresposedano@gmail.com (A.C.); fsas@telefonica.net (F.d.C.); robhor@tel.uva.es (R.H.)

² Section of Sleep Medicine, Department of Paediatrics, Pritzker School of Medicine, Biological Sciences Division, The University of Chicago, Chicago, IL 60637, USA; lgozal@peds.bsd.uchicago.edu (L.K.-G.); dgozal@uchicago.edu (D.G.)

³ Servicio de Neumología, Hospital Universitario Río Hortega, Valladolid 47012, Spain

* Correspondence: veronica.barroso@gib.tel.uva.es; Tel.: +34-983-423-000 (ext. 4716)

Received: 28 July 2017; Accepted: 24 August 2017; Published: 26 August 2017

Abstract: The aim of this paper is to evaluate the evolution of irregularity and variability of airflow (AF) signals as sleep apnoea-hypopnoea syndrome (SAHS) severity increases in children. We analyzed 501 AF recordings from children 6.2 ± 3.4 years old. The respiratory rate variability (RRV) signal, which is obtained from AF, was also estimated. The proposed methodology consisted of three phases: (i) extraction of spectral entropy (SE_1), quadratic spectral entropy (SE_2), cubic spectral entropy (SE_3), and central tendency measure (CTM) to quantify irregularity and variability of AF and RRV; (ii) feature selection with forward stepwise logistic regression (FSLR), and (iii) classification of subjects using logistic regression (LR). SE_1 , SE_2 , SE_3 , and CTM were used to conduct exploratory analyses that showed increasing irregularity and decreasing variability in AF, and increasing variability in RRV as apnoea-hypopnoea index (AHI) was higher. These tendencies were clearer in children with a higher severity degree (from $AHI \geq 5$ events/hour). Binary LR models achieved 60%, 76%, and 80% accuracy for the AHI cutoff points 1, 5, and 10 e/h, respectively. These results suggest that irregularity and variability measures are able to characterize paediatric SAHS in AF recordings. Hence, the use of these approaches could be helpful in automatically detecting SAHS in children.

Keywords: sleep apnoea-hypopnoea syndrome; airflow; respiratory rate variability; spectral entropy; central tendency measure; children

1. Introduction

The sleep apnoea-hypopnoea syndrome (SAHS) is a chronic respiratory disorder characterized by recurrent events of apnoea (complete absence of airflow) or hypopnoea (significant airflow reduction) during sleep time [1,2]. SAHS increases the risk for adverse and serious medical consequences in paediatric patients affected by this frequent condition. SAHS may affect cardiovascular and central nervous systems, as well as decrease somatic growth and promote nocturnal enuresis, all of which lead to decreases in health and quality of life [3]. Moreover, paediatric SAHS is an underdiagnosed disease with high prevalence, affecting up to 4% of the childhood population [3,4].

SAHS can affect both children and adults. However, its physical and cognitive consequences are different. In this sense, excessive daytime sleepiness and depression are more frequent in adults with

SAHS, whereas reduced growth and abnormalities in neurocognitive and behavioral development are common in affected paediatric patients [5,6]. Similarly, the main causes of SAHS also differ in adult and childhood population. Children suffering from SAHS usually present adenotonsillar hypertrophy that causes narrowing of the upper airways [5,6]. In contrast, narrowing of the upper airway is commonly associated with increased adipose tissue in adults, in the context of concurrent obesity [5,6]. In this regard, the majority of children affected by SAHS are not overweight, although both overweight and frank obesity increase the risk of suffering from the disease. The rules of the American Academy of Sleep Medicine (AASM) for scoring apnoeas and hypopnoeas are also somewhat different for adults and children [2]. Ten-second durations or longer are required to score an apnoea or a hypopnea event in the case of adults, whereas only six seconds (or, alternatively, two respiratory cycles) are required for children [2]. Thus, the less restrictive criteria for children are also evident when diagnosing and treating SAHS. The apnoea-hypopnoea index (AHI), calculated as the number of apnoeas and hypopnoeas per hour of sleep, is used to diagnose SAHS and determine its severity in both adults and children [7]. Nonetheless, adults with $AHI < 5$ events/hour (e/h) are considered free of SAHS, whereas $AHI \geq 1$ e/h has been often used as indicative of the presence of the disease in children [8–11]. Similarly, when a paediatric subject shows $AHI \geq 5$ e/h, a surgical treatment (adenotonsillectomy) is commonly recommended, whereas this is only the cutoff point of mild severity for an adult, and treatment is rarely if ever prescribed under these circumstances [8,11]. Similarly, 10 e/h still corresponds to mild SAHS in adults, but determines severe SAHS in children [8–11].

Due to the serious consequences of paediatric SAHS, early detection and treatment are critical. The gold standard diagnostic test is nocturnal polysomnography (PSG) [12]. PSG implies the overnight monitoring of the patient at a specialized sleep unit in order to record multiple biomedical signals during sleep. These include electrocardiogram (ECG), electroencephalogram (EEG), electrooculogram (EOG), electromyography (EMG), photoplethysmography (PPG), blood oxygen saturation (SpO_2), and airflow (AF), usually assessed using nasal pressure recordings and thermistor-derived airflow qualitative assessments. The recordings are used to calculate the AHI, whose evaluation lets the sleep specialist determine the presence and severity of SAHS. Despite its effectiveness, PSG requires expensive equipment to acquire the biomedical signals. It is also a labour-intensive task for physicians, who have to spend substantial time and effort evaluating the recordings [13]. These issues generate long waiting lists, as well as delays in diagnosis [13]. Moreover, PSG is particularly uncomfortable and inconvenient for children, due to the multiple sensors attached on their bodies and the need to spend a night at the laboratory facility [14].

Alternative diagnostic methods have been proposed to overcome the PSG limitations. One common approach has been the analysis of a reduced set of those signals involved in the PSG. In this regard, ECG, PPG, SpO_2 , and AF have been already manually evaluated [15] or, alternatively, by the use of automatic processing techniques [16–23]. Direct event-detection of apnoeas and hypopnoeas is a common approach [17–19]. In addition, the overall characterization of biomedical signals in time and frequency domain have been also adopted to help in the diagnosis of childhood SAHS [20–22]. These studies showed that there is no consensus in the threshold used to establish paediatric SAHS and its severity. Several of the published studies used the most conservative AHI cutoff (1 e/h) [15,16,23], whereas others applied less restrictive ones (3 e/h, 5 e/h, or 10 e/h) [8,17–22,24]. None of these studies however, assessed the evolution of irregularity and variability as AHI increases.

In this study, AF and the respiratory rate variability (RRV) signals were analyzed. The analysis of AF is justified since it is directly involved in the definition of apnoeas and hypopnoeas [1,2]. Several studies have already shown its usefulness in the diagnosis of SAHS in adults [22,25,26]. When an apnoeic event occurs, the amplitude of AF tends to values close to zero (apnoea) or decreases significantly (hypopnoea). These amplitude variations not only modify AF in the time domain, but their recurrence also causes alterations in the frequency domain. In addition to the AF signal, the analysis of RRV is also proposed. RRV is directly derived from single-channel AF, and is calculated as the elapsed time between consecutive breaths [27]. RRV has been already used to assist with SAHS

diagnosis in adults [25]. When a complete cessation or reduction of AF occurs, the elapsed time between two consecutive complete respiratory cycles varies [25]. Therefore, apnoeas and hypopnoeas also cause alterations in RRV time series. Moreover, as in the case of AF, the recurrence of these events also modifies its spectrum.

According to the aforementioned considerations, we propose the use of irregularity and variability measures applied to single-channel AF recordings, to assist in paediatric SAHS diagnosis. Our hypothesis was that these analyses are able to characterize the presence and severity of SAHS based on the AF signal obtained from children during sleep. Accordingly, our objectives were to evaluate the evolution of the irregularity and the variability of single-channel AF as AHI increases, as well as to assess the diagnostic ability of the information provided by these analyses.

In order to measure irregularity and variability of AF and RRV, spectral entropy (SE_1), quadratic spectral entropy (SE_2), cubic spectral entropy (SE_3), and central tendency measure (CTM) of both AF and RRV were proposed. Spectral entropies (SE) are commonly used to quantify irregularity in biomedical signals [20,28–31]. They have been successfully used to estimate the irregularity of EEG and magnetoencephalogram (MEG) recordings in Alzheimer’s disease [28], to identify different cardiac arrhythmias in heart rate variability (HRV) signals [29], and, in the context of SAHS, to help to diagnose adults and children with HRV [30], to identify affected children using SpO_2 recordings [20], as well as to help to detect apnoea events from ECG recordings [31]. CTM is a non-linear feature used to measure variability in biomedical time series [25,32–34]. It has been successfully used to quantify the variability of EEG recordings in Alzheimer’s disease [32], to help to identify congestive heart failure from ECG recordings [33], as well as to improve the diagnostic ability of SpO_2 in the detection of SAHS [34]. In order to evaluate the evolution of SE and CTM as AHI is higher, we divided our subject database in four groups of increasingly severity. We have already mentioned that there is no consensus on those AHI cutoff points that should be used to determine SAHS and its severity. However, 1 e/h, 5 e/h, and 10 e/h are commonly used thresholds [8–10,24]. In this regard, several studies suggest a classification of severity based on the association of $AHI < 1$ with no-SAHS, $1 \leq AHI < 5$ with mild SAHS, $5 \leq AHI < 10$ with moderate SAHS, and $AHI \geq 10$ with severe SAHS [8–10,24]. Therefore, our database division followed these severity groups.

After the irregularity and variability analyses of AF and RRV, an automatic feature selection stage was conducted to obtain an optimum feature subset that maximizes the diagnostic potential of the extracted information. In order to automatically select this feature subset, the well-known forward stepwise logistic regression (FSLR) method was proposed [35]. Finally, the diagnostic ability of logistic regression models (LR) built with the selected optimum features was evaluated.

2. Materials and Methods

2.1. Subjects and Signals

The population under study consisted of 501 paediatric subjects. All of them were suspected of suffering from SAHS since they presented common symptoms of the disease, such as snoring and breathing pauses during sleep time. PSGs of these children were recorded in the Paediatric Sleep Unit at the Comer Children’s Hospital of the University of Chicago. The study protocol was approved by the Ethics Committee and all legal caretakers of the children gave their informed consent. The paediatric subjects were diagnosed by medical specialists following the standards of the AASM [2]. Accordingly, apnoea was defined as a decrease $\geq 90\%$ of oronasal airflow during at least two respiratory cycles. Similarly, hypopnoea was defined as a decrease $\geq 30\%$ in the nasal pressure airflow signal lasting at least two respiratory cycles, accompanied by a desaturation of oxygen in blood $\geq 3\%$ and/or an arousal [2].

The population under study was randomly allocated into two groups, training (50%) and test (50%), to train and validate the proposed methodology, respectively. The subjects who did not suffer from SAHS ($AHI < 1$) composed the control group in both datasets. Table 1 shows the demographic

and clinical data of the paediatric subjects under study. No significant differences (p -value < 0.01) were found in age, gender, body mass index (BMI), and AHI between the training group and the test group when the non-parametric Mann-Whitney U test was conducted.

Table 1. Demographic and clinical data from the paediatric subjects under study. Data are presented as median [interquartile range] or n (%); BMI: body mass index; AHI: apnoea–hypopnoea index.

| Characteristics | All | Training Group | Test Group |
|--------------------------|----------------------|----------------------|----------------------|
| Subjects (n) | 501 | 250 | 251 |
| Age (years) | 6 [3, 9] | 6 [4, 8] | 6 [3, 9] |
| Males (n) | 314 (62.67%) | 160 (64%) | 154 (61.35%) |
| BMI (kg/m ²) | 17.81 [15.75, 22.21] | 17.39 [15.57, 21.97] | 18.14 [16.11, 22.44] |
| AHI (e/h) | 3.2 [0.95, 8.08] | 2.69 [0.77, 7.28] | 3.53 [1.37, 9.06] |
| AHI ≥ 1 (e/h) | 367 (73.25%) | 170 (68%) | 197 (78.49%) |
| AHI ≥ 5 (e/h) | 180 (35.93%) | 83 (33.2%) | 97 (38.65%) |
| AHI ≥ 10 (e/h) | 104 (20.76%) | 48 (19.2%) | 56 (22.31%) |

The recordings were acquired using a digital polysomnography (Polysmith, Nihon Kohden America Inc., Irvine, CA, USA). The AF signal used in the study was acquired with a thermistor during the PSG. The sampling frequency of AF was 100 Hz, as recommended by the AASM [2]. A preprocessing phase was carried out to remove the artifacts present on the AF signal. Artifacts were discarded using a methodology based on the comparison of the mean and the kurtosis of AF segments. Statistically significant differences between AF segments with noise or signal loss and the remaining segments were shown in these measures. The age of the subjects under study ranged from zero to 13 years old. In order to minimize possible differences among children due to this age span, AF recordings were normalized as proposed by Varady et al. [36]. RRV was obtained by computing the time between consecutive inspiratory onsets in the AF signal [27]. Hence, we looked for AF intervals containing turning points by finding positive segments followed by negative ones in its derivative [25]. Once these intervals were located, the elapsed times between consecutive maximum points were computed to obtain RRV [25]. Since AF-derived RRV samples do not follow a constant rate, a cubic spline interpolation was applied (100 Hz) in order to enable subsequent analysis in the frequency domain [25].

2.2. Methods

The proposed methodology consists of three phases: (i) extraction of *SE* and *CTM* to measure the irregularity and variability of AF and RRV; (ii) feature selection with FSLR, and (iii) classification of subjects using LR. Figure 1 shows the general scheme of the methodology carried out in our study.

2.2.1. Feature Extraction

SE requires the prior estimation of power spectral density (PSD) of the AF and RRV recordings. It was computed using the Welch method which uses the Fast Fourier Transform [37]. This method is mathematically described as follows [25,37]: the original time series is divided into M overlapped segments of length L . Each of the segments is averaged using a function $w[n]$ that smoothes the edges and improves the spectral resolution of the estimator. After that, the modified periodogram of each segment $v_m[n]$ is calculated by applying the FFT:

$$\hat{P}[f] = \frac{|V[f]|^2}{f_s L U} \quad (1)$$

where f_s is the sampling frequency, and V is the discrete Fourier transform (DFT) of N points of segment $v_m[n]$:

$$V[f] = \sum_{n=0}^{N-1} v_m[n] e^{-j(2\pi k/N)n} \quad k = 0, \dots, N-1 \quad (2)$$

and U is a normalization constant dependent of $w[n]$:

$$U = \frac{1}{M} \sum_{n=0}^{M-1} |w[n]|^2 \quad (3)$$

Finally, all periodograms are averaged to estimate PSD. In the present study, we used 2^{15} samples Hamming window, 50% overlap and 2^{16} point discrete Fourier transform length.

SE is calculated from PSD to characterize the oscillatory behavior of a signal through the flatness/peakedness of its spectral distribution, i.e., it quantifies irregularity [28]. Furthermore, higher-order SE may potentiate the changes in the breathing pattern due to apnoeic events. SE_1 , SE_2 , and SE_3 are based on Shannon's entropy and require normalization of PSD (PSDn) in a selected frequency range f_1 – f_2 . They can be computed as follows [28,30]:

$$SE_i = - \sum_{f=f_1}^{f_2} \text{PSDn}_i(f) \cdot \log[\text{PSDn}_i(f)] \quad (4)$$

where $\text{PSDn}_i(f)$ is [29,31]:

$$\text{PSDn}_i(f) = \frac{|\text{PSD}(f)|^i}{\sum_{f=f_1}^{f_2} |\text{PSD}(f)|^i} \quad i = 1, 2, 3 \quad (5)$$

and i is the order of entropy. The flatter the spectrum of the signal, the closer the value of SE to 1. This would indicate more irregularity for the signal in the time domain. By contrast, the higher spectral peakedness, the closer the value of SE to 0, which would indicate less irregularity in the time domain. Normal breathing in children is confined within a narrow band of PSD (0.20 Hz to 0.40 Hz approximately) [30,38]. As in the case of adults [39], it is expected that apnoeic-related events contribute to more frequency components, i.e., to broaden the spectrum. Thus, AF and RRV signals from more SAHS severity children are expected to have higher SE values. Hence, SE is presented as a suitable tool to evaluate the evolution of irregularity as AHI increases.

CTM is a non-linear parameter for measuring the variability of biomedical time series [25,32–34]. This feature is based on first-order difference plots, which are scatter diagrams centered on the origin that represent displaced subsequences of the original time series: $(x[i+2] - x[i+1])$ vs. $(x[i+1] - x[i])$, where each $x[i]$ is the value of the AF time series at time i [33,34]. CTM quantifies the dispersion of data as the quotient between the number of points located within a circular region of radius r centered on the origin and the total number of points. Thus, CTM can be calculated with the following expression [33,34]:

$$CTM = \frac{\sum_{i=1}^{N-2} \delta(i)}{N-2} \quad (6)$$

where N is the number of points of the time series and $\delta(i)$ is:

$$\delta(i) = \begin{cases} 1, & \text{if } ((x[i+2] - x[i+1])^2 + (x[i+1] - x[i])^2)^{1/2} < r \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

If data are widely scattered throughout the plot, a great number of points are located outside the circular region. When this happens, CTM value is close to 0, indicating high variability of data. Nevertheless, if data are concentrated around the origin, a high number of points fall within the circular region. In this case, CTM tends to 1, indicating low variability of the data. According to (6) and (7), it is expected that the absence of breathing will cause less variability (higher CTM) in AF, as well as higher variability (lower CTM) in the time between breaths, i.e., in RRV signal. Consequently, CTM is a priori a suitable tool to measure the evolution of variability as AHI increases. A good choice of r is fundamental for this purpose. Hence, it has to be experimentally optimized in the training set.

In this regard, we have assessed several r values for both AF and RRV. The CTM variables obtained for each of them where correlated with the AHI. This correlation is shown in Figure 2. Then we chose those r values showing the highest Spearman's correlation coefficient (RHO) between CTM and AHI ($r^{\text{AF}} = 0.0004$ and $r^{\text{RRV}} = 7$).

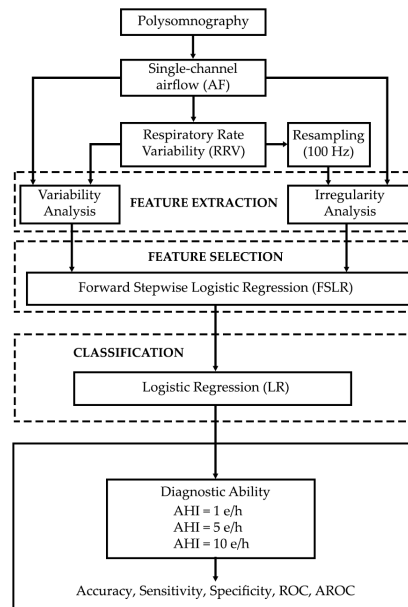


Figure 1. General scheme of the methodology carried out in the study. AHI: apnoea-hypopnoea index; ROC: Receiver Operating Characteristics curves; AROC: area under the ROC curve.

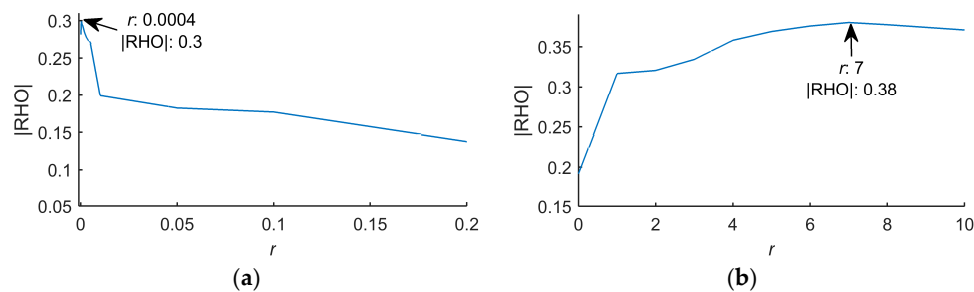


Figure 2. Spearman's correlation coefficient between central tendency measure (CTM) with different r values and the apnoea-hypopnoea index (AHI) for: (a) airflow (AF) and (b) respiratory rate variability (RRV).

2.2.2. Feature Selection

After feature extraction, eight characteristics (SE_1^{AF} , SE_2^{AF} , SE_3^{AF} , CTM^{AF} , SE_1^{RRV} , SE_2^{RRV} , SE_3^{RRV} , and CTM^{RRV}) were obtained. Next, a feature selection phase was then conducted to find an optimum subset of parameters that maximized the diagnostic potential of the irregularity and variability information extracted in the training set. In this regard, we applied the FSLR method proposed by Hosmer and Lemeshow, which is based on the LR classifier [35].

FSLR has already shown its usefulness in previous studies involving SAHS detection in adults [25,40]. This method allows us to explore the original feature space to find an optimal subset without having to evaluate all possible combinations [41]. It finds a feature subset as a result of the sequential incorporation of the most relevant variables, which is followed by the elimination of those that contribute with redundant information to a LR model [42]. The relevance and redundancy of a feature is defined according to the p -value of the likelihood ratio test when it is included or removed

from a specific model [35]. Thus, the algorithm includes a candidate feature in the LR model if the p -value of its likelihood ratio test is less than a certain input threshold α_I . Once included, the method accomplishes a redundancy analysis of the variables in the model, excluding those whose p -value is greater than a certain output threshold α_O . The algorithm ends when none of the candidate variables meet the input condition and none of the included variables meet the output condition [35].

2.2.3. Classification

LR is a standard for binary classification [35]. It allows estimation of the posterior probability of the occurrence of a certain event, defined by a dichotomous dependent variable, as a function of independent predictor variables [35]. In our case, the dependent variable was composed of the classification of the subjects under study in the positive class or negative class, whereas the predictors were the features of the optimum subset previously selected. Thus, LR allowed us to estimate the posterior probability that a subject had SAHS, without making any a priori assumption about the statistical nature of the data. LR follows the expression [35]:

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k}} \quad (8)$$

where $\pi(x)$ is the posterior probability of membership to the SAHS class, β_0 is interceptor, β_i ($i = 1, \dots, k$) are coefficients associated to each predictor variable, and k the number of features in the model. LR estimates β_0 and β_i by the maximum likelihood optimization method.

2.2.4. Statistical Analysis

The features extracted of AF and RRV did not pass the Lilliefors normality test. Hence, the nonparametric Mann–Whitney U test was used to assess statistically differences between each pair of SAHS severity groups. Boxplots were used to show differences in features according to SAHS severity degree. The diagnostic ability of our proposal was assessed in terms of sensitivity (percentage of subjects with the disease correctly classified), specificity (percentage of subjects without the disease correctly classified), and accuracy (percentage of subjects correctly classified) [43,44]. Receiver Operating Characteristics (ROC) curves were also used to evaluate the sensitivity vs. 1-specificity pairs resulting from the variation of the classification threshold obtained from LR [44]. Moreover, the area under the ROC curve (AROC) was calculated to estimate the diagnostic robustness of the model [44]. Ten-fold cross-validation was applied in order to ensure the generalization of our approach, i.e., the results of the logistic regression models are independent of the partition into training and test datasets [45]. Hence, the test group was randomly divided into 10 subsets, with the same proportion of subjects of each class than in the case of the original test set. Nine subsets constituted the training set, and one formed the test set. This process was repeated ten times, until each subset was used as test set.

3. Results

3.1. Training Group

3.1.1. Exploratory Analysis

Figure 3 shows the averaged PSDs of groups $AHI < 1$, $AHI [1, 5)$, $AHI [5, 10)$, and $AHI \geq 10$ of the AF and RRV recordings from the training group. Each of these groups was composed of 80, 87, 35, and 48 paediatric subjects, respectively. Notice that there was no substantial spectral information in frequencies greater than 0.6 Hz in AF and 0.2 Hz in RRV. Hence, the frequency range f_1 – f_2 used to compute SE in AF and RRV was 0–0.6 Hz and 0–0.2 Hz, respectively. In the case of AF, the averaged PSD of the groups $AHI < 1$ and $AHI [1, 5)$ shows scarce differences between them, whereas the group $AHI [5, 10)$ presented some differences around 0.20 Hz and 0.30 Hz. By contrast, large differences

could be qualitative observed between the $\text{AHI} \geq 10$ group and the other groups. It was also observed that the two groups with higher SAHS severity showed flatter spectra. In the case of RRV, only slight differences for the $\text{AHI} \geq 10$ group was observed (around 0.03 Hz and 0.10 Hz).

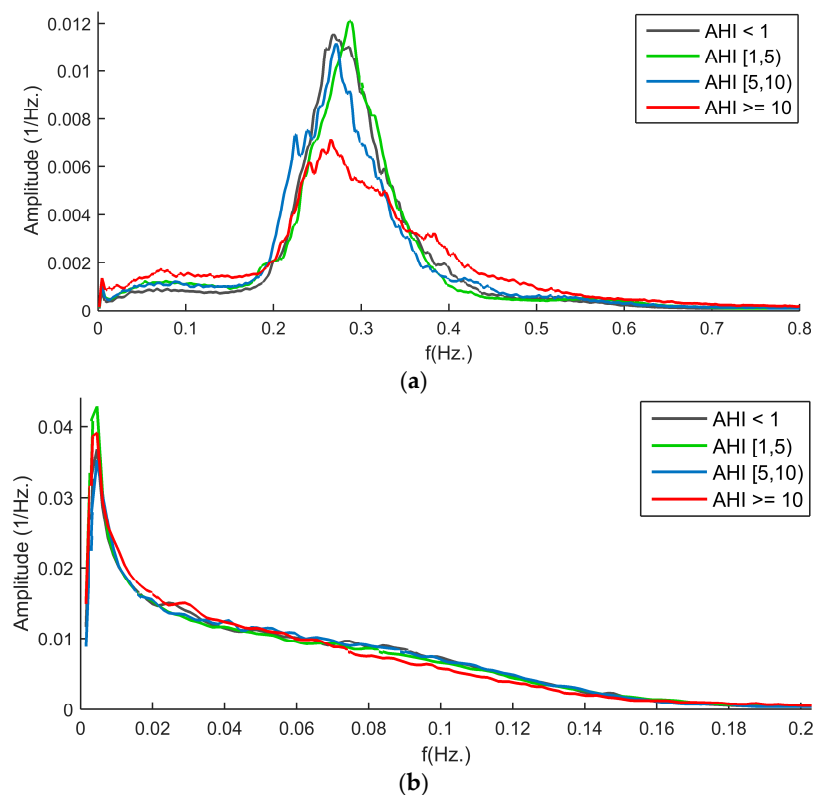


Figure 3. Averaged power spectral density (PSD) of the groups $\text{AHI} < 1$, $\text{AHI} [1, 5)$, $\text{AHI} [5, 10)$, and $\text{AHI} \geq 10$ for (a) AF and (b) RRV.

Figure 4 shows the boxplots of the features extracted from the AF signal in the training group (SE_1^{AF} , SE_2^{AF} , SE_3^{AF} , CTM^{AF}). Statistically significant differences between each pair of severity groups are also shown. Since the size of the groups involved in a hypothesis test comparison influences the p -value, and each of our comparison involved a different number of subjects, two significance thresholds were used (after Bonferroni correction for multiple comparisons): 0.05 (less conservative) and 0.01 (more conservative). The boxplots of the groups $\text{AHI} < 1$ and $\text{AHI} [1, 5)$ did not present visual differences between them in any of the four features. This agreed with the obtained p -values, since these did not show statistically significant differences (neither for p -value = 0.01 nor for p -value = 0.05), even though the number of subjects in this comparison was the highest (167 subjects). Conversely, the boxplots of the group $\text{AHI} \geq 10$ showed higher values than the remaining groups in all features. Thus, its SE and CTM values were closer to 1, which implies greater irregularity and less variability. This was also reflected in the p -values, since SE_2 from group $\text{AHI} \geq 10$ showed significant differences with all groups (p -value < 0.01 with $\text{AHI} < 1$ and $\text{AHI} [1, 5)$ and p -value < 0.05 with $\text{AHI} [5, 10)$). The group $\text{AHI} [5, 10)$ did not show as clear visual differences as the group $\text{AHI} \geq 10$, although its upward evolution in SE_1 , SE_2 , and CTM could be observed too. Moreover, this group had significant differences (p -value < 0.05) with the group $\text{AHI} \geq 10$, even though the number of subjects compared is the lowest (83 subjects).

Figure 5 shows the boxplots of the features extracted from the RRV signal of the training group as well as the statistical differences between each pair of SAHS severity groups. As in the case of AF, the boxplots and p -values of the groups $\text{AHI} < 1$ and $\text{AHI} [1, 5)$ did not show any differences. The highest visual differences were shown in CTM of the $\text{AHI} \geq 10$ group. This was also shown in the p -value

analysis, since the group $AHI \geq 10$ presented significant differences with all groups (p -value < 0.01 with $AHI < 1$ and $AHI [1, 5)$ and p -value < 0.05 with $AHI [5, 10)$). Similarly, the group $AHI [5, 10)$ significantly differed from $AHI < 1$ and $AHI \geq 10$ groups in CTM . However, no visual or significant differences among the spectral entropies of the four groups could be appreciated.

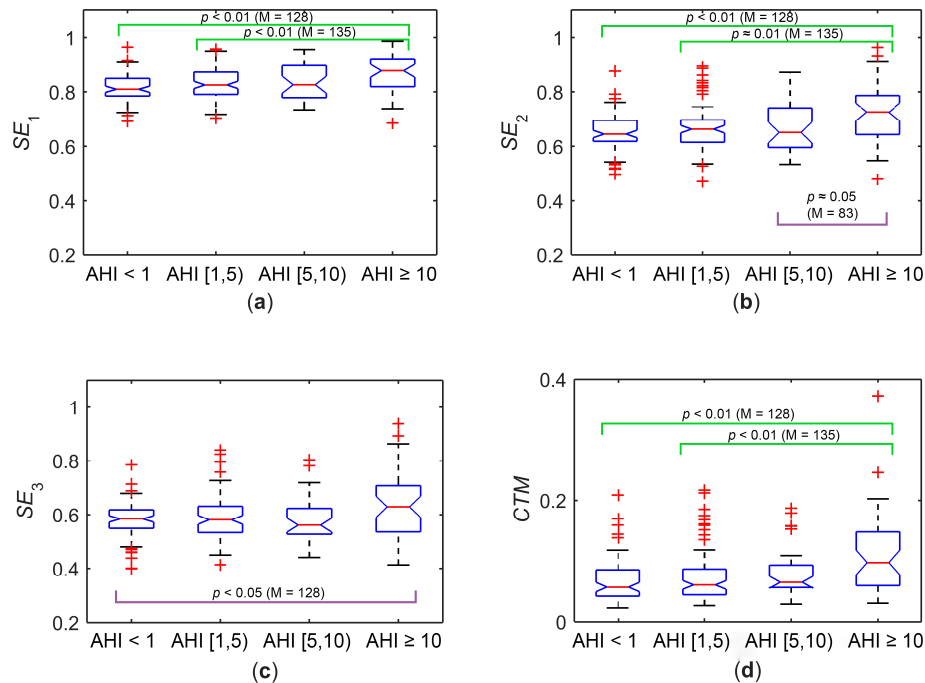


Figure 4. Boxplots of features extracted from AF signals from the training group: (a) Spectral Entropy (SE_1); (b) Quadratic Spectral Entropy (SE_2); (c) Cubic Spectral Entropy (SE_3); and (d) Central Tendency Measure (CTM).

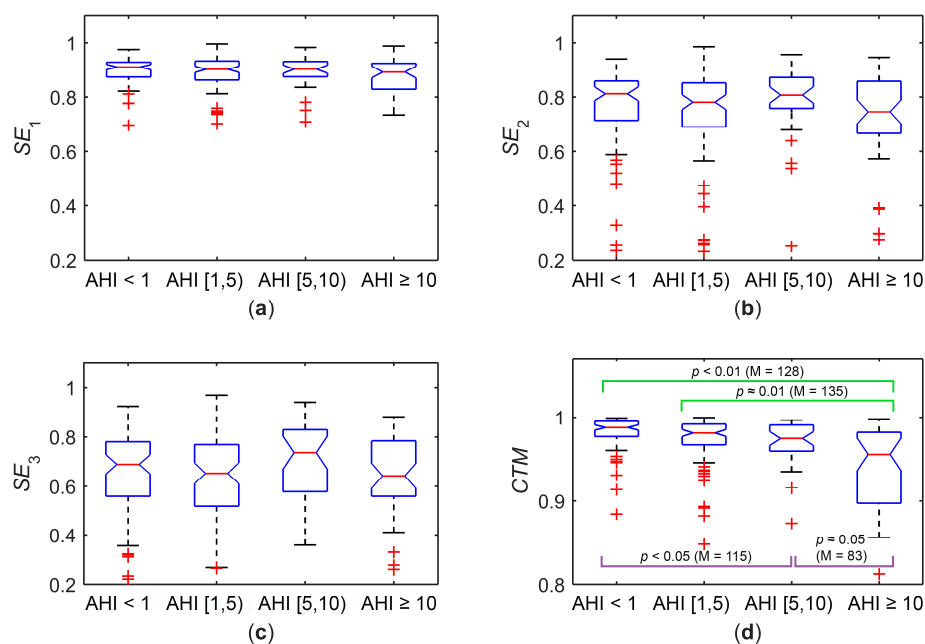


Figure 5. Boxplots of features extracted from RRV signals from the training group: (a) Spectral Entropy (SE_1); (b) Quadratic Spectral Entropy (SE_2); (c) Cubic Spectral Entropy (SE_3); and (d) Central Tendency Measure (CTM).

3.1.2. Selected Features and Model Training

The input to the FSLR selection algorithm were the eight features obtained in the extraction stage (SE_1^{AF} , SE_2^{AF} , SE_3^{AF} , CTM^{AF} , SE_1^{RRV} , SE_2^{RRV} , SE_3^{RRV} , and CTM^{RRV}). FSLR was used for the AHI cutoff points 1, 5, and 10 e/h. SE_1^{AF} and CTM^{RRV} were automatically selected for 1 e/h, while SE_2^{AF} and CTM^{RRV} were selected for 5 and 10 e/h. The optimal subsets obtained for each cutoff point fed three LR classifiers. Hence, for AHI = 1 e/h, a LR model (LR₁) was trained with SE_1^{AF} and CTM^{RRV} , while for AHI = 5 e/h and AHI = 10 e/h two LR models (LR₅ and LR₁₀, respectively) were built with the SE_2^{AF} and CTM^{RRV} values of the subjects from the training set.

3.2. Test Group

All subjects in the test set were considered to compute the diagnostic ability for each cutoff point (1, 5, and 10 e/h), i.e., LR_{cutoff} differentiates subjects with AHI < cutoff from subjects with AHI ≥ cutoff. The diagnostic ability results in the test set of LR₁, LR₅, and LR₁₀ are shown in Table 2 and Figure 6. LR₁ and LR₅ reached 60% Acc (0.590 AROC) and 76% Acc (0.780 AROC), respectively. LR₁₀ outperformed LR₁ and LR₅, reaching 80% Acc and 0.798 AROC.

Table 2. Diagnostic ability of the logistic regression (LR) models for the cutoffs 1 e/h, 5 e/h, and 10 e/h, considering all subjects in the test set. Se: sensitivity; Sp: specificity; Acc: accuracy.

| Model | Se (%) | Sp (%) | Acc (%) |
|------------------|--------|--------|---------|
| LR ₁ | 60.5 | 58.6 | 60.0 |
| LR ₅ | 65.0 | 80.6 | 76.0 |
| LR ₁₀ | 83.3 | 79.0 | 80.0 |

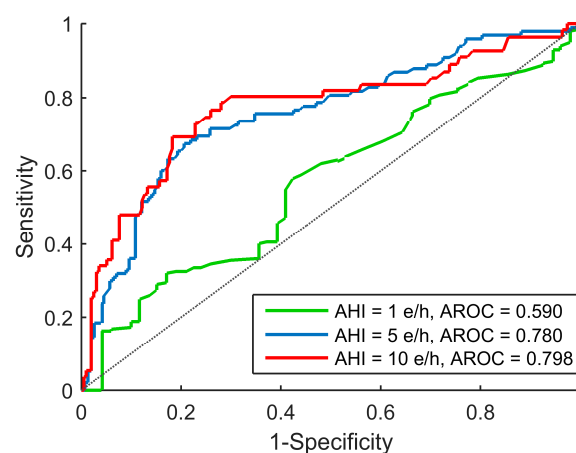


Figure 6. Receiver Operating Characteristics (ROC) curves of LR models obtained with optimum feature subsets in test group for 1, 5 and 10 e/h cutoff points.

4. Discussion

In this paper, we assessed the evolution of irregularity and variability of AF and RRV signals as SAHS severity increases in children. SE and CTM were used to conduct an exploratory analysis that showed increasing irregularity (increasing SE_1^{AF} , SE_2^{AF} and SE_3^{AF}) and decreasing variability (increasing CTM^{AF}) tendencies in AF as AHI is higher. However, the obtained *p*-values showed that these tendencies were only significant from AHI ≥ 5 e/h onwards. In contrast, no irregularity changes in RRV emerged among the SAHS severity groups, although variability of the most severe group was qualitatively higher (CTM^{RRV} was lower). This was supported by *p*-values, since the CTM^{RRV} of the group AHI ≥ 10 presented significant differences with all groups. Additionally, the group AHI [5, 10)

showed statistically significant differences with the $AHI < 1$ and $AHI \geq 10$ groups, even though the number of involved subjects was smaller than in other comparisons (115 and 83 subjects, respectively).

No qualitative differences were appreciated in the peakedness of RRV PSDs among severity groups and only the $AHI [5, 10)$ and $AHI \geq 10$ groups revealed flatter AF PSDs. This fact, along with the lack of significant differences in SE , suggests that apnoeic events do not affect irregularity of RRV signal. Additionally, the groups $AHI < 1$ and $AHI [1, 5)$ did not show significant differences between them in none of the AF or RRV features, even though the number of subjects in these comparisons was the largest (167 subjects). This suggests that when AHI is less than 5 e/h, there is an insufficient number of events to modify the irregularity and variability in AF and RRV. Since AF signals are directly involved in apnoea and hypopnoea definitions, our irregularity analysis may have been supporting those studies that used 5 e/h or 10 e/h as thresholds to establish SAHS or, at least, a high severity degree. The variability analysis of AF and RRV also agreed with this concept, since the CTM statistical differences among severity groups were present in $AHI [5, 10)$ and, especially, in $AHI \geq 10$. Further analyses would be required to assess direct relationships between irregularity and variability of AF and RRV with severity-related symptoms. However, our study of the evolution of irregularity and variability was consistent with those severity groups showing increased risk for major adverse health consequences. Indeed, it has been reported that the cutoff points 5 e/h and 10 e/h establish the limits of those groups that present higher risk of developing morbidities [46]. In this regard, an $AHI \geq 10$ has been associated with cardiac strain, while an $AHI \geq 5$ appears to be associated with an increased C-reactive protein level [10]. Moderate-severe SAHS have been associated with an increased blood pressure [10], as well as increased risk for neurocognitive deficits [47]. In addition, the threshold $AHI \geq 5$ is commonly used to determine whether a surgical treatment is recommended [46]. However, paediatric subjects with severe SAHS are at an increased risk for residual SAHS. In this regard, overnight observation is recommended after adenotonsillectomy in children with $AHI \geq 10$ e/h [12]. Hence, early detection is essential to start immediate treatment in these cases. Moreover, automatic detection of more severe cases would reduce the workload of physicians who are then able to focus on the more dubious cases.

The most notable differences among the severity groups were observed in SE_1^{AF} , SE_2^{AF} , CTM^{AF} , and CTM^{RRV} . Although CTM^{AF} was found redundant during the FSLR automatic selection, the complementarity of the irregularity and variability analyses was shown since SE_1^{AF} , SE_2^{AF} , and CTM^{RRV} were selected for LR_1 (SE_1^{AF} and CTM^{RRV}), LR_5 (SE_2^{AF} and CTM^{RRV}), and LR_{10} (SE_2^{AF} and CTM^{RRV}). This also highlighted the complementarity of the information contained in AF and RRV signals. As expected, LR_1 reached the lowest diagnostic performance: 60% Acc and 0.590 AROC. The LR_5 and LR_{10} models, however, achieved moderate-to-high diagnostic ability, reaching high Acc (76% and 80%) and AROC (0.780 and 0.798). In this regard, LR_{10} outperformed LR_1 and LR_5 in all the diagnostic statistics (Se, Sp, Acc, and AROC). These results agreed with those observed in the exploratory and p -value analyses, since the group $AHI \geq 10$ showed greater irregularity and variability differences when compared to the other groups.

Several previous studies have evaluated the use of a reduced set of biomedical signals to detect SAHS in children. Table 3 shows the performance of previous research in the context of detection of paediatric SAHS. The number of subjects involved in these state-of-the-art studies ranges from very low (21) to moderate (167), which limits the generalizability of their results. A relatively small sample size may be also behind the fact that most of them do not consider the assessment of their methodologies according to SAHS severity. In contrast, we have reliably shown irregularity and variability tendencies as AHI increases by the application of our methods to a large cohort ($n = 501$). In addition, we have shown that irregularity and variability information is able to achieve high diagnostic ability in the more severely affected children.

This study has some limitations. Although the number of subjects is the largest out of those found in the literature, an even greater database would reinforce the universal characteristics of our findings. In addition, the proportion of subjects belonging to each of the four SAHS severity groups may have

been better balanced. However, the proportion of subjects in our severity groups reflects realistic severity ratios of children derived from clinical practice in pediatric sleep units. Another limitation is the aforementioned lack of consensus on the cutoff points that should be used to limit each level of severity. Several studies used different cutoff points (usually 1, 3, 5, and 10 e/h) to discriminate the paediatric subjects with SAHS from those who do not suffer from it [8–10,15–24]. This hinders the comparison with the results from published studies. Nevertheless, our results regarding the evolution of irregularity and variability support those studies that chose 5 e/h and 10 e/h as cutoff points to discriminate moderate-to-severe SAHS [8–10,17–21,24]. Notwithstanding these considerations, AF features with ability to discriminate children with AHI below 1 e/h will be the object of future investigations. In order to complement our analyses, future investigation of AF signals when obtained in the patients' home would be an interesting follow-up study. Finally, the assessment of more complex automatic classifiers, with the ability to maximize the diagnostic ability of the AF-extracted information, is another future goal.

Table 3. Summary of the state-of-the-art in the context of detection of paediatric SAHS.

| Studies | Subjects (n) | Signal | AHI (e/h) | Methods (Analysis/Selection/Classifier) | Se (%) | Sp (%) | Acc (%) |
|------------------------------------|--------------|------------------------|-----------|--|--------|--------|---------|
| Shouldice et al. [16] (2004) | 50 | ECG | 1 | Temporal and spectral analysis/-/QDA | 85.7 | 81.8 | 84 |
| Gil et al. (2009) [17] | 21 | PPG | 5 | Spectral analysis of HRV and DAP events detection/Wrapper methodology/LDA | 87.5 | 71.4 | 80 |
| Lázaro et al. [18] (2014) | 21 | PPG | 5 | Spectral analysis of PRV and DAP events detection/Wrapper methodology/LDA | 100 | 71.4 | 86.7 |
| Gil et al. (2010) [19] | 21 | PPG | 5 | Analysis of PTTV/Wrapper methodology/LDA | 75 | 85.7 | 80 |
| Garde et al. [20] (2014) | 146 | SpO ₂ PRV | 5 | Temporal and spectral analysis/Selection algorithm optimizing the AROC/LDA | 88.4 | 83.6 | 84.9 |
| Dehkordi et al. [21] (2016) | 146 | PPG | 5 | Temporal and spectral analysis/LASSO/LASSO | 76 | 68 | 71 |
| Sahadan et al. [23] (2015) | 93 | PR | 1 | Automatic calculation and analysis of PR parameters/-/-/ | 18 | 97 | 49.5 * |
| Velasco-Suarez et al. [15] (2013) | 167 | SpO ₂ | 1 | Quantification of clusters of desaturations/-/-/ | 86.6 | 98.9 | 93.4 * |
| Gutiérrez-Tobal et al. [22] (2015) | 50 | AF SpO ₂ | 3 1 | Spectral features and oxygen desaturation index of 3% (ODI3)/FSLR/LR | 85.9 | 87.4 | 86.3 |
| Tsai et al. [24] (2013) | 148 | SpO ₂ | 5 | Oxygen desaturation index of 4% (ODI4) /-/-/ | 77.7 | 88.9 | 79 * |
| | | | 10 | | 83.8 | 86.5 | 85.1 * |
| | | | 10 | | 89.1 | 86 | 87.1 * |
| Tan et al. [8] (2014) | 100 | ECG | 1 | Comparison of the AHI obtained from PSG with the AHI directly estimated of respiratory polygraphic (RP)/-/-/ | 82.5 | 90 | 86 * |
| | | AF | 5 | | 62.5 | 100 | 85 * |
| | | SpO ₂ | 10 | | 65 * | 100 * | 93 * |
| | | RIP | 1 | | 60.5 | 58.6 | 60 |
| Our proposal | 501 | AF | 5 | SE and CTM/FSLR/LR | 65 | 80.6 | 76 |
| | | | 10 | | 83.3 | 79 | 80 |

QDA: Quadratic discriminant analysis; HRV: Heart rate variability; DAP: Decreases in amplitude fluctuations of the PPG signal; LDA: Linear discriminant analysis; PRV: Pulse rate variability; PTTV: Pulse transit time variability; AROC: Area under the receiver operating characteristics curves; LASSO: Least absolute shrinkage and selection operator; PR: Pulse rate; FSLR: Forward stepwise logistic regression; LR: Logistic regression model; RIP: Chest and abdominal movement by respiratory inductance plethysmography; SE: Spectral entropy; CTM: Central tendency measure. * Computed from reported data.

5. Conclusions

To the best of our knowledge, this is the first time that irregularity and variability measures from AF and RRV are automatically quantified and combined to assist in paediatric SAHS diagnosis. In addition, this is also the first time that the evolution of irregularity and variability measures as a function of SAHS severity is assessed in this context. We found higher irregularity and lower variability in the AF signals of children with AHI ≥ 5 e/h, as well as higher variability in the RRV signals from these same groups. Regarding the lack of consensus of AHI cutoffs to be used, these results support 5 e/h and 10 e/h as those related to the most important changes in the AF information associated with paediatric SAHS. Our results also showed complementarity between AF and RRV signals, as well as between irregularity and variability analyses. Therefore, LR models trained with optimum irregularity and variability features from both AF and RRV reached moderate-to-high diagnostic ability for AHI = 5 e/h and 10 e/h thresholds. Accordingly, our results suggest that irregularity and variability measures are able to characterize paediatric SAHS in AF recordings. Hence, the use of these

approaches could be helpful to automatically screen for moderate-to-severe SAHS cases in children with a high pre-test probability of disease.

Acknowledgments: This research was supported by ‘Ministerio de Economía y Competitividad’ and ‘European Regional Development Fund’ under project TEC2014-53196-R and RTC-2015-3446-1, by ‘Consejería de Educación de la Junta de Castilla y León and FEDER’ under projects VA037U16 and VA059U13, by ‘European Commission’ and ‘European Regional Development Fund’ under project ‘Análisis y correlación entre el genoma completo y la actividad cerebral para la ayuda en el diagnóstico de la enfermedad de Alzheimer’ (‘Cooperation Programme Interreg V-A Spain-Portugal POCTEP 2014–2020’), and the project 158/2015 de la Sociedad Española de Neumología y Cirugía Torácica (SEPAR). F. Vaquerizo-Villar received a ‘Promoción de Empleo Joven e Implantación de la Garantía Juvenil en I+D+i’ grant from MINECO and the University of Valladolid. D. Álvarez received a Juan de la Cierva grant (IJCI-2014-22664) from MINECO. L. Kheirandish-Gozal is supported by NIH grant 1R01HL130984.

Author Contributions: Verónica Barroso-García processed the signals, analyzed the data, and wrote the manuscript. Gonzalo C. Gutiérrez-Tobal and Roberto Hornero designed the study and interpreted the results. Daniel Álvarez, Fernando Vaquerizo-Villar, Andrea Crespo, and Félix del Campo interpreted the results. Leila Kheirandish-Gozal and David Gozal took part in the diagnosis of subjects and the collection of data. All authors have read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Patil, S.P.; Schneider, H.; Schwartz, A.R.; Smith, P.L. Adult obstructive apnea: Pathophysiology and diagnosis. *Chest* **2007**, *132*, 325–337. [[CrossRef](#)] [[PubMed](#)]
2. 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.; et al. Rules for scoring respiratory events in sleep: Update of the 2007 AASM manual for the scoring of sleep and associated events. *J. Clin. Sleep Med.* **2012**, *8*, 597–619. [[CrossRef](#)] [[PubMed](#)]
3. Kaditis, A.G.; Alonso-Alvarez, M.L.; Boudewyns, A.; Alexopoulos, E.I.; Ersu, R.; Joosten, K.; Larramona, H.; Miano, S.; Narang, I.; Trang, H.; et al. Obstructive sleep disordered breathing in 2–18 year-old children: Diagnosis and management. *Eur. Respir. J.* **2016**, *47*, 69–94. [[CrossRef](#)] [[PubMed](#)]
4. Lipton, A.J.; Gozal, D. Treatment of obstructive sleep apnea in children: Do we really know how? *Sleep Med. Rev.* **2003**, *7*, 61–80. [[CrossRef](#)] [[PubMed](#)]
5. Carroll, J.L.; Loughlin, G.M. Diagnostic criteria for obstructive sleep apnea syndrome in children. *Pediatr. Pulmonol.* **1992**, *14*, 71–74. [[CrossRef](#)] [[PubMed](#)]
6. Marcus, C.L. Obstructive sleep apnea syndrome: Differences between children and adults. *Sleep* **2000**, *23*, 140–141.
7. Spruyt, K. Pediatric Sleep-Disordered Breathing: Criteria and Spectrum of Disease. In *Sleep Disordered Breathing in Children: A Comprehensive Clinical Guide to Evaluation and Treatment*, 1st ed.; Kheirandish-Gozal, L., Gozal, D., Eds.; Humana Press: New York, NY, USA, 2012; pp. 245–260.
8. Tan, H.L.; Gozal, D.; Ramirez, H.M.; Bandla, H.P.R.; Kheirandish-Gozal, L. Overnight polysomnography versus respiratory polygraphy in the diagnosis of pediatric obstructive sleep apnea. *Sleep* **2014**, *37*, 255–260. [[CrossRef](#)] [[PubMed](#)]
9. Alonso-Álvarez, M.L.; Canet, T.; Cubell-Alarco, M.; Estivill, E.; Fernández-Julián, E.; Gozal, D.; Jurado-Luque, M.J.; Lluch-Roselló, M.A.; Martínez-Pérez, F.; Merino-Andren, M.; et al. Consensus Document on Sleep Apnea-Hypopnea Syndrome in Children. *Arch. Bronconeumol.* **2011**, *47*, 2–18.
10. Church, G.D. The Role of Polysomnography in Diagnosing and Treating Obstructive Sleep Apnea in Pediatric Patients. *Curr. Probl. Pediatr. Adolesc. Health Care* **2012**, *42*, 2–25. [[CrossRef](#)] [[PubMed](#)]
11. Epstein, L.J.; Kristo, D.; Strollo, P.J., Jr.; Friedman, N.; Malhotra, A.; Patil, S.P.; Ramar, K.; Rogers, R.; Schwab, R.J.; Weaver, E.M.; et al. Clinical guideline for the evaluation, management and long-term care of obstructive sleep apnea in adults. Adult obstructive sleep apnea Task Force of the American Academy of Sleep Medicine. *J. Clin. Sleep Med.* **2009**, *5*, 263–276. [[PubMed](#)]
12. Marcus, C.L.; Brooks, L.J.; Ward, S.D.; Draper, K.A.; Gozal, D.; Halbower, A.C.; Jones, J.; Lehmann, C.; Schechter, M.S.; Sheldon, S.; et al. Diagnosis and management of childhood obstructive sleep apnea syndrome. *Pediatrics* **2012**, *130*, 714–755. [[CrossRef](#)] [[PubMed](#)]

13. Ryan, P.J.; Hilton, M.F.; Boldy, D.A.; Evans, A.; Bradbury, S.; Sapiiano, S.; Prowse, K.; Cayton, R.M. Validation of British Thoracic Society guidelines for the diagnosis of the sleep apnoea/hypopnoea syndrome: Can polysomnography be avoided? *Thorax* **1995**, *50*, 972–975. [[CrossRef](#)] [[PubMed](#)]
14. Katz, E.S.; Mitchell, R.B.; D'Ambrosio, C.M. Obstructive Sleep Apnea in Infants. *Am. J. Respir. Crit. Care Med.* **2012**, *185*, 805–816. [[CrossRef](#)]
15. Velasco-Suarez, C.T.; Figueroa-Turienzo, J.M.; Len, F.; Mansilla, E. Pulse oximetry recording in children with adenotonsillar hypertrophy: Usefulness in the diagnostic of obstructive sleep apnea syndrome. *Arch. Argent. Pediatr.* **2013**, *111*, 196–201. [[CrossRef](#)] [[PubMed](#)]
16. Shouldice, R.B.; O'Brien, L.M.; O'Brien, C.; Chazal, P.; Gozal, D.; Heneghan, C. Detection of Obstructive Sleep Apnea in Pediatric Subjects using Surface Lead Electrocardiogram Features. *Sleep* **2004**, *27*, 784–792. [[CrossRef](#)] [[PubMed](#)]
17. Gil, E.; Mendez, M.; Vergara, J.M.; Cerutti, S.; Bianchi, A.M.; Laguna, P. Discrimination of Sleep-Apnea-Related Decreases in the Amplitude Fluctuations of PPG Signal in Children by HRV Analysis. *IEEE Trans. Biomed. Eng.* **2009**, *56*, 1005–1014. [[CrossRef](#)] [[PubMed](#)]
18. Lázaro, J.; Gil, E.; Vergara, J.M.; Laguna, P. Pulse rate variability analysis for discrimination of sleep-apnea-related decreases in the amplitude fluctuations of pulse photoplethysmographic signal in children. *IEEE J. Biomed. Health Inform.* **2014**, *18*, 240–246. [[CrossRef](#)]
19. Gil, E.; Bailón, R.; Vergara, J.M.; Laguna, P. PTT Variability for Discrimination of Sleep Apnea Related Decreases in the Amplitude Fluctuations of PPG Signal in Children. *IEEE Trans. Biomed. Eng.* **2010**, *57*, 1079–1088. [[CrossRef](#)] [[PubMed](#)]
20. Garde, A.; Dehkordi, P.; Karlen, W.; Wensley, D.; Ansermino, J.M.; Dumont, G.A. Development of a Screening Tool for Sleep Disordered Breathing in Children Using the Phone Oximeter™. *PLoS ONE* **2014**, *9*, 1–15. [[CrossRef](#)] [[PubMed](#)]
21. Dehkordi, P.; Garde, A.; Karlen, W.; Petersen, C.L.; Wensley, D.; Dumont, G.A.; Ansermino, J.M. Evaluation of cardiac modulation in children in response to apnea/hypopnea using the Phone Oximeter™. *Physiol. Meas.* **2016**, *37*, 187–202. [[CrossRef](#)] [[PubMed](#)]
22. Gutiérrez-Tobal, G.C.; Alonso-Álvarez, M.L.; Álvarez, D.; del Campo, F.; Terán-Santos, J.; Hornero, R. Diagnosis of pediatric obstructive sleep apnea: Preliminary findings using automatic analysis of airflow and oximetry recordings obtained at patients' home. *Biomed. Signal Process. Control* **2015**, *18*, 401–407. [[CrossRef](#)]
23. Sahadan, D.Z.; Davey, M.J.; Horne, R.S.C.; Nixon, G.M. Improving detection of obstructive sleep apnoea by overnight oximetry in children using pulse rate parameters. *Sleep Breath.* **2015**, *19*, 1409–1414. [[CrossRef](#)] [[PubMed](#)]
24. Tsai, C.M.; Kang, C.H.; Su, M.C.; Lin, H.C.; Huang, E.Y.; Chen, C.C.; Hung, J.C.; Niu, C.K.; Liao, D.L.; Yu, H.R. Usefulness of desaturation index for the assessment of obstructive sleep apnea syndrome in children. *Int. J. Pediatr. Otorhinolaryngol.* **2013**, *77*, 1286–1290. [[CrossRef](#)] [[PubMed](#)]
25. Gutiérrez-Tobal, G.C.; Hornero, R.; Álvarez, D.; Marcos, J.V.; del Campo, F. Linear and nonlinear analysis of airflow recordings to help in sleep apnoea-hypopnoea syndrome diagnosis. *Physiol. Meas.* **2012**, *33*, 1261–1275. [[CrossRef](#)] [[PubMed](#)]
26. Han, J.; Shin, H.B.; Jeong, D.U.; Park, K.S. Detection of apnoeic events from single channel nasal airflow using 2nd derivative method. *Comput. Methods Programs Biomed.* **2008**, *98*, 199–207. [[CrossRef](#)]
27. Cysarz, D.; Zerm, R.; Bettermann, H.; Frühwirth, M.; Moser, M.; Kröz, M. Comparison of respiratory rates derived from heart rate variability, ECG amplitude, and nasal/oral airflow. *Ann. Biomed. Eng.* **2008**, *36*, 2085–2094. [[CrossRef](#)] [[PubMed](#)]
28. Poza, J.; Hornero, R.; Abásolo, D.; Fernández, A.; García, M. Extraction of spectral based measures from MEG background oscillations in Alzheimer's disease. *Med. Eng. Phys.* **2007**, *29*, 1073–1083. [[CrossRef](#)] [[PubMed](#)]
29. Chua, K.C.; Chandran, V.; Acharya, U.R.; Min, L.C. Cardiac state diagnosis using higher order spectra of heart rate variability. *J. Med. Eng. Technol.* **2008**, *32*, 145–155. [[CrossRef](#)] [[PubMed](#)]
30. Gutiérrez-Tobal, G.C.; Álvarez, D.; Gómez-Pilar, J.; del Campo, F.; Hornero, R. Assessment of Time and Frequency Domain Entropies to Detect Sleep Apnoea in Heart Rate Variability Recordings from Men and Women. *Entropy* **2015**, *17*, 123–141. [[CrossRef](#)]
31. Atri, R.; Mohebbi, M. Obstructive sleep apnea detection using spectrum and bispectrum analysis of single-lead ECG signal. *Physiol. Meas.* **2015**, *36*, 1963–1980. [[CrossRef](#)] [[PubMed](#)]

32. Abásolo, D.; Hornero, R.; Gómez, C.; García, M.; López, M. Analysis of EEG background activity in Alzheimer's disease patients with Lempel–Ziv complexity and central tendency measure. *Med. Eng. Phys.* **2006**, *28*, 315–322. [[CrossRef](#)] [[PubMed](#)]
33. Cohen, M.E.; Hudson, D.L.; Deedwania, P.C. Applying continuous chaotic modeling to cardiac signals. *IEEE Eng. Med. Biol. Mag.* **1996**, *15*, 97–102. [[CrossRef](#)]
34. Álvarez, D.; Hornero, R.; García, M.; del Campo, F.; Zamarrón, C. Improving diagnostic ability of blood oxygen saturation from overnight pulse oximetry in obstructive sleep apnea detection by means of central tendency measure. *Artif. Intell. Med.* **2007**, *41*, 13–24. [[CrossRef](#)] [[PubMed](#)]
35. Hosmer, D.W.; Lemeshow, S. *Applied Logistic Regression*; John Wiley & Sons: London, UK, 2000.
36. Varady, P.; Micsik, T.; Benedek, S.; Benyo, Z. A novel method for the detection of apnea and hypopnea events in respiration signals. *IEEE Trans. Biomed. Eng.* **2002**, *49*, 936–942. [[CrossRef](#)] [[PubMed](#)]
37. Welch, P.D. The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on time Averaging Over Short, Modified Periodograms. *IEEE Trans. Audio Electroacoust.* **1967**, *15*, 70–73. [[CrossRef](#)]
38. Fleming, S.; Thompson, M.; Stevens, R.; Heneghan, C.; Plüddemann, A.; Maconochie, I.; Tarassenko, L.; Mant, D. Normal ranges of heart rate and respiratory rate in children from birth to 18 years of age: A systematic review of observational studies. *Lancet* **2011**, *377*, 1011–1018. [[CrossRef](#)]
39. Gutiérrez-Tobal, G.C.; Álvarez, D.; del Campo, F.; Hornero, R. Utility of AdaBoost to Detect Sleep Apnea-Hypopnea Syndrome From Single-Channel Airflow. *IEEE Trans. Biomed. Eng.* **2016**, *63*, 636–646. [[CrossRef](#)] [[PubMed](#)]
40. Álvarez, D.; Hornero, R.; Marcos, J.V.; del Campo, F. Multivariate analysis of blood oxygen saturation recordings in obstructive sleep apnea diagnosis. *IEEE Trans. Biomed. Eng.* **2010**, *57*, 2816–2824. [[CrossRef](#)] [[PubMed](#)]
41. Sutter, J.M.; Kalivas, J.H. Comparison of forward selection, backward elimination and generalized simulated annealing for variable selection. *Microchem. J.* **1993**, *47*, 60–66. [[CrossRef](#)]
42. Guyon, I.; Elisseeff, A. An introduction to variable and feature selection. *J. Mach. Learn. Res.* **2003**, *3*, 1157–1182.
43. Flemons, W.W.; Littner, M.R. Measuring agreement between diagnostic devices. *Chest* **2003**, *124*, 1535–1542. [[CrossRef](#)] [[PubMed](#)]
44. Zweig, M.H.; Campbell, G. Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. *Clin. Chem.* **1993**, *39*, 561–577. [[PubMed](#)]
45. Bishop, C.M. *Pattern Recognition and Machine Learning*, 1st ed.; Springer: New York, NY, USA, 2006.
46. Kaditis, A.; Kheirandish-Gozal, L.; Gozal, D. Pediatric OSAS: Oximetry can provide answers when polysomnography is not available. *Sleep Med. Rev.* **2016**, *27*, 96–105. [[CrossRef](#)] [[PubMed](#)]
47. Hunter, S.J.; Gozal, D.; Smith, D.L.; Philby, M.F.; Kaylegian, J.; Kheirandish-Gozal, L. Effect of Sleep-disordered Breathing Severity on Cognitive Performance Measures in a Large Community Cohort of Young School-aged Children. *Am. J. Respir. Crit. Care Med.* **2016**, *194*, 739–747. [[CrossRef](#)] [[PubMed](#)]

