Heart Rate Variability Analysis in Pediatric Obstructive Sleep Apnea

Automated Signal Processing to Identify Biomarkers and Help in the Diagnosis

Doctoral Dissertation

Adrián Martín-Montero

Advisors: Roberto Hornero Sánchez
Gonzalo C. Gutiérrez Tobal

Universidad de Valladolid

Doctoral Program of Information and Telecommunication Technologies

Doctoral Thesis:

Heart Rate Variability Analysis in Pediatric Obstructive Sleep Apnea: Automated Signal Processing to Identify Biomarkers and Help in the Diagnosis

Thesis presented by **Adrián Martín Montero**
to apply for the *Ph.D. degree*
from the *University of Valladolid*

Directed by:
Dr. Roberto Hornero Sánchez y Dr. Gonzalo C. Gutiérrez Tobal.

2024
Valladolid, Spain

A mi familia

Universidad de Valladolid

School of Telecommunications Engineering
Department of Signal and Communications Theory and Telematic Engineering

Research Stay for the International Mention

City:	Berlin (Germany)
Institution:	*Charité - Universitätsmedizin Berlin*
Center:	Interdisciplinary Center of Sleep Medicine
Dates:	31/08/2022-02/12/2022
Duration:	94 days (3 months)
Supervisor:	Prof. Dr. Thomas Penzel

CHARITÉ
UNIVERSITÄTSMEDIZIN BERLIN

National Research Stay

City:	Zaragoza (Spain)
Institution:	University of Zaragoza
Research Group:	Biomedical Signal Interpretation and Computational Simulation
Dates:	15/01/2022-18/02/2022
Duration:	35 days (1 month)
Supervisor:	Prof. Dr. Pablo Laguna Lasaosa

BSICoS
Biomedical Signal Interpretation & Computational Simulation

1542 Universidad Zaragoza

Agradecimientos

A lo largo de los últimos años, he sido testigo de que realizar una Tesis Doctoral es un proceso arduo del que, sin embargo, y a pesar de algunos altibajos, tengo que reconocer que he disfrutado mucho. Esto no hubiera sido así sin el apoyo de una serie de personas a las que lo menos que puedo hacer es dedicar aquí unas palabras.

En primer lugar, a mis directores de tesis, el Dr. Roberto Hornero Sánchez y el Dr. Gonzalo C. Gutiérrez Tobal, por su supervisión, consejo y apoyo a lo largo de todo el proceso. Cuando llegué a este grupo de investigación, prácticamente no sabía lo que era un artículo científico. Gracias por abrirme los ojos desde el principio, haciéndome ver que esto no era un camino de rosas, pero que, si me decidía a recorrerlo, no iba a hacerlo solo. Si hemos llegado hasta aquí, ha sido gracias a vuestra sabiduría y experiencia. También, por qué no decirlo, habéis demostrado tener bastante paciencia. Ha sido un placer trabajar bajo vuestra dirección y espero que el placer haya sido mutuo.

También tengo que extender el agradecimiento por el apoyo tanto científico como personal al resto de miembros del Grupo de Ingeniería Biomédica (GIB) con los que he coincidido durante esta etapa. A Jesús, Carlos, María G., Javier, Pablo, Saúl, Víctor R., Aarón, Víctor G., Marcos, Roberto R., María H., Víctor M., Eduardo, Sergio, Selene, Diego y, especialmente, a mis compañeros de la línea de investigación de apneas, Daniel, Fernando, Verónica, Jorge, Clara y Enrique. En este documento están plasmados los resultados del trabajo científico que hay detrás de todos estos años en el GIB, pero trabajar aquí es infinitamente más que eso. Me habéis ayudado con vuestro consejo profesional, pero también me habéis hecho sentir como en casa desde el primer día, haciendo que muchas veces venir a trabajar no pareciese trabajar. Ha sido un auténtico placer compartir con vosotros este proceso.

También tengo que agradecer todo lo conseguido al Dr. David Gozal y a la

Dra. Leila Kheirandish-Gozal, de la *University of Missouri School of Medicine*, así como al Dr. Félix del Campo del Hospital Universitario Río Hortega de Valladolid. Esta investigación no habría sido posible sin su colaboración y supervisión en los aspectos más clínicos.

Tampoco puedo olvidarme de dar las gracias a los miembros del grupo de investigación Biomedical Signal Interpretation and Computational Simulation (BSICoS) de la Universidad de Zaragoza. Al Dr. Pablo Laguna, al Dr. Eduardo Gil, a la Dra. Raquel Bailón, al Dr. Jesús Lázaro y al bueno de Pablo Armañac. Muchísimas gracias por acogerme en vuestro grupo de investigación, por darme la oportunidad de colaborar con vosotros y por hacer de Zaragoza un segundo hogar para mí.

I would also like to thank Dr. Thomas Penzel, Dr. Niels Wessel, Jan Kraemer and the rest of the members from the Interdisciplinary Center of Sleep Medicine from the *Charité - Universitätsmedizin Berlin* and the Cardiovascular Physics Group from the Humboldt University for letting me do the international internship at their institution. My stay in Berlin was a unique experience, helping me to grow both professionally and personally.

En este sentido, tampoco puedo olvidarme de dar las gracias a Diego, Carmen, y en especial a Laura. Cuando llegué a Berlín para tres meses, pensaba que se me iban a hacer eternos. Gracias a vosotros, esos tres meses se me quedaron demasiado cortos. Me llevo recuerdos para toda la vida.

También estaré siempre agradecido a mis amigos, que sin saber muy bien qué andaba haciendo, me han dado su apoyo y me han ayudado a desconectar cuando más me ha hecho falta. Además, llevan diciéndome "doctor" desde que empecé esta tesis. Verás como ahora se les olvida.

Por último, a quienes más les debo todo esto y nunca podré devolver todo lo que me han dado es a mis padres, Carmen y Ricardo; a mi hermano Richi; a mis sobrinos, Antonio y Olivia; y al resto de miembros de mi familia. Desde que me fui de casa para empezar a formarme, no os habéis cansado nunca de darme vuestro apoyo y cariño incondicionales, acompañasen o no los resultados. Habéis creído en mi cuando yo más dudaba y siempre me habéis dicho que puedo conseguir todo lo que me proponga. Hoy, doce años después, parece que al final teníais algo de razón.

En definitiva, a todos aquellos que me habéis apoyado de una manera u otra, de corazón, gracias.

Abstract

Pediatric obstructive sleep apnea (OSA) is a respiratory disease defined by episodes of complete airflow cessation (apneas) or significant airflow reduction (hypopneas). This disorder is highly prevalent among children, affecting 5.7% of the pediatric population. If left untreated, it can lead to various adverse cardiovascular consequences, including hypertension, hypercholesterolemia, and ventricular hypertrophy, as well as cognitive and developmental impairments. Given the potential threat to the health of children, the early diagnosis and management of OSA are of the utmost importance, however being an underdiagnosed disease.

Overnight polysomnography (PSG) is currently considered the gold standard diagnostic technique for pediatric OSA. This procedure involves children spending an entire night in a sleep laboratory, where numerous sensors are attached to their bodies to record up to 32 biomedical signals. Subsequently, medical experts visually analyze and score apneic events based on these recordings. The number of events of apneas and hypopneas per hour of sleep (e/h) is then calculated, resulting in the apnea-hypopnea index (AHI), which is used for diagnosing and determining the severity of OSA in children. Despite being the gold standard, PSG has certain drawbacks. It is time-consuming, complex, expensive, and highly uncomfortable for the pediatric population, leading to issues such as limited access and long waiting lists. As a result, there has been growing interest in exploring alternative techniques to simplify OSA diagnosis.

In recent years, various approaches have been proposed as alternatives to PSG. One widely studied approach involves analyzing a reduced set of signals recorded during PSG, with a primary focus on the continuous measurement of blood oxygen saturation (SpO_2), derived from photoplethysmography signal. However, there is a noticeable gap in the literature regarding research that explores other biomedical signals, particularly cardiovascular signals, as potential alternatives to PSG. To address this gap, the present Doctoral Thesis aims to com-

prehensively examine the behavior of heart rate variability (HRV) in children, seeking valuable insights into cardiac alterations associated with pediatric OSA that could aid in its diagnosis. Specifically, during apneic events occurring due to OSA, recurrent cardiac behavior, often in the form of bradycardia-tachycardia patterns, has been documented. These alterations are regulated by the autonomic nervous system (ANS), whose activity is typically analyzed through HRV. However, existing HRV analyses in pediatric OSA have primarily relied on conventional approaches designed to evaluate standard ANS behavior, ignoring the specific alterations resulting from the disease. Therefore, in this Doctoral Thesis, we hypothesize that characterizing overnight HRV using novel approaches could uncover previously unknown information about ANS behavior in relation to pediatric OSA, thereby aiding to simplify the diagnosis of the disease. Consequently, the primary objective of this Doctoral Thesis is to conceive, develop, and evaluate novel automated techniques for processing HRV signals, enabling a comprehensive characterization of the alterations in overnight ANS functioning caused by OSA and ultimately contributing to the diagnosis of the disease.

To reach this goal, a total of 2593 electrocardiogram (ECG) recordings from two different databases were analyzed. These recordings were obtained from children aged 0 to 13 years old who had clinical suspicion of OSA. The databases used for this study were the Childhood Adenotonsillectomy Trial (CHAT, 1612 children), which is publicly available, and a private database from the University of Chicago (UofC, 981 children). Prior to analysis, all recordings underwent automated pre-processing to ensure they met the specific requirements of each study, thus following two main approaches: a whole-night analysis and a segment-level analysis, the last aimed to evaluate the influence of sleep stages and the presence of apneic events. Following the pre-processing step, the feature-engineering approach is the unifying methodology among the studies we performed, conducted in three phases. First, in the feature-extraction phase, a variety of HRV parameters specific to pediatric OSA, along with conventional HRV metrics, were extracted using spectral, bispectral, and time domain approaches. The second phase involved feature selection, aiming to identify subsets of features that contained relevant and non-redundant information regarding pediatric OSA alterations. This phase was applied within the bispectral approach, as it was the only study where a large number of features were considered. The final phase of the feature-engineering methodology involved evaluating the clinical utility of the HRV characterization through pattern recognition techniques. This evaluation included assessing the ability of the HRV characterization to diagnose pediatric

OSA and classify sleep stages. To emphasize the importance of the HRV-derived features for sleep staging and OSA detection, an explainable artificial intelligence (XAI) analysis was conducted within the segment-level approach. Additionally, the results obtained from the feature-extraction phase using HRV spectral analysis underwent causal mediation analysis (CMA) to assess their ability to reflect OSA treatment effects.

Each of the HRV characterization approaches has provided valuable insights into previously unknown behaviors of the ANS in the context of pediatric OSA. Through HRV spectral analysis, we identified three OSA-specific frequency bands: BW1 (0.001 – 0.005 Hz), BW2 (0.028 – 0.074 Hz), and BWRes (0.04 Hz around an individual adaptive respiratory peak). The activity of these bands was found to be related to the severity of pediatric OSA and its associated alterations. Importantly, these OSA-specific bands outperformed the classic HRV spectral frequency bands (very low frequency, VLF: 0 - 0.04 Hz; low frequency, LF: 0.04 - 0.15 Hz; and high frequency, HF: 0.15 - 0.4 Hz) in pediatric OSA diagnosis both individually and jointly. Furthermore, the spectral features obtained were used to evaluate the effects of OSA treatment on HRV using CMA, revealing that OSA treatment causally influences HRV activity in both the classic and OSA-specific frequency ranges. Among these bands, BW2 was the only one that showed significant differences in HRV activity between children with and without OSA resolution. Regarding the bispectral analysis, this approach contemplated the inclusion of up to 80 bispectral features. Accordingly, the fast correlation-based filter algorithm allowed us to identify two optimal subsets of bispectral features. One subset was derived from bispectral regions defined based on the classic HRV frequency ranges, while the other was defined based on the HRV OSA-specific frequency ranges. This analysis demonstrated that pediatric OSA leads to alterations in the non-Gaussianity, nonlinearity, and irregularity of overnight HRV. To end with HRV characterization, a segment-level analysis, which considered both the influence of apneic events and sleep stages, revealed that the existent basal sympathetic activation during the rapid-eye movement (REM) sleep stage seems to mask the sympathetic excitation caused by apneic events. It resulted in greater differentiation between apneic event segments during the non-rapid eye movement (NREM) sleep stage compared to REM.

When assessing the clinical utility of this characterization for diagnosing pediatric OSA at three different AHI severity cutoffs (1, 5, and 10 e/h), the multilayer perceptron models trained using bispectral features from the optimal subsets achieved the highest overall diagnostic performance at the 5 and 10 e/h cut-

offs. However, at the lowest cutoff (1 e/h), a least-square boost model trained using segment-level HRV features exhibited the highest performance, making it suitable for detecting pediatric OSA in its mildest form (sensitivity of 90.76% and a positive predictive value of 86.26%). Furthermore, when comparing with state-of-the-art studies using cardiovascular measures, our proposed methodologies achieved similar or higher overall OSA diagnostic performance, especially to detect OSA presence. Regarding the classification of sleep stages in the context of pediatric OSA, the adaptive boosting model reached the highest overall performance in NREM sleep stage segments, thus aligning with the increased differentiation of HRV apneic segments observed during NREM. Furthermore, the use of XAI techniques revealed that two of the novel OSA-specific HRV features developed in this Doctoral Thesis, BW2 and BWRes, reached the highest importance for pediatric OSA diagnosis and sleep stage classification, respectively.

Based on the aforementioned considerations, the different approaches undertaken in this Doctoral Thesis have proven to be valuable in characterizing ANS behavior through HRV analysis in the context of pediatric OSA. These approaches have successfully uncovered specific information regarding the HRV alterations associated with OSA, surpassing the capability of conventional HRV analysis methods used so far to detect the presence and severity of pediatric OSA. Consequently, it can be concluded that the HRV characterization utilizing the novel pediatric OSA-specific HRV features enables the description of ANS alterations resulting from OSA, thereby aiding in the simplification of pediatric OSA diagnosis.

Acronyms

AASM	American academy of sleep medicine
Acc	Accuracy
ACME	Averaged causal mediation effect
AdaBoost	Adaptive boosting
ADE	Averaged direct effect
AF	Airflow
AHI	Apnea-hypopnea index
AI	Apnea index
ANN	Artificial neural networks
ANS	Autonomic nervous system
AT	Adenotonsillectomy
AUC	Area under the receiver-operating characteristics curve
BMI	Body mass index
bpm	Beats per minute
CHAT	Childhood Adenotonsillectomy Trial
CMA	Causal mediation analysis
ECG	Electrocardiogram
EDR	Electrocardiogram-derived respiration
EEG	Electroencephalogram
EMG	Electromyogram
EOG	Electrooculogram
FCBF	Fast correlation-based filter
FFT	Fast Fourier transform
FN	False negatives
FP	False positives
HOS	Higher-order spectra

HR	Heart rate
HRV	Heart rate variability
JCR	Journal Citation Reports
LDA	Linear discriminant analysis
LR+	Positive likelihood ratio
LR-	Negative likelihood ratio
LSBoost	Least-squares boosting
LZC	Lempel-Ziv complexity
MLP	Multi-layer perceptron
MSE	Mean squared error
NPV	Negative predictive value
NREM	Non-rapid eye movement
OAHI	Obstructive apnea-hypopnea index
OSA	Obstructive sleep apnea
PNS	Parasympathetic nervous system
PPG	Photoplethysmography
PPV	Positive predictive value
PRV	Pulse rate variability
PSD	Power spectral density
PSG	Polysomnography
PTT	Pulse transit time
REM	Rapid eye movement
RMSSD	Root mean square of the normal-to-normal successive differences
ROC	Receiving-operating characteristics
RP	Relative power
Se	Sensitivity
SNS	Sympathetic nervous system
Sp	Specificity
SpO_2	Blood oxygen saturation
SU	Symmetrical uncertainty
TAI	Total arousal index
TN	True negatives
TP	True positives
UofC	University of Chicago
WASO	Wake after sleep onset
WWSC	Watchful waiting with supportive care
XAI	Explainable artificial intelligence

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The present Doctoral Thesis is focused on the analysis of heart rate variability (HRV) to characterize and help diagnose pediatric obstructive sleep apnea (OSA). To this purpose, HRV signals have been analyzed through several approaches that have allowed us to evaluate the alterations that take place in the autonomic nervous system (ANS) in OSA presence. Thus, our methodologies have made possible to unravel patterns from HRV signal caused by the disease and unknown until date, being proposed as potential biomarkers of pediatric OSA. The outcomes derived from this research work have been published in four journals indexed in the Journal Citation Reports (JCR) from the Web of Science™. Therefore, this Doctoral Thesis is presented as a compendium of publications.

Section 1.1 of this Doctoral Thesis offers an explanation of the thematic consistency of the included articles. In Section 1.2, the general context of biomedical engineering, biomedical signal processing, and machine learning are concisely outlined. Section 1.3 presents a description of pediatric OSA, including the associated risks and adverse consequences. The limitations of the gold standard for pediatric OSA diagnosis, polysomnography (PSG), are discussed in Section 1.4. Finally, Section 1.5 offers a comprehensive state-of-the-art review of previous research on the simplification of pediatric OSA diagnosis using cardiovascular signal analysis.

1.1 Compendium of publications: thematic consistency

Pediatric OSA is a respiratory sleep disorder characterized by periods of total airflow interruption (apnea) and/or significant airflow decrease (hypopnea) (American Thoracic Society, 1996). This disease is highly prevalent, affecting around 6% of the general pediatric population (Marcus et al., 2012), and has been linked to increased risk for cardiovascular problems, such as left and right ventricular hypertrophy, alterations in autonomic regulation, or changes in blood pressure (Marcus et al., 2012). OSA has also been associated with adverse cognitive consequences, such as learning deficits and reduced academic performance (Hunter et al., 2016), making early detection and treatment essential for maintaining long-term cardiovascular health and academic potential in children. The accepted standard diagnostic test for pediatric OSA is overnight PSG. During this test, children spend a night in a sleep laboratory, being time-consuming, complex and uncomfortable for pediatric population (Tan et al., 2014). Therefore, the drawbacks linked to the PSG have motivated the search for alternatives to diagnose pediatric OSA and study its consequences (Alonso-Álvarez et al., 2015; Marcus et al., 2012; Tan et al., 2014).

In this sense, the research work undertaken during the present Doctoral Thesis has been directed to scrutinize the nature of overnight HRV signal to describe pediatric OSA and help to simplify the diagnosis of the disease. Thus, as shown in Figure 1.1, the analysis of HRV behavior in presence of the disease following different feature-engineering methodologies is the link that connects the four articles included in the compendium of publications.

Different techniques were used to describe HRV, enabling the extraction of pertinent and valuable information from the particular changes induced by pediatric OSA in the signal. In this regard, it was noticed that all the previous works focused on HRV frequency analysis in pediatric OSA context analyzed the classic frequency bands, defined initially based on the standard activity of the ANS (Malik et al., 1996): very low frequency (VLF, 0-0.04 Hz), low frequency (LF, 0.04-0.15 Hz), and high frequency (HF, 0.15-0.4 Hz). Therefore, the first article was intended to define novel frequency bands that were specific of the disease (Martín-Montero et al., 2021b). Following this approach, we were able to identify three frequency ranges (BW1: 0.001-0.005 Hz; BW2: 0.028-0.074 Hz; and BWRes: 0.04 Hz around the individual adaptive respiratory peak) that were associated with pe-

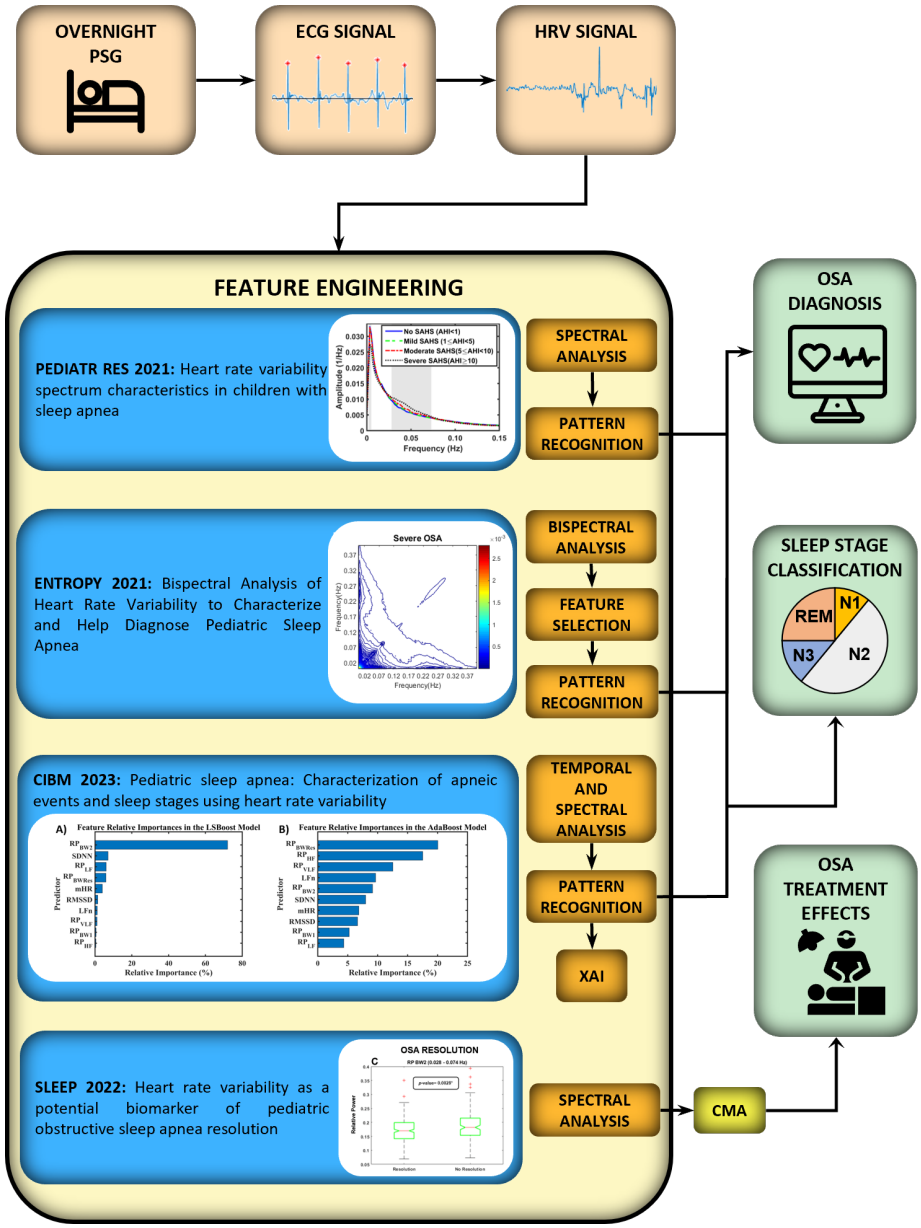

Figure 1.1. Overview of the key contributions made by the papers included in the compendium of publications comprising this Doctoral Thesis. CIBM: Computers in Biology and Medicine, CMA: causal mediation analysis, PEDIATR RES: Pediatric Research, XAI: explainable artificial intelligence.

diatric OSA, and whose activity was found to surpass both individual and combined diagnostic performance of the conventional HRV bands examined so far. In the second article of the Doctoral Thesis (Martín-Montero et al., 2021a), both classic and OSA-specific HRV frequency ranges were used to define regions in the bispectral domain. Through this analysis, we were able to observe alterations in the non-Gaussianity, nonlinearity and irregularity of overnight HRV due to pediatric OSA. In the light of the promising results achieved with the new OSA-specific frequency ranges to characterize OSA alterations, in the third article we decided to carry out a causal mediation analysis (CMA) to evaluate changes in HRV activity due to OSA treatment (Martín-Montero et al., 2022). This methodology showed that OSA treatment causally affects HRV activity in both classic and OSA-specific spectral bands through changes in OSA severity and OSA resolution. Finally, in the last paper of the compendium (Martín-Montero et al., 2023), classic HRV features, as well as the new features developed, were computed in a segment-level analysis to assess the joint influence of apneic events and sleep stages presence. By following this methodology, we observed an increase in basal sympathetic activity during rapid eye movement (REM) sleep stage, which might mask the sympathetic excitation in response to apneic events, causing higher differentiation between apneic segments during non-rapid eye movement (NREM) sleep stage compared to REM.

Hence, the results of the four research works have shown the usefulness of the novel HRV features evaluated in this Doctoral Thesis, to both help in pediatric OSA diagnosis and classify sleep stages (Martín-Montero et al., 2023, 2021a,b), as well as to assess OSA treatment effects (Martín-Montero et al., 2022), highlighting the importance of considering those new OSA-specific features when evaluating HRV in pediatric OSA context.

Below are the titles, authors, and abstracts of the articles included in the compendium of publications for this Doctoral Thesis, as well as the journals in which they were published. To ensure proper comprehension of this document, the access information to the complete articles have been included in Chapter 8.

Martín-Montero et al. (2021b): Heart rate variability spectrum characteristics in children with sleep apnea.

Adrián Martín-Montero, Gonzalo C. Gutiérrez-Tobal, Leila Kheirandish-Gozal, Jorge Jiménez-García, Daniel Álvarez, Félix del Campo, David Gozal, and Roberto Hornero. *Pediatric Research*, vol. 89 (7), p. 1771-1779, 2021. Impact factor in 2021: 3.953, Q1 in "PEDIATRICS" (JCR-WOS).

Abstract: Background: Classic spectral analysis of heart rate variability (HRV) in pediatric sleep apnea-hypopnea syndrome (SAHS) traditionally evaluates the very low frequency (VLF: 0-0.04 Hz), low frequency (LF: 0.04–0.15 Hz), and high frequency (HF: 0.15–0.40 Hz) bands. However, specific SAHS-related frequency bands have not been explored. Methods: One thousand seven hundred and thirty-eight HRV overnight recordings from two pediatric databases (0–13 years) were evaluated. The first one (981 children) served as training set to define new HRV pediatric SAHS-related frequency bands. The associated relative power (RP) were computed in the test set, the Childhood Adenotonsillectomy Trial database (CHAT, 757 children). Their relationships with polysomnographic variables and diagnostic ability were assessed. Results: Two new specific spectral bands of pediatric SAHS within 0–0.15 Hz were related to duration of apneic events, number of awakenings, and wakefulness after sleep onset (WASO), while an adaptive individual-specific new band from HF was related to oxyhemoglobin desaturations, arousals, and WASO. Furthermore, these new spectral bands showed improved diagnostic ability than classic HRV. Conclusions: Novel spectral bands provide improved characterization of pediatric SAHS. These findings may pioneer a better understanding of the effects of SAHS on cardiac function and potentially serve as detection biomarkers.

Martín-Montero et al. (2021a): Bispectral Analysis of Heart Rate Variability to Characterize and Help Diagnose Pediatric Sleep Apnea.

Adrián Martín-Montero, Gonzalo C. Gutiérrez-Tobal, David Gozal, Verónica Barroso-García, Daniel Álvarez, Félix del Campo, Leila Kheirandish-Gozal, and Roberto Hornero. *Entropy*, vol. 23 (8), p. 1016, 2021. Impact factor in 2021: 2.738, Q2 in "PHYSICS, MULTIDISCIPLINARY" (JCR-WOS).

Abstract: Pediatric obstructive sleep apnea (OSA) is a breathing disorder that alters heart rate variability (HRV) dynamics during sleep. HRV in children is commonly assessed through conventional spectral analysis. However, bispectral analysis provides both linearity and stationarity information and has not been applied to the assessment of HRV in pediatric OSA. Here, this work aimed to assess HRV using bispectral analysis in children with OSA for signal characterization and diagnostic purposes in two large pediatric databases (0–13 years). The first database (training set) was composed of 981 overnight ECG recordings obtained during polysomnography. The second database (test set) was a subset of the Childhood Adenotonsillectomy Trial database (757 children). We characterized three bispectral regions based on the classic HRV frequency ranges (very low frequency: 0–0.04 Hz; low frequency: 0.04–0.15 Hz; and high frequency: 0.15–0.40 Hz), as well as three OSA-specific frequency ranges obtained in recent studies (BW1: 0.001–0.005 Hz; BW2: 0.028–0.074 Hz; BWRes: a subject-adaptive respiratory region). In each region, up to 14 bispectral features were computed. The fast correlation-based filter was applied to the features obtained from the classic and OSA-specific regions, showing complementary information regarding OSA alterations in HRV. This information was then used to train multi-layer perceptron (MLP) neural networks aimed at automatically detecting pediatric OSA using three clinically defined severity classifiers. Both classic and OSA-specific MLP models showed high and similar accuracy (Acc) and areas under the receiver operating characteristic curve (AUCs) for moderate (classic regions: Acc = 81.0%, AUC = 0.774; OSA-specific regions: Acc = 81.0%, AUC = 0.791) and severe (classic regions: Acc = 91.7%, AUC = 0.847; OSA-specific regions: Acc = 89.3%, AUC = 0.841) OSA levels. Thus, the current findings highlight the usefulness of bispectral analysis on HRV to characterize and diagnose pediatric OSA.

Martín-Montero et al. (2022): Heart rate variability as a potential biomarker of pediatric obstructive sleep apnea resolution.

Adrián Martín-Montero, Gonzalo C. Gutiérrez-Tobal, Leila Kheirandish-Gozal, Fernando Vaquerizo-Villar, Daniel Álvarez, Félix del Campo, David Gozal, and Roberto Hornero. *SLEEP*, vol. 45 (2), p. zsab214, 2022. Impact factor in 2022: 5.599, Q1 in "CLINICAL NEUROLOGY" (JCR-WOS).

Abstract: Study Objectives: Pediatric obstructive sleep apnea (OSA) affects cardiac autonomic regulation, altering heart rate variability (HRV). Although changes in classical HRV parameters occur after OSA treatment, they have not been evaluated as reporters of OSA resolution. Specific frequency bands (named BW1, BW2, and BWRes) have been recently identified in OSA. We hypothesized that changes with treatment in these spectral bands can reliably identify changes in OSA severity and reflect OSA resolution. Methods: Four hundred and four OSA children (5–9.9 years) from the prospective Childhood Adenotonsillectomy Trial were included; 206 underwent early adenotonsillectomy (eAT), while 198 underwent watchful waiting with supportive care (WWSC). HRV changes from baseline to follow-up were computed for classical and OSA-related frequency bands. Causal mediation analysis was conducted to evaluate how treatment influences HRV through mediators such as OSA resolution and changes in disease severity. Disease resolution was initially assessed by considering only obstructive events, and was followed by adding central apneas to the analyses. Results: Treatment, regardless of eAT or WWSC, affects HRV activity, mainly in the specific frequency band BW2 (0.028–0.074 Hz). Furthermore, only changes in BW2 were specifically attributable to all OSA resolution mediators. HRV activity in BW2 also showed statistically significant differences between resolved and nonresolved OSA. Conclusions: OSA treatment affects HRV activity in terms of change in severity and disease resolution, especially in OSA-related BW2 frequency band. This band allowed to differentiate HRV activity between children with and without resolution, so we propose BW2 as potential biomarker of pediatric OSA resolution.

Martín-Montero et al. (2023): Pediatric sleep apnea: Characterization of apneic events and sleep stages using heart rate variability.

Adrián Martín-Montero, Pablo Armañac-Julián, Eduardo Gil, Leila Kheirandish-Gozal, Daniel Álvarez, Jesús Lázaro, Raquel Bailón, David Gozal, Pablo Laguna, Roberto Hornero and Gonzalo C. Gutiérrez-Tobal. *Computers in Biology and Medicine*, vol. 154, p. 106549, 2023. Impact factor in 2022 (last year available): 7.669, D1 in "MATHEMATICAL AND COMPUTATIONAL BIOLOGY" (JCR-WOS).

Abstract: Heart rate variability (HRV) is modulated by sleep stages and apneic events. Previous studies in children compared classical HRV parameters during sleep stages between obstructive sleep apnea (OSA) and controls. However, HRV-based characterization incorporating both sleep stages and apneic events has not been conducted. Furthermore, recently proposed novel HRV OSA-specific parameters have not been evaluated. Therefore, the aim of this study was to characterize and compare classic and pediatric OSA-specific HRV parameters while including both sleep stages and apneic events. A total of 1,610 electrocardiograms from the Childhood Adenotonsillectomy Trial (CHAT) database were split into 10-minute segments to extract HRV parameters. Segments were characterized and grouped by sleep stage (wake, W; non-rapid eye movement, NREMS; and REMS) and presence of apneic events (under 1 apneic event per segment, e/s; 1-5 e/s; 5-10 e/s; and over 10 e/s). NREMS showed significant changes in HRV parameters as apneic event frequency increased, which were less marked in REMS. In both NREMS and REMS, power in BW2, a pediatric OSA-specific frequency domain, allowed for the optimal differentiation among segments. Moreover, in the absence of apneic events, another defined band, BWRes, resulted in best differentiation between sleep stages. The clinical usefulness of segment-based HRV characterization was then confirmed by two ensemble-learning models aimed at estimating apnea-hypopnea index and classifying sleep stages, respectively. We surmise that basal sympathetic activity during REMS may mask apneic events-induced sympathetic excitation, thus highlighting the importance of incorporating sleep stages as well as apneic events when evaluating HRV in pediatric OSA.

1.2 Framework: biomedical engineering, biomedical signal processing, and machine learning

Biomedical Engineering is an interdisciplinary field that focuses on the application of engineering principles to understand, control, or modify biological systems. It also involves the monitoring of physiological functioning, as well as aiding in the diagnosis and treatment of patients (Bronzino and Peterson, 2014). The scope of this field is broad and covers a wide range of activities such as designing therapeutic and rehabilitation devices and procedures, processing biomedical signals and medical images, creating biological products, or developing biomaterials. In recent years, interest in this interdisciplinary field has grown significantly due to the substantial benefits it offers, providing appropriate tools for the development of a more effective and efficient healthcare system (Bronzino and Peterson, 2014). The present Doctoral Thesis have been aimed at the analysis of HRV signal to extract pertinent information related to pediatric OSA. Accordingly, the primary focus of this work lies within the domain of biomedical signal processing area.

Biomedical signal processing has assumed a critical role in comprehending biological systems through an engineering perspective. The physiological systems of the human body generate a multitude of biomedical signals that reflect their temporal behavior (Bronzino and Peterson, 2014). As such, by analyzing these signals, biomedical engineers can assess the physiological state of the user, identifying significant changes in the biological systems, and detecting pathological conditions. However, understanding the information embedded in raw biomedical signals is challenging and frequently beyond the capabilities of visual inspection. A processing stage, therefore, is needed to extract and decode relevant information. Signal processing techniques aim to extract features that characterize the signals as functions conveying information on the behavior or attributes of the underlying phenomena, using mathematical and information theory methodologies (Najarian and Splinter, 2016; Rangayyan, 2015; Sörnmo and Laguna, 2005). Consequently, signal processing enables the automatic extraction of physiological signal features, which is crucial to understand the information obtained from biological systems. Subsequently, statistical techniques like hypothesis testing or CMA can be used to assess the dissimilarities in the behavior exhibited by features derived from healthy and pathological systems. Furthermore, this feature-extraction step is the first stage for a methodology known as feature engineering,

the ultimate goal of which is the development of automatic diagnostic methods (Najarian and Splinter, 2016).

Feature-engineering techniques are composed of three main stages. The first one is the aforementioned feature-extraction stage. Sometimes, this stage results in a considerable amount of information, which can lead to the evaluation of redundant information and an unnecessary high computational cost, i.e., sub-optimal features set (Rangayyan, 2015). When it occurs, the second stage of the feature-engineering approach, known as feature-selection stage, becomes necessary. In the feature-selection step, automatic statistically based techniques are used to search for significant and non-redundant information among the initial feature group, leading to an optimum feature set (Rangayyan, 2015). Afterwards, the final information can be used in the third stage to train machine-learning models, providing prediction models for automatic diagnosis (Najarian and Splinter, 2016; Rangayyan, 2015).

Machine learning is a specialized area of artificial intelligence that aims to create algorithms and statistical models capable of automatically learning and improving from experience without human intervention (Alpaydin, 2014; Bishop, 2006). Machine-learning algorithms are designed to recognize patterns and relationships in extensive data sets and use that knowledge to make predictions or decisions about new data (Alpaydin, 2014; Bishop, 2006). In a supervised learning approach, this procedure involves training a model on a set of labeled data, where the output is already known, and then using the trained model to make predictions on new, unseen data, assigning a class among different categories (classification), or providing a continuous output (regression) (Alpaydin, 2014). The development of machine-learning algorithms has been a significant driving force behind recent advances in artificial intelligence and has opened up new opportunities for automated decision-making and problem-solving. Notwithstanding, since its origins, machine-learning models have been perceived as a 'black box', which is a clear drawback when the predictions are used in fields such as the health context (Adadi and Berrada, 2018). To solve this limitation, in the last years the concept of 'eXplainable Artificial Intelligence' (XAI) has gained relevance, as methodologies utilized to create transparency and comprehensibility in the decision-making process of artificial intelligence models for human interpretation (Adadi and Berrada, 2018).

The purpose of the present Doctoral Thesis was to provide an exhaustive characterization of the behavior of overnight HRV to obtain significant and valuable information of the alterations caused by pediatric OSA. Accordingly, all the pa-

pers of the compendium have addressed, at least, one of the feature-engineering stages aforementioned to help in the diagnosis of pediatric OSA or evaluate its treatment effects. Furthermore, in the last of these studies, XAI techniques have also been used (Martín-Montero et al., 2023). Therefore, the context above-presented reflects the framework of this research work.

1.3 Pediatric Obstructive Sleep Apnea (OSA)

OSA is a medical condition that occurs when the upper airway is blocked during sleep leading to apnea events, in which there is a total absence of airflow, or hypopnea events, in which there is partial obstruction only. OSA affects both adults and children, but there are significant differences that demands distinction between the two groups (Marcus et al., 2012). The American Academy of Sleep Medicine (AASM) has established rules for scoring apneic events (Berry et al., 2020), whereby an event is considered apneic when there is a complete absence or a reduction of over 90% of airflow for at least 10 seconds in adults. For hypopnea, the reduction in airflow should be between 30-90% for at least 10 seconds. To deal with this complications, the preferred treatment for OSA in adults is continuous positive airway pressure (CPAP) (Marcus et al., 2012; Vlahandonis et al., 2013). In contrast, for children, an event is scored when there is a minimum duration of 2 breathing cycles, instead of 10 seconds (Berry et al., 2020). Additionally, OSA in children is primarily caused by enlarged tonsils and adenoids, leading to the preferred treatment of adenotonsillectomy (AT). Given these differences in disease definition and management, it is essential to study OSA and its consequences in children and adults separately.

OSA has been proven to threaten children quality of life, normal development, and cardiovascular health. The repeated occurrences of apneas and hypopneas result in intermittent hypoxia and hypercapnia, which can trigger systemic inflammation, impaired autonomic regulation, oxidative stress, and endothelial dysfunction (Bhattacharjee et al., 2009; Marcus et al., 2012). These factors contribute to the development of atherosclerosis and increase the risk of hypertension, myocardial infarction, stroke, and other cardiovascular disorders (Bhattacharjee et al., 2009; Marcus et al., 2012; Tauman and Gozal, 2011). Additionally, children diagnosed with OSA have an increased likelihood of developing hypertension, obesity, and insulin resistance, all of which are significant cardiovascular risk factors (Smith and Amin, 2019). These findings emphasize that OSA in children is not merely a respiratory disorder but also a significant risk factor

for cardiovascular disease. Furthermore, OSA in children has also been linked to other adverse outcomes, such as cognitive impairment, decreased academic performance, or increased daytime sleepiness (Hunter et al., 2016). Therefore, the prompt identification and treatment of OSA in children are crucial to anticipate long-term complications.

Apart from the aforementioned consequences, pediatric OSA also poses a significant menace to children due to its high prevalence and underdiagnosis. Epidemiological studies have reported that the prevalence of OSA in children ranges from 1.2% to 5.7%, with the highest prevalence occurring in preschool-aged children and up to 30% in certain high-risk groups, such as those with obesity or craniofacial abnormalities (Lumeng and Chervin, 2008; Marcus et al., 2012). Despite these high numbers, it is estimated that 90% of children with OSA have not been diagnosed yet (Kheirandish-Gozal, 2010). This combination of high prevalence and underdiagnosis highlights the substantial number of children who may experience adverse health effects due to OSA.

1.4 Pediatric OSA diagnosis: Polysomnography (PSG)

The gold standard diagnostic test for pediatric OSA is overnight PSG, which requires the child to spend a night in a specialized pediatric sleep unit while multiple neurophysiological and cardiorespiratory signals are recorded using different body sensors (Marcus et al., 2012). These signals include electroencephalogram (EEG), electromyogram (EMG), electrooculogram (EOG), electrocardiogram (ECG), oronasal airflow (AF), abdominal and chest wall movements, respiratory effort, photoplethysmography (PPG), and continuous measurement of blood oxygen saturation (SpO_2) from PPG (Jon, 2009; Tan et al., 2014). Sleep medical specialists visually inspect the recordings for manual scoring of apnea and hypopnea events, and the number of apneas and hypopneas per hour of sleep, known as the apnea-hypopnea index (AHI), is used to diagnose and classify OSA severity in children (Berry et al., 2020). AHI thresholds of 1, 5, and 10 events per hour (e/h) are frequently utilized to classify the severity of obstructive sleep apnea (OSA) in children as none (AHI $<$ 1 e/h), mild (1 e/h $\leq$ AHI $<$ 5 e/h), moderate (5 e/h $\leq$ AHI $<$ 10 e/h), or severe OSA (AHI $\geq$ 10 e/h) (Hornero et al., 2017; Luz Alonso-Álvarez et al., 2011; Tan et al., 2014). It is noteworthy that these thresholds are more cautious than those used in adults to distinguish

between different levels of OSA severity (Kapur et al., 2017).

Although PSG is widely regarded as the gold standard for diagnosing childhood OSA, it presents certain limitations. Due to the need for recording multiple signals, PSG is a complex test, requiring the child to spend at least one night in a sleep laboratory with qualified technicians present, resulting in high hospital expenses (Alonso-Álvarez et al., 2015; Tan et al., 2014). Visual scoring of apneic events is also labor-intensive and may lead to subjective diagnoses (Tan et al., 2015). In addition, PSG requires the attachment of multiple body sensors to children, which can be uncomfortable (Alonso-Álvarez et al., 2015; Jon, 2009). Furthermore, the fact that children sleep in an unfamiliar environment may lead to sleep recordings that are not a true representation of natural sleep, thus resulting in the need to repeat the diagnostic test (Katz et al., 2012).

The diagnosis and treatment of children affected by OSA are hindered by long waiting lists, complexity, cost, time required to analyze sleep signals, and the rest of drawbacks of the PSG aforementioned (Tan et al., 2015). As a result, there is a need for simplified screening tests to address the unavailability of PSG. The use of portable monitoring equipment has been proposed as the primary substitute for PSG in diagnosing pediatric OSA, addressing the limitations of the latter (Kaditis et al., 2016; Marcus et al., 2012). The devices used for sleep studies are categorized into four types, based on the number and type of the recorded signals, as defined by the Portable Monitoring Task Force of the AASM (Standards of Practice Committee of the American Sleep Disorders Association, 1994):

1. **Type I:** is the conventional PSG equipment and is considered the gold standard to which the other types must be compared. This equipment is capable of recording up to 32 biomedical signals.

2. **Type II:** the comprehensive portable PSG, records at least seven channels that allows for the identification of sleep stages and the calculation of the AHI. The channels required to this purpose are AF, ECG or heart rate (HR), EEG, EMG (located on the chin), EOG, respiratory effort, and SpO_2.

3. **Type III:** the modified portable sleep apnea testing, involves the recording of ventilation (i.e., at least two respiratory movement recordings or a respiratory movement along with AF), ECG or heart rate, and SpO_2. These studies are also known as respiratory polygraphy.

4. **Type IV:** characterized by the continuous recording of one or two physiological signals, usually AF and/or SpO_2. Any device that fails to satisfy the

criteria of Type III is also classified under this type.

Therefore, the advantages that portable monitoring offers compared to PSG are a reduction in the set of channels used, and its realization at children home, leading to a reduction in costs and complexity, and an increase in the comfortability of the test (Chiner et al., 2020; Tan et al., 2014). However, the American Academy of Pediatrics guidelines recommend caution when using alternative tests, as more conclusive evidence is needed about their efficacy (Marcus et al., 2012). Thus, there has been a search and development of alternative methods to simplify and accelerate pediatric OSA diagnosis before the consequences get worse (Alonso-Álvarez et al., 2015).

1.5 Literature review: cardiovascular signal analysis in pediatric OSA context

The seek for alternative diagnostic methods for pediatric OSA has focused on analyzing a reduced set of signals involved in PSG (Gutiérrez-Tobal et al., 2022). In this regard, the evaluation of cardiovascular signals has gained relevance to characterize pediatric OSA and aid in its diagnosis. The rationale behind studying cardiovascular signals to evaluate OSA is that apneic events are typically associated with progressive bradycardia followed by an abrupt tachycardia at the end of the event (Guilleminault et al., 1984). While these cyclic variations are well-established in adults, they also occur in children, albeit with a high degree of variability depending on the presence, type, and duration of the apneic events, leading to increased cardiovascular risks (Aljadeff et al., 1997; Gozal et al., 2013; O'Driscoll et al., 2009; Vitelli et al., 2016). Therefore, these characteristic changes in heart rate during sleep reflect the direct impact of OSA on ANS behavior.

The ANS plays a crucial role in regulating cardiac activity. This is composed of two branches: the sympathetic nervous system (SNS), and the parasympathetic nervous system (PNS). While SNS raises heart rate and contractility, PNS diminishes it (Acharya et al., 2006). The interplay between both branches maintains appropriate cardiac output and blood pressure in response to physiological demands, such as the above mentioned OSA alterations (Acharya et al., 2006; Gozal et al., 2013). Thus, evaluating ANS functioning during sleep can improve the understanding of the cardiovascular implications of childhood OSA. The most widespread method to assess ANS status non-invasively is HRV, a signal derived from ECG that quantifies the fluctuations in the time between consecutive heart-

beats, allowing the monitorization of both ANS branches modulation (Acharya et al., 2006). Figure 1.2 shows HRV signals for different situations that are displayed in both the time and frequency domains. It can be observed that there exists a distinct HRV pattern that appears during apneic events, which is not present during normal sleep (see Figures 1.2.a and 1.2.b). Additionally, the alterations due to these periodic patterns are captured through HRV spectral analysis, resulting in differences reflected in the power spectral density (PSD) (see Figure 1.2.c).

Alternative to the ECG acquisition, several studies have examined the use of PPG signals. PPG is a non-invasive optical method that measures changes in blood volume through the skin surface. By using a light source and a detector, it is possible to measure pulse rate and blood oxygen saturation, being an easy-

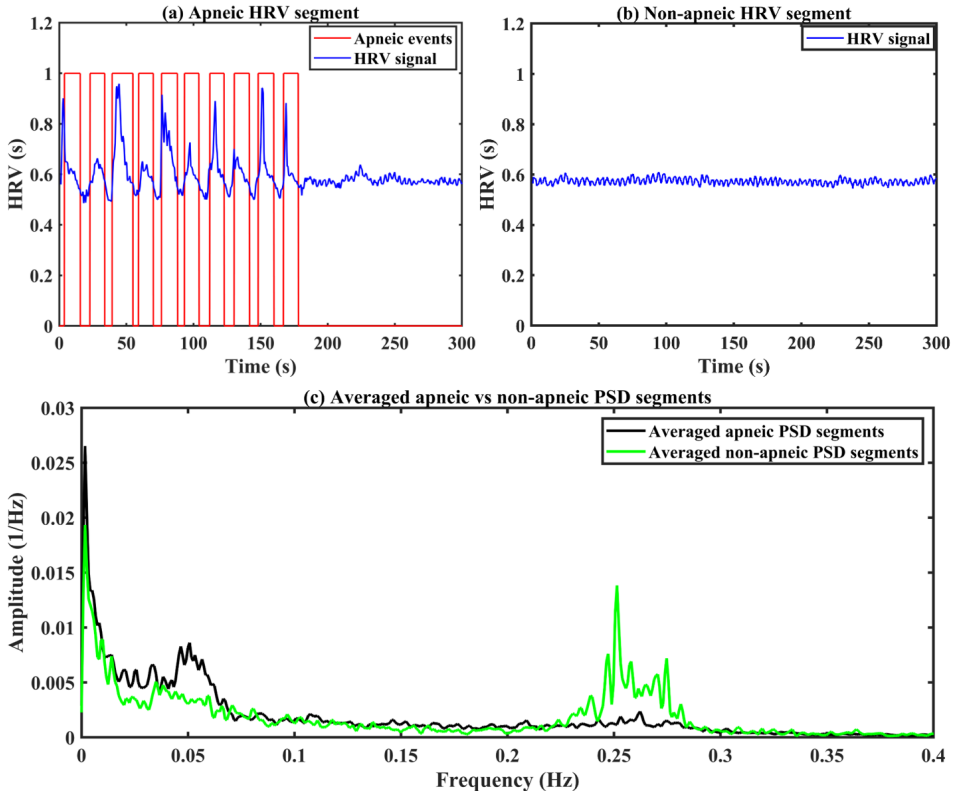

Figure 1.2. Examples of HRV from a OSA child corresponding to: (a) 5-minute segment of HRV during apneic events, (b) 5-minute segments of HRV without apneic events, (c) averaged PSD of all HRV segments across night, separated by those with presence of apneic events (black line) and those without apneic events (green line).

to-acquire signal (Allen, 2007). As PPG is related to arterial vasodilatation or vasoconstriction generated by the ANS, it can capture ANS alterations too (Allen, 2007; Gil et al., 2009). PPG has been widely investigated to extract pulse rate variability (PRV), which can serve as a surrogate measure of HRV (Georgiou et al., 2018). Similar to HRV, PRV is determined by the variation of the pulse-to-pulse intervals calculated from the PPG, allowing for the analysis of ANS alterations under different conditions (Georgiou et al., 2018; Mejía-Mejía et al., 2020). Finally, by combining the information provided by ECG and PPG signals, pulse transit time (PTT) has also been assessed. PTT refers to the time interval between the ECG R-peak and the corresponding wave-peak in the PPG signal (Mukkamala et al., 2015), and it can be influenced by factors such as continuous blood pressure or HR, being altered accordingly to the presence of apneic events (Katz et al., 2003; Lázaro et al., 2014). Given the usefulness of the cardiovascular signals mentioned above to assess ANS alterations, several studies have explored their potential for characterizing, evaluating the implications of treatment, and automatically diagnosing pediatric OSA.

1.5.1 Cardiovascular signal processing methods to characterize pediatric OSA

The assessment of pediatric OSA through the analysis of cardiovascular signals has mainly relied on traditional metrics, particularly those derived from HRV in the frequency domain. The most widespread evaluated parameters in this domain are the activity in the LF and HF frequency bands, as well as the LF/HF ratio in HRV (Baharav et al., 1999; Liao et al., 2010a,b; Nisbet et al., 2013; Van Eyck et al., 2016; Vlahandonis et al., 2014; Walter et al., 2013; Wu et al., 2022). The LF band is believed to reflect both PNS and SNS activity, whereas HF is considered to reflect PNS activity (Acharya et al., 2006; Shaffer and Ginsberg, 2017). Hence, LF/HF is commonly reported as an indicator of sympathovagal balance (Malik et al., 1996), although this assumption should be approached with caution (Stein and Pu, 2012). In a less degree, some previous studies have also examined, alongside the previously mentioned parameters, the VLF band in both HRV (Kwok et al., 2011) and PRV (Dehkordi et al., 2016; Garde et al., 2014, 2019), whose interpretation is more controversial, being assumed to reflect slow regulatory mechanisms (Shaffer and Ginsberg, 2017). Also in the frequency domain, Horne et al. (2018) conducted a study comparing cardiovascular activity between normal weight and overweight children with OSA and controls by combining

HRV evaluation with averaged PTT across the night.

In addition to the conventional frequency metrics, several studies have analyzed cardiovascular signals in the temporal domain to characterize pediatric OSA. The principle behind the assessment of HRV and PRV in the frequency domain is similar, as the former is based on R-R intervals and the latter on the pulse-to-pulse intervals. Thus, extracted from the normal-to-normal (NN) intervals from HRV (Liao et al., 2010a,b; Nisbet et al., 2013; Van Eyck et al., 2016) and PRV (Dehkordi et al., 2016; Garde et al., 2014), previous studies have analyzed its mean (*mNN*), directly influenced by mean HR, standard deviation (*SDNN*), as an overall ANS measure, and root mean square of its successive differences (*RMSSD*), mainly influenced by PNS (Malik et al., 1996). Although less common, several studies have also included the evaluation of the standard deviation of 5-min means of NN intervals (*SDANN*) and the percentage of differences between adjacent NN intervals that are over 50 msec (*pNN50*) in their analysis (Kwok et al., 2011; Wu et al., 2022), measuring long-term and short-term components of HRV, respectively (Malik et al., 1996). Besides the thoroughly cardiovascular analysis using conventional metrics from the temporal and frequency domain in pediatric OSA context, a non-linear approach has also been performed in the studies conducted by Aljadeff et al. (1997) and Vitelli et al. (2016) through Poincaré plots to investigate ANS imbalance in children suffering from OSA. The non-stationarity of PRV was also evaluated by Dehkordi et al. (2016) via detrended fluctuation analysis, in addition to the spectral analysis aforementioned. Despite some controversy in the findings among all the previous studies, there is a consistent trend of sympathetic dominance, along with parasympathetic inhibition, leading to a progressive sympathovagal imbalance during apneic events in pediatric OSA. However these findings were based on the conventional HRV metrics that were originally designed to assess the typical ANS behavior. In contrast, this Doctoral Thesis has introduced novel HRV metrics that are specifically designed to evaluate the alterations caused by pediatric OSA (Martín-Montero et al., 2021a,b).

Several of the previous studies on cardiovascular activity in pediatric OSA have assessed alterations during sleep using a whole night approach (Aljadeff et al., 1997; Garde et al., 2014, 2019; Liao et al., 2010a; Van Eyck et al., 2016; Vitelli et al., 2016). Notwithstanding, cardiac behavior varies between wakefulness and sleep and is influenced by sleep stages (Bonnet and Arand, 1997; Qin et al., 2021). In this regard, some works have compared HRV patterns across sleep stages between children with OSA and healthy children to analyze alterations related to pediatric OSA. Among those assessing pediatric OSA across sleep stages, some

studies did not differentiate between signals segments with and without respiratory episodes (Baharav et al., 1999; Horne et al., 2018; Walter et al., 2013; Wu et al., 2022), whereas others excluded segments containing apneic events, thereby losing the modulation caused by these episodes (Liao et al., 2010b; Nisbet et al., 2013; Vlahandonis et al., 2014). Among all the previous work, the study developed by Dehkordi et al. (2016) was the only one considering simultaneously sleep stage (REM or NREM), and differentiating apneic and non-apneic segments, but mixing all apneic segments without differentiating its severity. In contrast, Kwok et al. (2011) followed a different approach, evaluating HRV alterations in pediatric OSA during the daytime. Differing from these previous studies, in one of our works HRV was characterized in pediatric OSA by considering both sleep stages and presence of apneic events, while distinguishing between several degrees of severity (Martín-Montero et al., 2023).

1.5.2 Pediatric OSA treatment effects in ANS

In light of the abnormalities that occur in the ANS in the presence of childhood OSA, several works have investigated the ability of OSA treatment to reverse these alterations. As it has been previously commented, the preferred treatment for pediatric OSA is AT. Thus, some works have evaluated AT effects on traditional cardiovascular metrics. The work developed by Şaylan et al. (2011) reported that, in a reduced cohort of 15 OSA and 15 healthy children, altered HRV in pre-AT patients apparently did not differ significantly after AT. However, other studies analyzing AT effects on cardiovascular signals have reported that the alterations caused by OSA in the ANS revert following AT, as reflected by a reduction of the sympathetic tone measured through HRV analysis (Kaditis et al., 2011; Muzumdar et al., 2011), PRV (Constantin et al., 2008) or a combination of HRV and PTT (El-Hamad et al., 2017). The study performed by Pavone et al. (2017) posed an interesting approach, trying to predict OSA patients that were referred to AT through PRV and pulse rate analysis. In addition to the analysis of AT effects on cardiovascular measures, Kirk et al. (2020) conducted a study to examine the impact of non-invasive ventilation treatment on HRV. The study reported that, despite a sympathetic-parasympathetic imbalance around arousal events in 12 obese OSA children, no overall treatment effect was found in the classic HRV measures.

Finally, to assess treatment effects, prior works have also taken advantage of the benefits offered by the Childhood Adenotonsillectomy Trial database

(CHAT). This was a prospective randomized controlled trial developed to assess the efficacy of AT versus a strategy of watchful waiting with supportive care (WWSC) for pediatric OSA treatment (Marcus et al., 2013; Redline et al., 2011). Liu et al. (2018) analyzed symbolic pattern dynamics directly from the ECG to evaluate both treatment arm effects. Quante et al. (2015) evaluated heart rate and cardiometabolic measures by assessing differences between both treatment groups. Baumert et al. (2016) evaluated differences in respiratory rate, respiratory sinus arrhythmia and heart rate. In all the previous works, there exist an overall absence of significant differences between treatment arms, highlighting that cardiovascular improvements were due to OSA resolution rather than the intervention itself. Curiously, the original CHAT study found that whereas only a portion of the AT population resolved OSA, half of the children from the WWSC group also resolved OSA (Marcus et al., 2013). Thus, Isaiah et al. (2020) evaluated changes in HRV conventional measures related to pediatric OSA treatment using CMA and reported no causal mediation effects in those parameters, regardless treatment arm. As the novel OSA-specific HRV spectral measures developed in this Doctoral Thesis had not been evaluated for assessing pediatric OSA treatment effects, a CMA has been conducted using the CHAT database to assess those parameters as potential treatment effect biomarkers on HRV (Martín-Montero et al., 2022).

1.5.3 Automatic cardiovascular signal analysis to diagnose pediatric OSA

The information contained in the cardiovascular signals about the alterations in the ANS consequence of pediatric OSA can be used to aid in automatic diagnosis of this disease. Therefore, several works have scrutinized the efficacy of automated analysis of these signals in pediatric OSA context. These investigations have been focused in two directions: detection of apneic episodes and pediatric OSA diagnosis.

Framed in the first approach, Cohen and de Chazal (2015) conducted automated event detection by distinguishing between 30-second epochs of apnea or normal breathing. They achieved this by using temporal and frequency-based features extracted from HRV through a linear discriminant classifier. Similarly, Garde et al. (2017) compared performance of multivariate logistic regression models developed based on correntropy spectral density against PSD analysis of PRV to detect 1-minute epochs with apneic events. In both studies, the authors

noted that combining information obtained from cardiovascular signals along with SpO_2 features led to an improvement in the diagnostic performance of the models.

Other studies have used apneic event detection as a preliminary step to diagnose pediatric OSA. Shouldice et al. (2004) combined segmentation and characterization in the temporal and frequency domain of HRV and electrocardiogram-derived respiration (EDR) signals, achieving the highest accuracy through 1-minute segments model of HRV alone. In addition, three studies from a research group performed per-subject classification based on event detection. Specifically, Gil et al. (2009) diagnosed patients using HRV spectral parameters extracted from decreases in amplitude fluctuations of the PPG. In a subsequent study, Gil et al. (2010) analyzed temporal evolution of time-frequency PTT variability measures to classify events as apneic or non-apneic, providing later a clinical diagnosis for each subject. On the other hand, Lázaro et al. (2014) used conventional frequency analysis of PRV signal to the same purpose.

Lastly, some authors directly reported per-subject classification of pediatric OSA patients. Dehkordi et al. (2016) utilized PRV characterization via temporal and frequency analysis, along with detrended fluctuation analysis, to evaluate its clinical usefulness in diagnosing pediatric OSA. Garde et al. (2014, 2019) extracted PRV and SpO_2 features in the temporal and frequency domains to detect the presence of pediatric OSA (Garde et al., 2014), and to establish different degrees of the disease (Garde et al., 2019). Again, these studies achieved the highest diagnostic performance when combining cardiovascular information with SpO_2 signals. However, in a recent systematic review (Gutiérrez-Tobal et al., 2022), it has been emphasized that the vast majority of studies on machine-learning applications for diagnosing pediatric OSA were based on SpO_2 signal analysis, revealing a gap in research focusing on other biomedical recordings, such as cardiovascular signals, as alternatives to PSG. To help filling this gap in the literature, in this Doctoral Thesis the analysis of exclusively HRV features to aid in pediatric OSA diagnosis has been developed. Thus, we have conducted an evaluation of individual and joint diagnostic performance of the novel metrics presented to diagnose pediatric OSA, and a comparison against traditional HRV metrics (Martín-Montero et al., 2021a,b). Additionally, the potential of HRV segments characterization to diagnose pediatric OSA has been evaluated in Martín-Montero et al. (2023), as well as its potential to classify sleep stages, a task that, to the best of our knowledge, had not been performed earlier through HRV analysis in the context of pediatric OSA.

Chapter 2

Hypotheses and objectives

The preceding chapter has been devoted to highlight the negative impact of pediatric OSA on cardiovascular health, emphasizing the significance of early diagnosis and disease management to prevent its complications. Accordingly, this doctoral thesis aims to characterize specific overnight HRV behavior using different methodological approaches to enhance the understanding of OSA implications in ANS and facilitate disease diagnosis. Thus, Section 2.1 outlines the hypotheses that have been stated across the realization of this Doctoral Thesis. Likewise, Section 2.2 delineates the primary goal of this research, along with the specific objectives established to achieve it.

2.1 Hypotheses

The occurrence of apneic events results in recurrent oscillations in the HRV signal (see Figure 1.2.a). As stated in Section 1.5, these cyclic variations have been well-documented in adults, taking the form of a bradycardia-tachycardia pattern (Guilleminault et al., 1984), which also occurs in children but with greater variability based on the characteristics of the breathing pauses (Aljadeff et al., 1997; Gozal et al., 2013; O'Driscoll et al., 2009; Vitelli et al., 2016). Despite these specific cardiac alterations, the periodic behavior of HRV in pediatric OSA has typically been assessed using spectral analysis within the fixed boundaries of classic HRV frequency bands (Malik et al., 1996). However, a prior study in adults revealed that OSA-related alterations are reflected in a spectral range extending between the VLF and LF bands, suggesting the existence of specific HRV frequency bands for OSA (Gutiérrez-Tobal et al., 2015). Accordingly, at the starting point of this

Doctoral Thesis, it was hypothesized that *novel feature-extraction approaches would enable the characterization of overnight HRV patterns in children with OSA in a more specific manner*.

Additionally, as mentioned in Subsection 1.5.2, the proved alterations in the ANS resulting from pediatric OSA have stimulated the investigation of treatment effects on cardiovascular signals, with some researchers reporting a reversal of the effects that manifests mainly as a return of sympathetic tone to normal levels (Constantin et al., 2008; El-Hamad et al., 2017; Kaditis et al., 2011; Muzumdar et al., 2011). As these findings were derived from the conventional analysis of cardiovascular signals, we assume that *improvements in the severity and resolution of pediatric OSA due to treatment may induce modifications in the HRV features specific to this disease*.

Finally, automated methods have been developed as a potential alternative to PSG for the diagnosis of pediatric OSA, with most of these methods focusing on features derived from the SpO_2 signal (Gutiérrez-Tobal et al., 2022). However, due to the limited presence of alternatives in the literature for the automated diagnosis of OSA in children that extend beyond SpO_2-based models, we wonder if *the novel features extracted from HRV can provide complementary information that, in combination with machine-learning models, can be useful to help in the automatic detection and determination of the severity of OSA in pediatric patients*.

The aforementioned statements constitute the principal hypotheses that comprise the fundamental basis of the present Doctoral Thesis, which can be synthesized into the subsequent global hypothesis:

> "*The characterization of HRV signal through new approaches could reveal unknown information of the overnight ANS behavior linked to pediatric OSA, helping to simply the diagnosis of the disease*"

2.2 Objectives

The primary aim of the present Doctoral Thesis is the development of new HRV signal processing techniques that allow a comprehensive description of the alterations arising in the overnight ANS of OSA children, and aiding in the diagnosis of the disease. In order to reach this main goal, the next specific objectives have been established:

I. To identify novel HRV features that contain significant and non-redundant information about ANS in the presence of apneic events in children with

OSA.

II. To assess the capability of the novel OSA-specific HRV features to reflect the impacts of OSA treatment in the ANS of the affected children.

III. To conceive and assess machine-learning models to maximize the clinical usefulness of the novel HRV characterization, to ascertain the existence and severity of the disease, as well as to classify sleep stages through the employment of optimal sets of HRV features.

Chapter 3

Subjects and signals under study

During the present Doctoral Thesis, we investigated two different databases of pediatric ECG recordings: a private and a public one. The private database, named University of Chicago (UofC) database, was obtained from the Comer Children's Hospital of the University of Chicago (Chicago, IL, USA), while the public database was the CHAT database (Marcus et al., 2013; Redline et al., 2011). Both of them included ECG recordings of children ranging from 0 to 13 years of age who had clinical suspicion of OSA due to a range of symptoms, including snoring, apneas, excessive daytime sleepiness, and tonsillar hypertrophy, among others. The CHAT database consisted of 1612 sleep studies from multiple centers, while the UofC database included 981 children, all of them undergoing a complete overnight PSG at the pediatric sleep unit of the corresponding hospitals. Pediatric sleep studies were scored in two databases by trained medical sleep specialists in accordance with the AASM rules (Berry et al., 2012). The AHI was derived from these annotations to determine the severity of OSA in the pediatric subjects. Based on previous studies (Church, 2012; Hornero et al., 2017; Tan et al., 2014), three frequent AHI thresholds (1, 5, and 10 e/h) were used to categorize the severity of OSA into four groups: no-OSA (AHI $<$ 1 e/h), mild OSA ($1 \leq$ AHI $<$ 5 e/h), moderate OSA ($5 \leq$ AHI $<$ 10 e/h), and severe OSA (AHI $\geq$ 10 e/h). Tables 3.1 and 3.2 provide details of both databases, including the total population, age, sex, body mass index (BMI), and the amount of children in each severity group.

3.1 Childhood Adenotonsillectomy Trial (CHAT) database

The CHAT database was a randomized, single-blind controlled trial comprising 1639 sleep studies from children between the ages of 5 and 10 years with symptoms of OSA at the onset of the study (Marcus et al., 2013; Redline et al., 2011). PSGs were conducted at six different pediatric sleep centers located in the United States (Boston Children's Hospital, Boston, MA; Cardinal Glennon Children's Hospital, St. Louis, MI; Cincinnati Children's Medical Center, Cincinnati, OH; Montefiore Medical Center, Bronx, NY; Pennsylvania Children's Hospital, Philadelphia, PA; Rainbow Babies and Children's Hospital, Cleveland, OH) (Marcus et al., 2013; Redline et al., 2011). All children caretakers provided written consent for permission for the research, and assent was obtained from those children over the age of 7 years as part of the research protocol (Marcus et al., 2013). Out of the 1639 original sleep recordings, 1612 contained ECG recordings and fulfilled our studies requirements, thereby being used in the articles of the compendium. The dataset is internally split into three subsets:

- Baseline group, comprised of 451 ECG recordings from children diagnosed with OSA, satisfying the eligibility criteria outlined in Section 3.1.1 to be randomized to early AT or WWSC strategies (Marcus et al., 2013; Redline et al., 2011). The complete set of 451 recordings was used in Martín-Montero et al. (2023). In contrast, and according to the study design, we included only those children who completed the entire trial in Martín-Montero et al. (2022). As a result, a subset of 404 out of the 451 ECG recordings was used in this research.

- Follow-up group, formed by 404 ECG recordings from those children that performed a second nocturnal PSG, completing the entire trial procedure. This subset was used for the purpose of the studies presented in Martín-Montero et al. (2022) and in Martín-Montero et al. (2023).

- Nonrandomized group, consisted of ECG recordings from 757 children who underwent an overnight PSG but did not meet the inclusion criteria (see Section 3.1.1) set in the primary CHAT studies (Marcus et al., 2013; Redline et al., 2011). The complete subset was used in both Martín-Montero et al. (2021a) and Martín-Montero et al. (2021b). However, two subjects were later removed from the subset based on a reviewer's recommendation due to a

low ECG sampling frequency of 50Hz. Hence, in Martín-Montero et al. (2023), 755 children from this group were included.

The standardized procedure outlined in Redline et al. (2011) was followed during nocturnal PSG, which included the collection of ECG recordings through the use of gold cup electrodes or Ag/AgCl patches at sampling frequencies ranging from 200 to 512 Hz. The complete protocol for the CHAT database, with clinical trial identifier NCT00560859, can be found in the supplementary material of Marcus et al. (2013). Sociodemographic and clinical data from the database are presented in Table 3.1.

3.1.1 CHAT database eligibility criteria

As aforementioned, the original CHAT study was developed to include both female and male children at ages ranging from 5.0 to 9.9 years with reported OSA. Accordingly, the inclusion criteria for the original study were as follows (Redline et al., 2011):

- Children aged between 5 and 9.9 years at the screening moment.
- Children with confirmed OSA, defined as those children presenting OAI $\geq$ 1 or AHI $\geq$ 2 after PSG evaluation, and with habitual snoring reported by their legal caretakers.
- Children presenting tonsillar size $\geq$ 1 regarding to a standardized scale from 0 to 4.
- Children considered candidates for surgical AT after otorhinolaryngological evaluation.

Table 3.1. Sociodemographic and clinical data from children of the CHAT database.

	All	no-OSA	mild OSA	moderate OSA	severe OSA
Subjects (n)	1612 (100%)	346 (21.5%)	798 (49.5%)	253 (15.7%)	215 (13.3%)
Age (years)	7.0 [6.0, 8.0]	7.0 [6.0, 8.3]	6.8 [6.0, 8.0]	7.0 [6.0, 8.0]	6.6 [5.8, 8.0]
Males (n)	774 (48.0%)	163 (47.1%)	380 (47.6%)	123 (48.6%)	108 (50.2%)
BMI (kg/m^2)	17.3 [15.6, 21.6]	17.1 [15.5, 19.6]	17.2 [15.6, 20.8]	18.5 [15.4, 23.3]	18.6 [15.9, 23.8]
AHI (e/h)	2.5 [1.2, 6.0]	0.6 [0.4, 0.8]	2.2 [1.6, 3.3]	7.1 [6.0, 8.5]	17.1 [12.7, 26.7]

Data are shown as median [25th percentile, 75th percentile], or number (percentage). AHI = apnea-hypopnea index, BMI: body mass index, OSA: obstructive sleep apnea.

Similarly, the exclusion criteria applied to remove children from the original study were the next (Redline et al., 2011):

- Recurrent tonsillitis, defined as over three events in each of three years, five events repeated two consecutive years, or seven events in a year.
- Children with craniofacial anomalies that could hamper general anesthesia or surgical AT.
- Obstructive respiration during awake that requires immediate AT, evaluated by a physician.
- Children suffering from severe OSA (OAI $>$ 20 or AHI $>$ 30) or significant hypoxemia ($SpO_2 < 90\%$ during $> 2\%$ sleep duration).
- Normal values of AHI (OAI $<$ 1 or AHI $<$ 2).
- Children evidencing significant cardiac arrhythmias after PSG evaluation.
- Children with severe health conditions that could worsen because of the delay in AT application, such as cardiopulmonary affections,epilepsy or diabetes, among others.
- Behavioral or psychiatric disorders that may lead to the initiation of medical treatments during trial duration.
- Genetic, neurological, psychiatric or craniofacial alterations that could disturb airway, behavior or cognition.
- Children under: ADHD or psychotropic drugs, hypnotics, antihypertensives, insulin or hypoglycemic agents, growth hormone, anticonvulsants, corticosteroids, anticoagulants.

3.2 University of Chicago (UofC) database

The private UofC sample, which was also used in the realization of this Doctoral Thesis, comprises 981 ECG recordings of pediatric patients aged between 0 and 13 years. All children included in this database presented OSA symptoms, being derived to the pediatric sleep unit from the Comer Children's Hospital of the University of Chicago Sleep Medicine (Chicago, IL, USA) to undergo an overnight PSG. A digital polysomnography system (Polysmith; Nihon Kohden

America Inc., CA, USA) was used to conduct the PSGs, which enabled the registration of ECG signals at a sampling rate of either 200 Hz or 500 Hz. The study protocol of the UofC sample was approved by the Ethics Committee of the Comer Children's Hospital (numbers of approval: (#11-0268-AM017, #09-115-B-AM031, and #IRB14–1241), and informed consent was obtained from the legal guardians of all participating children. This written consent provided by the legal caretakers conformed to the Declaration of Helsinki. The UofC database was used in Martín-Montero et al. (2021a), as well as in Martín-Montero et al. (2021b). In both studies, these sample serve as the training set of the research works. Sociodemographic and clinical data from the UofC database can be found in Table 3.2.

3.2.1 UofC database eligibility criteria

For children to be referred for polysomnography in the UofC database, they had to meet the following inclusion criteria:

- Children ranging from 0 up to 17 years of age.
- Children under high suspicion of OSA because of showing one or several symptoms among snoring, apneas, arousals, excessive daytime sleepiness, restless sleep, hyperactivity, tonsillar hypertrophy, increase in neck circumference, developmental disorder depression and low self-esteem, enuresis, obesity, attention deficit, behavioral problems, and headache.
- Legal caretakers of the children involved in the study gave their informed consent.

Table 3.2. Sociodemographic and clinical data from children of the UofC database.

	All	no-OSA	mild OSA	moderate OSA	severe OSA
Subjects (*n*)	981 (100%)	173 (17.6%)	401 (40.9%)	178 (18.1%)	229 (23.3%)
Age (years)	6 [3, 9]	7 [4, 10]	6 [4, 9]	5 [2, 8]	4 [2, 8]
Males (*n*)	602 (61.4%)	107 (61.8%)	247 (61.6%)	109 (61.2%)	139 (60.7%)
BMI (kg/m^2)	18.0 [16.1, 21.9]	17.6 [15.6, 21.0]	17.9 [16.1, 21.2]	18.6 [16.5, 24.0]	18.3 [16.1, 23.2]
AHI (e/h)	3.8 [1.5, 9.3]	0.5 [0.1, 0.8]	2.5 [1.7, 3.5]	6.8 [5.8, 8.3]	19.1 [13.9, 31.1]

Data are shown as median [25th percentile, 75th percentile], or number (percentage). AHI = apnea-hypopnea index, BMI: body mass index, OSA: obstructive sleep apnea.

Similarly, the following exclusion criteria were also applied:

- Children suffering from other sleep disorders.
- Children with crania-encephalic abnormalities, epilepsy, neuromuscular disease, genetic syndromes, and congenital cardiac or pulmonary disease.
- Legal caretakers of the children involved in the study gave their informed consent. This consent included the aims of the study, a description about the diagnostic process and a clause stating that the inclusion of the children was voluntary.

Chapter 4

Methods

This chapter provides a comprehensive overview of the methodology used during the Doctoral Thesis. Figure 4.1 depicts the overall methodology steps, beginning with ECG signal acquisition, followed by a pre-processing stage and a segmentation procedure (right branch) to meet the requirements of the different study approaches. Once the ECG has been optimized, the next step involves R-peak detection and HRV extraction. The information related to these initial steps are collected in Section 4.1. Following this pre-processing stage, the feature-engineering approach starts with feature-extraction techniques. In this Doctoral thesis, time, spectral and bispectral analysis techniques have been used, all of them detailed through Section 4.2. Among all the analysis performed, only the bispectral analysis approach included a considerably large number of features, being likely that it included suboptimal feature sets, containing redundant and non-relevant information. For this reason, this is the only branch that contemplated a feature-selection stage, based on FCBF algorithm (see Section 4.3). Afterwards, Section 4.4 provides a comprehensive explanation of the different pattern recognition techniques fed with the different feature sets for pediatric OSA diagnosis or sleep stage classification. Later, Section 4.5 reflects the XAI techniques applied to enhance the understanding of the decisions made by some of the machine-learning algorithms considered. Additionally, one of the studies used spectral HRV characterization with a CMA analysis, described in Section 4.6, to assess whether OSA treatment had any causal effect on changes in HRV. Finally, Section Section 4.7 describes the statistical analysis methods used to evaluate the results.

All the methods detailed in this chapter, but CMA, have been developed us-

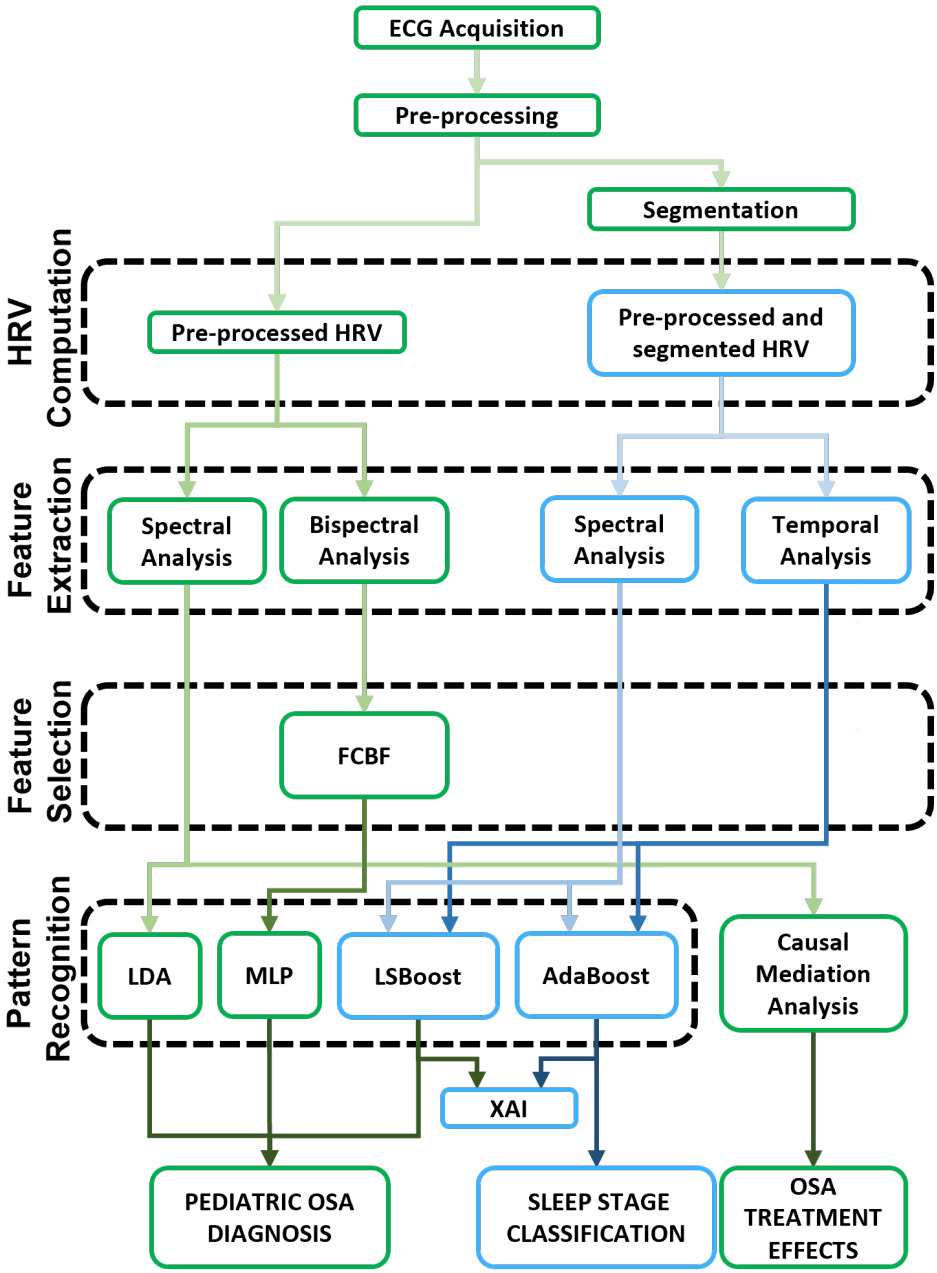

Figure 4.1. Scheme of the overall methodology performed through the Doctoral Thesis. Green boxes indicates that the corresponding step of the protocol was applied per subject, while blue boxes means that the step was applied per segment. AdaBoost: adaptive boosting, ECG: electrocardiogram, FCBF: fast correlation-based filter, HRV: heart rate variability, LDA: linear discriminant analysis, LSBoost: least-squares boosting, MLP: multi-layer perceptron, OSA: obstructive sleep apnea, XAI: explainable artificial intelligence.

ing Matlab software (Mathworks Inc, Natick, MA, USA). The methodology applied in Martín-Montero et al. (2021b) was implemented in Matlab environment version R2018b. The computation of the methods from Martín-Montero et al. (2021a) and Martín-Montero et al. (2023) were implemented in Matlab version R2021b. This version of MATLAB software was also applied to perform the research presented in Martín-Montero et al. (2022), but CMA from this work was performed using R statistical software (v4.0.5; R Core Team 2021) through the R package developed by (Tingley et al., 2014).

4.1 ECG pre-processing and HRV extraction

ECG recordings from the different databases used in this research were acquired during the PSG tests. The electrode was placed on the torso below the right clavicle, while the reference electrode was aligned to the left hip, following the AASM guidelines (Iber et al., 2007). The employment of this single modified ECG lead II using electrode torso placement rather than limb was first proposed by Mason and Likar (1966), as this montage allows to reduce artifacts due to movement (Mason and Likar, 1966). The primary focus of this work was HRV, which is based on the detection of the QRS complex in the ECG. However, raw ECG signals often contain artifacts, such as motion or muscle noise, signal loss, or baseline wander, hindering QRS complex detection (Benitez et al., 2001). Accordingly, ECG signals underwent a pre-processing stage to remove noise, improve QRS detection, and standardize HRV signals.

The pre-processing procedures applied to the ECG signals differed between the three first articles from the compendium and the last one, as they pursued different research objectives. In Martín-Montero et al. (2021a), Martín-Montero et al. (2021b) and Martín-Montero et al. (2022), a whole-night approach was adopted. Thus, the first step was to remove the initial and final 15 minutes, eliminating frequent artifacts during these periods. In addition, to ensure enough and representative data for each child included in the study, each ECG signal was required to contain at least 3 hours of sleep; otherwise, it was excluded from the study, consistent with previous research on pediatric OSA (Barroso-García et al., 2020, 2021; Garde et al., 2014; Xu et al., 2019). Conversely, in Martín-Montero et al. (2023), a segment-level approach was taken, involving the split of the original ECG recordings into 10-minute segments prior to HRV computation (just following R peaks detection, explained below).

The QRS complex detection algorithm used in all the studies was based on the

Hilbert transform. This algorithm, which was previously introduced by Benitez et al. (2001), includes a preliminary two-step filtering stage. First, two median filters were applied to reduce the ripple of QRS complex, and P and T oscillations, followed by a zero-phase low-pass Butterworth filter of 5th order with a cut-off frequency of 0.8 Hz to estimate baseline wander, removing it later from the original signal (Benitez et al., 2001). Figure 4.2 shows an example of an ECG signal before and after pre-processing, highlighting the correction of baseline wander (see Figure 4.2.a). Following this, the QRS complex detection methodology was applied (Benitez et al., 2001). The first differential of the corrected ECG was computed, and then the Hilbert transform was applied to obtain regions with high likelihood of having R peaks inside (Benitez et al., 2001). An adaptive threshold was then used to obtain the final R peak for each region. An example of the R peaks detected has also been depicted in Figure 4.2.b, over the corrected signal.

The final step in obtaining HRV signals is computing the R-R intervals, which can be accomplished once the R peaks have been detected. However, to ensure the accuracy of the results, we also implemented an artifact rejection process aimed at removing physiologically implausible beats detected by the algorithm. To this end, we discarded R peaks that did not satisfy the following criteria (Penzel et al., 2003): (*i*) the duration of R-R intervals had to be between 0.33s and 1.5s, and (*ii*) the maximum variation between an R-R interval and the preceding interval was 0.66s. The outcome of the aforementioned procedures resulted

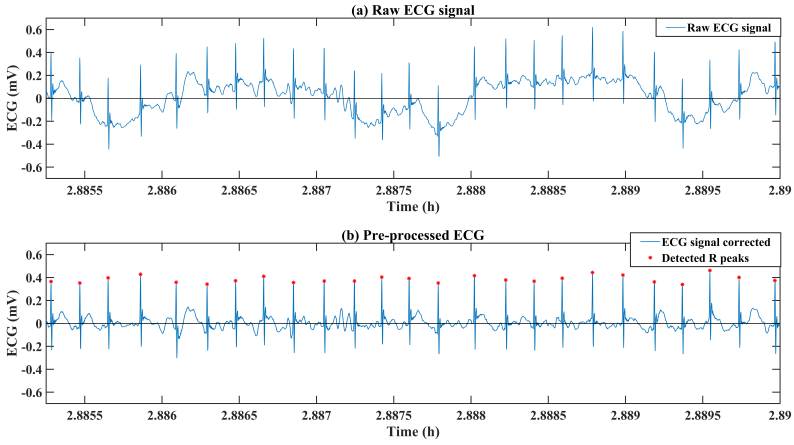

Figure 4.2. Example of the pre-processing applied over an ECG and the R-peaks detection: (a) raw ECG recording, with a clear presence of baseline wander, (b) ECG signal corrected, with the R-peaks located depicted in red over it.

in obtaining normal-to-normal (N-N) intervals, which were subsequently used for HRV estimation. Furthermore, in Martín-Montero et al. (2023), an additional step was taken to minimize the amount of excessively noisy 10-minute segments included. This involved the rejection of segments with less than 500 normal R peaks, equivalent to a HR of 50 beats per minute (bpm). The 50-bpm threshold was chosen as the minimum heart rate for the age range of the children under study (Fleming et al., 2011). Despite the segment-level approach, the criterion of having at least 3 hours of analyzed data for each child was still applied and confirmed after artifact rejection in all the studies included in the compendium.

4.2 Feature extraction

As aforesaid, the Doctoral Thesis aimed at comprehensively characterize overnight HRV in the presence of pediatric OSA. Therefore, the feature-extraction stage from HRV was crucial in achieving this goal. In this regard, different signal processing methods were applied to obtain features from the HRV signal, including both time and frequency domain approaches. Also, novel approaches were proposed in this study to better capture the intrinsic properties of the HRV signal and provide unique information compared to classical methods. All these techniques are described in detail below.

4.2.1 Temporal domain

Conventional HRV metrics in the temporal domain were computed in Martín-Montero et al. (2023) to characterize the activity in the different segments considered across the study. These metrics are based on the assessment of the temporal variations of R-R intervals and offer valuable information on general variability and short-term HRV fluctuations. They are commonly measured to evaluate ANS function and cardiovascular health (Malik et al., 1996). Specifically, three widely used metrics were selected:

- Mean heart rate (mHR), which is the average HR in each segment, measured in bpm. This parameter is related to the inverse of mNN, so when comparing both measures, results should be interpreted in the opposite direction.
- Standard deviation of normal-to-normal interval time series ($SDNN$), which globally reflects the variability in the N-N time series, reflecting global ANS activity (Malik et al., 1996).

- Root mean square of the successive differences ($RMSSD$), which reflects short-term HRV (Malik et al., 1996).

4.2.2 Frequency domain

Frequency-domain measures evaluate HRV at specific frequency ranges associated with particular physiological processes (Acharya et al., 2006). Prior to compute frequency domain analysis, it was necessary to detect and remove all abnormal heartbeats and artifacts, as outlined in Section 4.1. After removing the abnormal beats, the resulting HRV signal was irregularly sampled, and it needed to be resampled to allow for HRV evaluation in the frequency domain (Acharya et al., 2006). In all the publications of the compendium, we resampled the HRV signal to a constant frequency rate of 3.41 Hz, enabling us to compute the corresponding frequency analysis using a window length power of two (1024 samples for 5-min HRV segments), making spectral estimation computationally efficient without adding unnecessary amount of estimated data (Gutiérrez-Tobal et al., 2015; Penzel et al., 2003). The resampling of HRV signal was performed using linear interpolation (Gnauck, 2004; Morelli et al., 2019). The frequency domain was examined in this Doctoral Thesis using two distinct approaches: power spectral density (PSD) and bispectral analysis.

4.2.2.1 Power Spectral Density

One of the main drawbacks of using time domain parameters extracted from HRV is that they are more limited when differentiating between the contribution of the PNS and SNS to the HRV. To address this limitation, PSD analysis is commonly used, which allows for the decomposition of HRV spectral components, requiring at least 5 minutes of recording (Acharya et al., 2006). Therefore, in Martín-Montero et al. (2021b) and Martín-Montero et al. (2022), PSD of the HRV signal was estimated using Welch's approach (Welch, 1967) with a Hamming window of 2^{10} (5 minutes at a sampling frequency of 3.41 Hz), using 50% overlapping and a fast Fourier transform (FFT) of 2^{11} samples. However, in Martín-Montero et al. (2023), due to the study design, we aimed to evaluate differences between the HRV activity of the different 10-minute segment types considered. Consequently, in this work, we estimated PSD using a periodogram method with a Hamming window of 2^{11} samples (the entire 10-minute segment at 3.41 Hz). In all instances, the amplitude values at each frequency of the PSDs were divided by the corresponding total spectral power to normalize the PSD, obtaining normalized PSD

(PSDn).

OSA-specific frequency bands definition

The traditional analysis of HRV spectral features focuses on the well-known VLF, LF, and HF frequency bands. However, the definition of these ranges was originally based mainly on the normal behavior of ANS branches (Acharya et al., 2006; Malik et al., 1996). As one of the main aims of this Doctoral Thesis was to assess ANS alterations in pediatric OSA, alternative frequency bands that specifically reflect such changes were found in Martín-Montero et al. (2021b). Two different approaches were used depending on the frequency range. The first approach focused on the 0-0.15 Hz range, which covered VLF and LF ranges. As commented in the hypothesis section, the rationale for the search of OSA-specific frequency bands between the fixed range from VLF to LF was a previous work in adults suggesting that OSA-related alterations may be affecting HRV activity between both ranges (Gutiérrez-Tobal et al., 2015). The second approach involved an adaptive analysis within the HF range (0.15-0.4 Hz), which is known to be respiratory-modulated (Acharya et al., 2006), thus being strongly influenced by age (Michels et al., 2013). To reduce inter-subject variability, the maximum inside the HF range was first located as a surrogate measure of the individual respiratory peak (Goren et al., 2006), and an adaptive range of 0.15 Hz width was defined centered on this peak (Milagro et al., 2018, 2019). The specific frequency bands were then selected in both approaches based on statistical differences between the PSDns from children of the OSA severity groups established in the training set of this work (UofC database). We performed Mann-Whitney U tests to compare the amplitude values from the PSDns between pairs of severity groups. In the first approach, we compared by frequencies, while in the adaptive approach, we compared by samples. The application of these six statistical tests between OSA severity groups provided us a p-value at each frequency point (0-0.15 Hz range) or at each sample (adaptive range). After performing Bonferroni correction, we established as band of interest those ranges where, at least, two of the tests showed statistically significant differences (fixed at p<0.01). The outcomes of both approaches are shown in Figure 4.3, with the shadow areas corresponding to the frequency bands selected. As can be seen, the first approach led to the identification of two OSA-specific frequency ranges: BW1 (0.001-0.005 Hz) and BW2 (0.028-0.074 Hz). In the adaptive range, three frequency ranges were obtained using the applied methodology. However, after verification (detailed in the next chapter), it was concluded

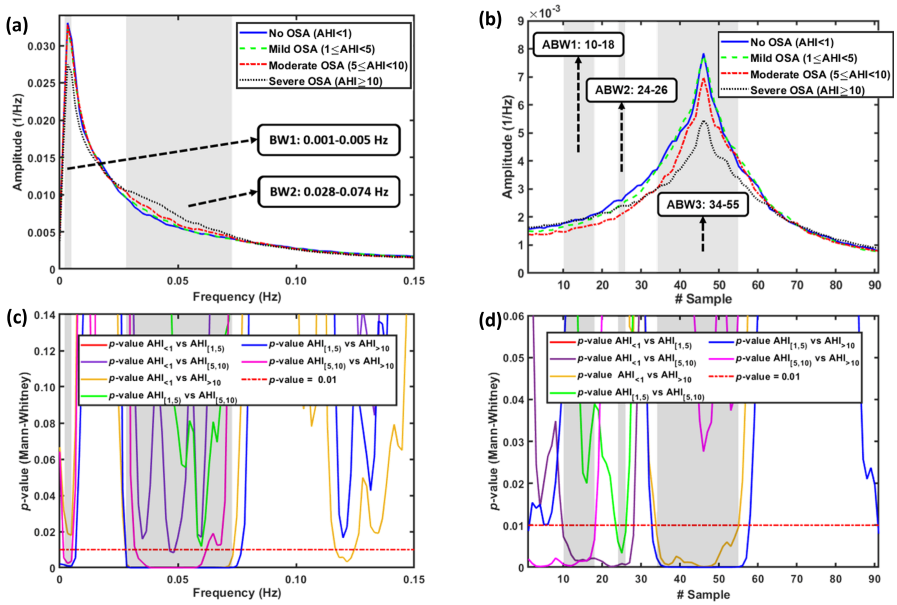

Figure 4.3. Methodology applied to the extraction of pediatric OSA-specific frequency bands in the training set of Martín-Montero et al. (2021b). The resulting frequency bands of interest are represented by the gray-shaded areas. (a) Averaged PSDns for each OSA severity group in the range 0-0.15 Hz. (b) Averaged PSDns for each OSA severity group in the adaptive range. (c) *p*-values from the statistical test for each pair of severity groups in the range 0-0.15 Hz, with the dashed red line highlighting the significance threshold. (d) *p*-values from the statistical test for each pair of severity groups in the adaptive band, with the dashed red line highlighting the significance threshold. This figure has been adapted from Martín-Montero et al. (2021b).

that a range of 0.04 Hz surrounding the respiratory peak (corresponding roughly with ABW3 in Figure 4.3.b) is sufficient for obtaining the OSA-specific frequency band within respiratory range. Thus, ABW3 was renamed BWRes, and was the only band from the adaptive approach used in subsequent analyses, considering ABW1 and ABW2 as spurious.

PSD feature extraction

After defining the OSA-specific frequency bands, the activity in both conventional and novel frequency ranges was characterized by calculating relative power (RP), which is the sum of the amplitude values of the PSDn within a specified frequency range. Therefore, in this Doctoral Thesis, the PSDns of the children were used to extract the following HRV spectral features (Martín-Montero et al., 2023, 2021b, 2022):

- RP within VLF band (RP_{VLF}, 0-0.04 Hz). Its physiological significance remains controversial. It has been suggested that it may be related to slow regulatory mechanisms (Acharya et al., 2006). However, the magnitude and frequency of its oscillations may also be affected by the SNS. This parameter correlates with SDNN (Shaffer and Ginsberg, 2017).

- RP within LF band (RP_{LF}, 0.04-0.15 Hz). It may reflect PNS, SNS and baroreceptor activity (Acharya et al., 2006; Shaffer and Ginsberg, 2017). It is also assumed to correlate with SDNN (Shaffer and Ginsberg, 2017).

- RP within HF band (RP_{HF}, 0.15-0.40 Hz). It reflects HR changes due to respiration, being linked to PNS activity, and correlating with RMSSD (Acharya et al., 2006; Shaffer and Ginsberg, 2017).

- LF/HF ratio (LF/HF). It is usually used to measure the balance between SNS and PNS, where a decrease might reflect either increased PNS or decreased SNS, and vice versa (Acharya et al., 2006). This parameter was computed in Martín-Montero et al. (2021b) and Martín-Montero et al. (2022).

- Normalized power in LF (LF_n). It is measured as the proportion between RP_{LF} and the sum of RP_{LF} and RP_{HF}. As LF/HF, is also used as a measure of sympathovagal balance, but the normalization applied emphasizes the balanced activation of both PNS and SNS (Malik et al., 1996). This feature was computed in Martín-Montero et al. (2023).

- RP within BW1 band (RP_{BW1}, 0.001-0.005 Hz). The HRV activity inside the first out of three novel pediatric OSA-related frequency bands presented in this Doctoral Thesis.

- RP within BW2 band (RP_{BW1}, 0.028-0.074 Hz). The HRV activity inside the second out of three novel pediatric OSA-related frequency bands presented in this Doctoral Thesis.

- RP within BWRes band (RP_{BWRes}, adaptive band of $\pm$ 0.02 Hz around the respiratory peak). The HRV activity inside the last of the novel pediatric OSA-related frequency bands identified.

4.2.2.2 Bispectrum

For decades, power spectrum estimation has been a fundamental tool for analyzing biological signals (Chua et al., 2010). However, this technique only retains in-

formation about the autocorrelation sequence, which is adequate for characterizing Gaussian signals, but does not capture phase relationships among frequencies or phase coupling (Chua et al., 2008, 2010). HRV signals, along with many other biological signals, are nonlinear, non-Gaussian, and non-stationary (Chua et al., 2010). Indeed, the nonlinear dynamics of HRV signals could be exacerbated during sleep (Martín-González et al., 2018; Penzel et al., 2003). Additionally, OSA-induced cardiac changes may also contribute to increased HRV nonlinearity (Atri and Mohebbi, 2015; Guilleminault et al., 1984; Martín-González et al., 2018; Penzel et al., 2003). As a result, PSD analysis may be insufficient to fully characterize changes in HRV series (Chua et al., 2008). In contrast, higher-order spectra (HOS) analysis incorporates both phase and amplitude information and can be used to evaluate diversions in Gaussianity, stationarity, and linearity (Chua et al., 2010). Specifically, the bispectrum is a type of HOS that is derived from the third-order cumulant of a random process, and provides a spectral decomposition of the skewness of the signal across frequency (Chua et al., 2010). Due to the ability of bispectral analysis to overcome PSD limitations, we performed bispectral HRV analysis in Martín-Montero et al. (2021a) to describe alterations due to the disease. The computation of the bispectrum is based on the two-dimensional Fourier transform of the third-order cumulant, and it can be mathematically defined as (Chua et al., 2008, 2010):

$$B(f_1, f_2) = X(f_1) \cdot X(f_2) \cdot X^*(f_1 + f_2), f_1, f_2 = 0, ..., f_N, \tag{4.1}$$

where $X(f)$ is the discrete Fourier transform of a given signal, f_1 and f_2 are the frequency indices of the bispectral axes, and f_N is the signal Nyquist frequency. The same configuration parameters used for spectral analysis were also applied for bispectrum computation, which included a Hamming window length of 2^{10} samples, 50% overlapping, and FFT of 2^{11} samples (Martín-Montero et al., 2021b, 2022). Following bispectrum computation, a normalization step was carried out to constrain all values in the bispectral matrix to a range between 0 and 1, dividing each value of the bispectral matrix by the sum of all its elements as follows (Barroso-García et al., 2021; Chua et al., 2010):

$$B_N(f_1, f_2) = \frac{B(f_1, f_2)}{BP}, f_1, f_2 = 0, ..., f_N, \tag{4.2}$$

where BP represents total bispectral power, computed as the sum of all the matrix values.

The bispectral matrix, by definition, presents symmetric properties, allowing

for the evaluation of a non-redundant triangular area called the region of interest. This triangular area, which is enough for whole bispectrum description, is bounded by $0 \leq f1 \leq f2 \leq f1 + f2 \leq fN$ (Chua et al., 2008, 2010). Previous studies in other research areas have performed bispectral HRV analysis within this non-redundant triangular region of interest (Atri and Mohebbi, 2015; Chua, 2009; Chua et al., 2008). However, as the bispectral matrix axis ranges from 0 to f_N, it is possible to define sub-regions within specific frequency ranges. Thus, we identified six sub-regions from the bispectral matrix, three based on classic frequency ranges (i.e., VLF, LF, and HF), and the remaining three based on the OSA-related frequency ranges, which were obtained in Martín-Montero et al. (2021b) (i.e., BW1, BW2, and BWRes). Of note, all bispectral regions but BWRes are located on the main diagonal of the bispectral matrix, while BWRes was defined by searching for the maximum value within the HF region and fixing its limits around this peak, it may or may not being centered in the main diagonal. Accordingly, those parameters (detailed below) that are based on the diagonal elements of the bispectral regions cannot be computed for this adaptive region, as they lose its physiological meaning.

Once that the bispectral sub-regions were defined, we followed four different feature-extraction approaches to characterize them: bispectral region amplitude, bispectral region entropy, bispectral region moment, and weighted center of bispectrum (WCOB) region features. Additionally, a novel bispectral feature was introduced. Those parameters are described below.

- Features derived from bispectral regions amplitudes:
  - Maximum amplitude (B_{max}) quantified as the highest magnitude value found within each of the evaluated regions (Barroso-García et al., 2021):
    $$B_{max} = max(|B_N(f_1, f_2)|), f1, f2 \in \Omega, \tag{4.3}$$
    where Ω is one out of the six regions under study.
  - Minimum amplitude (B_{min}) computed as the lowest magnitude value within each of the considered regions (Barroso-García et al., 2021):
    $$B_{min} = min(|B_N(f_1, f_2)|), f1, f2 \in \Omega, \tag{4.4}$$
  - Total bispectral power (B_{total}), which enables to measure Gaussianity deviations, and is quantified as the sum of all the values within each

region (Barroso-García et al., 2021):

$$B_{total} = \sum_{f_1,f_2 \in \Omega} |B_N(f_1,f_2)|. \tag{4.5}$$

- Features derived from bispectral regions entropy:
  - Normalized bispectral entropy of first (BE_1), second (BE_2), and third order (BE_3), which measure the irregularity of the bispectral distribution in each region based on Shannon's entropy as follows (Barroso-García et al., 2021; Chua, 2009):

$$BE_i = -\sum_{j \in \Omega} p_j \cdot log(p_j), i = 1,2,3, \tag{4.6}$$

    where p is the amplitude distribution of the values of the region, computed as:

$$p_j = \frac{|B_N(f_1,f_2)|^i}{\sum_{f1,f2 \in \Omega} |B_N(f_1,f_2)|^i}, i = 1,2,3. \tag{4.7}$$

    The bispectral entropies are sensitive to the degree of randomness in a process. Therefore, the changes in HRV irregularity resulting from OSA can be captured by the bispectral regions entropies (Atri and Mohebbi, 2015).
  - Phase entropy (PE), which measures irregularity in the phase of the bispectral region, also based on Shannon's entropy (Barroso-García et al., 2021; Chua et al., 2010). PE is analogous to the bispectral entropies in the sense that higher values indicate greater randomness, while a harmonious, recurrent and predictable process would have a PE of zero (Chua et al., 2008). PE is calculated by applying Shannon's entropy to the normalized distribution of the phase angles in the region as (Barroso-García et al., 2021; Chua et al., 2010):

$$PE = -\sum_{n \in \Omega} p(\Psi_n) \cdot log(p(\Psi_n)) \tag{4.8}$$

    where:

$$p(\Psi_n) = \frac{1}{L} \sum_{f1,f2 \in \Omega} Ind(\phi[B_N(f_1,f_2)] \in \Psi_n)|, \tag{4.9}$$

$$\Psi_n = \{\phi| -\pi + \frac{2\pi n}{N} \leq \phi < -\pi + \frac{2\pi(n+1)}{N}\}. \tag{4.10}$$

In the computation of the phase entropy, the indicator function $Ind(\cdot)$ is used to determine if the phase angle ϕ is within the range of the histogram bins Ψ_n, where Ψ_n represents the n-th bin of the histogram. If ϕ is within the range of Ψ_n, then $Ind(\cdot)$ equals 1, otherwise it equals 0. The value of N corresponds to the number of bins in the histogram.

- Features derived from bispectral regions moments:
  - Sum of logarithmic amplitude values of each bispectral region ($H1$), the sum of logarithmic amplitude values of the diagonal components of each region ($H2$), and the first- and second-order spectral moments of the amplitude values of the diagonal components of each region ($H3$ and $H4$, respectively). These parameters were selected for their ability to capture nonlinear effects within the regions, measured as (Atri and Mohebbi, 2015; Barroso-García et al., 2021):

$$H1 = \sum_{f_1, f_2 \in \Omega} log(|B_N(f_1, f_2)|), \tag{4.11}$$

$$H2 = \sum_{f_k \in \Gamma_{diag}} log(|B_N(f_k, f_k)|), \tag{4.12}$$

$$H3 = \sum_{f_k \in \Gamma_{diag}} k \cdot log(|B_N(f_k, f_k)|), \tag{4.13}$$

$$H4 = \sum_{f_k \in \Gamma_{diag}} (k - H3)^2 \cdot log(|B_N(f_k, f_k)|), \tag{4.14}$$

  where Γ_{diag} represents the diagonal components of the bispectral region being evaluated (unless BWRes region).

- Features derived from bispectral WCOB:
  - WCOB is a feature that reflects the interaction between different frequency components within a bispectral region. This is achieved by assigning a weight to each bispectral element (Ji-Wu Zhang et al., 2000; Wang et al., 2015). WCOB is made up of two vectors, $f1m$ and $f2m$, which are used to measure the coupling focus of the bispectral region and summarize the interaction between frequencies. WCOB parameters are correlated with the bispectral maximum values, where a decrease in the values of $f1m$ and $f2m$ indicates a shift in activity towards lower frequencies (Barroso-García et al., 2021; Wang et al., 2015). The

two vectors extracted from the WCOB are computed as follows (Ji-Wu Zhang et al., 2000; Wang et al., 2015):

$$f1m = \frac{\sum_{f1,f2\in\Omega} f1 \cdot B_N(f_1, f_2)}{\sum_{f1,f2\in\Omega} B_N(f_1, f_2)}, \tag{4.15}$$

$$f2m = \frac{\sum_{f1,f2\in\Omega} f2 \cdot B_N(f_1, f_2)}{\sum_{f1,f2\in\Omega} B_N(f_1, f_2)}, \tag{4.16}$$

- Novel bispectral feature: Relative Power of the diagonal.
  - Relative power of the diagonal (RP_{Diag}), measured as the sum of the bispectral magnitudes of the diagonal elements of the region, after applying a normalization over the whole diagonal from the bispectral matrix. This novel bispectral feature, evaluated for the first time during the development of the present doctoral thesis, quantifies the relative bispectral amplitude within the diagonal of the region with reference to the whole bispectral diagonal ($Diag$), and is computed as:

$$RP_{Diag} = \sum_{f_k\in\Gamma_{diag}} |Diag_N(f_k)|, \tag{4.17}$$

where $Diag_N$ is the normalized diagonal os the bispectral matrix after applying the following normalization:

$$Diag_N(f_k) = \frac{Diag_N(f_k)}{DP}, f_k = 0, ..., f_N \tag{4.18}$$

where DP is the power inside the diagonal of the region, computed as the sum of all the magnitude values of the $Diag$ components. The diagonal elements are a specific case when $f_1 = f_2$; therefore, this feature is intended to measure the phase coupling between the harmonic components of HRV signals, such that $f_3 = f_1 + f_2 = 2f_1$ and $\phi_3 = \phi_1 + \phi_2 = 2\phi_1$. Through the normalization, the sum of all diagonal elements is equal to 1, so evaluating RP_{Diag} over a specific region reflects the proportion of total diagonal bispectral power contained within the region. Thus, the underlying principle of this feature is that the normalization applied allows to reflect shifts in the bispectral power concentration from other regions to those elements within

the diagonal of the region being evaluated.

4.3 Feature selection

After characterizing HRV signals through feature extraction, a set of variables is obtained that may contain irrelevant or redundant information (Guyon and Elisseeff, 2003). Extracting a large number of features does not necessarily improve pattern recognition results, as redundant or irrelevant features can adversely impact predictability and even lead to overfitting in the classifier (Guyon and Elisseeff, 2003). To address this issue, a feature-selection stage can be included to obtain a more concise representation of the information and reduce input features for classifiers. Using feature-selection techniques not only enhances understanding of the feature-extraction process but also improves computation time, efficiency of classifiers, and prediction capability (Guyon and Elisseeff, 2003). In Martín-Montero et al. (2021a), up to eighty features were initially extracted from the different bispectral regions. Consequently, the fast correlation-based filter (FCBF) algorithm was used to obtain two optimal subsets: one for the features extracted in the classic HRV bispectral regions and the other for the OSA-related regions. It is also noteworthy that, as explained below in Section 4.4.3, the least-squares boosting (LSBoost) algorithm used in Martín-Montero et al. (2023) for the pattern recognition stage also performs an internal *de facto* feature-selection process.

4.3.1 Fast Correlation-Based Filter (FCBF)

The application of the FCBF algorithm has been extensively reported in the field of pediatric OSA diagnosis (Barroso-García et al., 2020, 2021; Hornero et al., 2017; Jiménez-García et al., 2020; Vaquerizo-Villar et al., 2018a,b,c). This filter method operates through a two-stage process, where the symmetrical uncertainty (SU) is used to evaluate the relevance and redundancy of features. In the first stage, the algorithm calculates the SU between the feature vector X_i and a dependent variable Y to determine the relevance of the feature as follows (Yu and Liu, 2004):

$$SU(X_i, Y) = 2 \cdot \left(\frac{H(X_i) - H(X_i|Y)}{H(X_i) + H(Y)}\right), \tag{4.19}$$

where $H(X_i)$ and $H(Y)$ denote the Shannon entropies of X_i and Y, respectively, and $H(X_i|Y)$ represents the Shannon entropy of X_i when Y is observed. In pediatric OSA context, the dependent variable Y is usually the AHI. In the second stage, the redundancy of X_i is computed as $SU(X_i, X_j)$, where X_j represents

other feature vector. Based on its relevance and redundancy, FCBF removes X_j if $SU(X_i, Y) \geq SU(X_j, Y)$, and $SU(X_i, X_j) \geq SU(X_j, Y)$ (Yu and Liu, 2004). Consequently, the FCBF algorithm produces an optimum subset of non-redundant and relevant features.

In this Doctoral Thesis, the FCBF algorithm was applied in Martín-Montero et al. (2021a) to select features subsets that were relevant and non-redundant using 1000 bootstrap-derived replicates from the corresponding training dataset. The features selected more than 500 times were included in the optimal subsets.

4.4 Pattern recognition

The pattern recognition stage involved machine-learning techniques to extract complex patterns from the sets of features and make predictions based on this information (Alpaydin, 2014; Bishop, 2006). As the datasets considered is labeled, several supervised learning approaches were used for different purposes. Binary classification tasks were conducted using linear discriminant analysis (LDA) and multi-layer perceptron (MLP) to diagnose pediatric OSA based on different AHI thresholds. An adaptive boosting (AdaBoost) ensemble-learning method was used for multiclass classification to distinguish between wake, NREM, and REM sleep stages. Finally, in a regression task, a LSBoost model was developed to estimate the AHI of the children from HRV signals.

4.4.1 Binary classification

4.4.1.1 Linear Discriminant Analysis (LDA)

The LDA classifier is a statistical approach used to distinguish between two or more classes of events or objects by identifying a linear combination of features that characterizes or separates them (Bishop, 2006). The objective of the LDA algorithm is to decrease the dimensionality of the feature space by projecting the original data onto a new feature subspace while retaining as much class-discriminatory information as possible (Bishop, 2006; Friedman, 1989). To achieve this, a set of linear discriminant functions is determined, which maximizes the proportion of between-class variance to within-class variance (Bishop, 2006; Friedman, 1989). When using LDA, the discriminant score for each class is calculated as follows (Friedman, 1989; Marcos et al., 2009):

$$y_j(x) = \mu_j^T \sum^{-1} x - \frac{1}{2} \mu_j^T \sum^{-1} \mu_j + lnP(C_j), \tag{4.20}$$

where μ_j is the mean vector for class C_j, $P(C_j)$ is the prior probability of C_j (i.e., the ratio of input feature vectors x_j belonging to C_j), and $\sum$ is the covariance matrix. The resulting discriminant scores are then used to assign new instances into one of the pre-defined classes based on the decision boundaries established (Bishop, 2006; Friedman, 1989).

In Martín-Montero et al. (2021b), two LDA models were trained for each of the binary severity AHI thresholds (1, 5 and 10 e/h). In each case, one of the models was developed based on the RPs in the classic HRV frequency bands, while the other model was developed based on the RPs in the OSA-specific frequency bands.

4.4.1.2 Multi-layer perceptron (MLP) neural network

Artificial neural networks (ANNs) were developed to simulate biological systems information processing by mathematical models (Bishop, 2006). ANNs are composed of basic interconnected processing units, also called neurons, which can learn complex patterns from the input data. Among the various types of ANNs, the multi-layer perceptron (MLP) is the most commonly used in the context of pediatric OSA (Gutiérrez-Tobal et al., 2022). MLP is a feed-forward ANN typically consisting of an input, a hidden, and an output layer. The neurons in MLP are known as perceptrons and each perceptron is connected to every perceptron in the subsequent layer, assigning a weight to each link. Perceptrons are defined by an activation function that performs a nonlinear conversion of the data (Bishop, 2006). In the input layer, the number of perceptrons corresponds to the number of features in the input data (N). The number of neurons in the hidden layer (N_H) is a parameter that needs to be optimized. Each hidden layer neuron receives a linear combination of the outputs from the input layer neurons and provides a non-linear function of it as output (Bishop, 2006). In the output layer, the number of perceptrons is adapted based on the task of the MLP. For binary classification, as in Martín-Montero et al. (2021a), a single perceptron is used, which receives the hidden layer outputs and provide the posterior likelihood of belonging to each severity group under study. The values of the output units are then calculated as (Bishop, 2006):

$$y_k = g_o\{\sum_{j=1}^{N_H} w_{jk} g_h\{\sum_{i=1}^{N} w_{ij} x_i + b_j\} + b_k\}, \tag{4.21}$$

where x_i is the feature input vector, $g_h(\cdot)$ and $g_o(\cdot)$ represent the activation functions of the perceptrons in the hidden and output layers, respectively, w_{ik} indicates the weight associated with the connection between the i-th input feature and the j-th hidden layer perceptron, and w_{jk} represents the weight associated with the link that connects the j-th hidden layer perceptron to the output perceptron y_k. Additionally, b_j and b_k are the biases of the hidden and the output layers, respectively. Furthermore, we introduced a regularization parameter (λ) in the tunning of the MLP weights to minimize overfitting (Bishop, 2006), that was randomly initialized and optimized later.

In Martín-Montero et al. (2021a), the two optimum subsets obtained from the FCBF algorithm were utilized to train MLP models for binary classification across each severity thresholds, with up to six MLP neural networks being optimized. The optimization of the hidden layer design parameters (i.e, N_H and λ) was carried out by performing 1000 bootstrap replicates from the training set. The combination of N_H and λ values chosen for each model was the one that yielded the highest Cohen's kappa coefficient (k) for each specific case.

4.4.2 Multiclass classification: Adaptive Boosting

The concept of ensemble learning refers to the merging of multiple models to improve the overall performance beyond what each individual model can achieve (Freund and Schapire, 1997; Witten et al., 2011). Boosting is one of the most profitable strategies for implementing ensemble-learning algorithms, as it can significantly enhance the prediction capability on new data (Bishop, 2006; Witten et al., 2011). Boosting algorithms combine models iteratively to complement each other by fitting weighted votes assigned to the same type of base classifiers at each iteration (Bishop, 2006; Witten et al., 2011). AdaBoost is an ensemble-learning boosting algorithm where weak classifiers are commonly used as base classifiers. Weak classifiers are preferred because complex classifiers may cause overfitting, thereby reducing the generalizability of the algorithm (Bishop, 2006; Freund and Schapire, 1997). Following a similar approach than in previous research in OSA context (Gutierrez-Tobal et al., 2019; Jiménez-García et al., 2020), in Martín-Montero et al. (2023), an AdaBoost algorithm with LDA as weak classifier was implemented.

The AdaBoost algorithm functions by assigning a weight (w_i^m) at each iteration (m) to each observation (x_i), i.e., the feature vector of each instance in the training group. The model for that iteration is then trained with the associated

weighted observation, and its performance is evaluated, with an error (ϵ_m) being computed. This error is used to assign the weighted vote (α_m) of the classifier being trained in that iteration, with a smaller error resulting in a higher contribution to the final prediction (Witten et al., 2011). At the end of each iteration, the weights are updated, giving more importance to those observations that were misclassified, thus increasing the chance of proper classification in the next iteration (Freund and Schapire, 1997; Witten et al., 2011).

Various versions of the AdaBoost algorithm exist, each with a specific purpose. In Martín-Montero et al. (2023), we utilized AdaBoost.M2, the AdaBoost algorithm specifically designed for multiclass classification, for the purpose of classification of sleep stages. In this version of AdaBoost, the ϵ_m value is computed using the next equation (Freund and Schapire, 1997):

$$\epsilon_m = \frac{1}{2}\sum_{k=1}^{N}\sum_{l\neq l_{true}} w_{i,l}^{m}(1 - c_m(x_i, l_{true}) + c_m(x_i, l)), \tag{4.22}$$

where l refers to a categorical variable containing the different classes, l_{true} indicates the original class assigned to x_i, and c_m represents the LDA confidence prediction for a feature vector and a given class. The predicted class is determined as the class that receives the largest sum of weighted votes from the LDA classifiers, with the weight of the predictions calculated as follows (Freund and Schapire, 1997):

$$\alpha_m = ln(\beta_m), \tag{4.23}$$

where the calculation of β_m is determined by ϵ_m as $\frac{(1-\epsilon_m)}{\epsilon_m}$ (Freund and Schapire, 1997). To minimize overfitting, a learning rate (ν) can also be introduced in the computation of β_m. Thus, for the multiclass classification task addressed in this Doctoral Thesis, there were certain parameters for the design of the AdaBoost algorithm that required tuning. These parameters included the number of LDA classifiers used (N_{AB}) and the regularization parameter ν. The optimization process was carried out by identifying the (N_{AB},ν) pair that maximized the multiclass k on the validation set. To do this, ν was increased from 0.1 to 1 in increments of 0.1, while N_{AB} was modified from 1 to 10,000 in increments of multiples of 10, starting with 1 and increasing up to 9 in increments of 1, then from 10 to 100 in increments of 10, and so on.

4.4.3 Regression: Least-Squares Boosting

In the present Doctoral Thesis, the LSBoost algorithm was used to perform a regression task. LSBoost, like AdaBoost, is a type of ensemble-learning boosting algorithm that sequentially combines weak base learners, with the aim of minimizing the sum of squared errors between the predicted and actual values of the regression problem (Bühlmann and Hothorn, 2007). In Martín-Montero et al. (2023), we used as weak learners decision stumps, which are decision trees with one parent and two child nodes, thus conducting a *de facto* selection process at each iteration (Gutiérrez-Tobal et al., 2021). In particular, we used LSBoost to estimate the number of apneic events present in each HRV segment considered. The generic workflow of LSBoost algorithm can be summarized as follows (Bühlmann and Hothorn, 2007; Bühlmann and Yu, 2003):

1. Considering $f^m(x)$ as the predicted output, the number of learners (m) is initialize to 0, computing the corresponding $f^0(x)$.

2. m is increased by 1, computing the residuals as: $U_i = y_i - f^{m-1}(x_i)$ for $i = 1, 2, ...N$, where y_i is the target variable, x_i is the corresponding instance, and N is the number of instances in the training group.

3. The residual vector is fitted using least squares loss function, the base learner h, and the predictor for every instance x_i : $(\lambda_m, a_m) = argmin_{\lambda,a} \sum_{i=1}^{m} [U_i^m - \lambda h(x_i; a)]^2$, being a the parameters of h, and λ a regularization factor.

4. $f^m(x)$ is updated as: $f^m(x) = f^{m-1}(x) + \lambda_m h(x; a_m)$.

5. Steps 2 to 4 are sequentially repeated until $m = N_{LSB}$, where N_{LSB} is the number of learners considered.

Thus, in our study, y_i was the original number of apneic event scored to each segment, x_i was the feature vector of each segment, and N was the number of segments included in the training set. Following the prediction of apneic events presence in every HRV segment, the estimated AHI for each subject can be measured as the proportion between the estimation and the total recording time. Nevertheless, the original AHI is constructed based on the total sleep time scored by the medical experts. Thus, the use of the total recording time causes an underestimation in the AHI reconstruction (Deviaene et al., 2019). To solve this limitation, a linear regression model can be fitted using training data, helping to deal with this

bias introduced by the use of total recording time, and improving AHI estimation (Deviaene et al., 2019; Vaquerizo-Villar et al., 2021).

Similar to the optimization of hyperparameters for the AdaBoost model, the LSBoost algorithm also required the tuning of two hyperparameters: λ and N_{LSB}. The optimization process for LSBoost followed the same procedure as for AdaBoost, which involved varying the values of λ and N_{LSB} over the same range as those used for ν and N_{AB}, respectively.

4.5 Explainable Artificial Intelligence (XAI)

As mentioned in Section 1.2, the concept of XAI has become increasingly important, particularly in health research. XAI enhances the interpretability of AI predictions, reducing the perception of AI models as 'black boxes' and ensuring that they work safely and ethically (Adadi and Berrada, 2018). Therefore, in the final research included in the compendium of publications, we aimed to incorporate some XAI techniques that would help us to better understand the decision-making processes of the AdaBoost and LSBoost algorithms, focused in the analysis of feature importance.

The XAI techniques were chosen based on the selected base learners for each case. As previously mentioned, in the case of AdaBoost, we used LDA as base learners. Accordingly, to improve the interpretability of the predictions of the algorithm, we utilized local interpretable model-agnostic explanation (LIME), as an XAI technique that can be applied to any type of model regardless of its internal functioning (Ribeiro et al., 2016). LIME is one of the most widespread XAI techniques due to its simplicity and its ability to create interpretable explanations of machine-learning approaches predictions. It operates by constructing an interpretable model of the classifier around the prediction of each instance, which is known as local explanation. This local model is fitted around altered samples close to the target instance, with these slight modifications created through the addition of noise or direct manipulation of the instance's features (Ribeiro et al., 2016). Therefore, we applied LIME to generate local models around each instance from the test set, resulting in a weighted coefficient W_{ij} for each instance and feature (Ribeiro et al., 2016). Using this coefficient, we computed the total importance of the j-th feature as the square root of the sum of the absolute values of all the coefficients obtained for this feature in the n instances from the test set (Ribeiro et al., 2016):

$$I_j = \sqrt{\sum_{i=1}^{n} |W_{ij}|}. \tag{4.24}$$

After computing I_j for all features, the total importance of each parameter can be scaled as a proportion of its contribution to the sum of the total importance of all the features considered. This provides the relative importance for each feature to the AdaBoost model as a percentage of contribution, allowing us to measure the role of each feature in the sleep stage classification task.

In contrast to AdaBoost, the selection of an appropriate XAI strategy for the LSBoost algorithm was different due to the use of decision stumps as base learners. Each stump tree $h(x; a_m)$ in the LSBoost algorithm depends on a single feature. Consequently, as explained in Section 4.4.3, LSBoost conducts an internal feature-selection process for each decision (Bühlmann and Hothorn, 2007; Bühlmann and Yu, 2003). Hence, the importance of each feature in the LSBoost algorithm can be determined using the mean squared error (MSE), as follows (Friedman and Meulman, 2003; Gutiérrez-Tobal et al., 2021):

$$\hat{I}_j^2 = \frac{1}{N_{LSB}} \sum_{m=1}^{N_{LSB}} MSE_m(x_j) w_m - (MSE_m^p(x_j) w_m^p + MSE_m^r(x_j) w_m^r), \tag{4.25}$$

where w_m represents the weight of the parent node likelihood, MSE_m corresponds to the MSE of the m-th tree linked to the j-th feature, and p and r are the parameters linked to the children nodes. Once the importance of each feature to the LSBoost model has been calculated, it can be scaled as a proportion of the total contribution (Elith et al., 2008), measuring the relative importance of each feature in the pediatric OSA diagnosis task.

4.6 Causal Mediation Analysis (CMA)

CMA is a statistical methodology applied to evaluate the causal pathways by which an action, typically a treatment, affects an outcome (Imai et al., 2010a,b). The purpose of CMA is to measure the impact of intermediate variables (named mediators) and the extent of their influence on the outcome due to the treatment, as well as the remaining effect of the treatment that is not linked to the mediator being evaluated. The impact of the treatment on the outcome through the mediator is known as the averaged causal mediation effect (ACME), while the

treatment effect that is not linked to the mediator is known as the averaged direct effect (ADE) (Imai et al., 2010a,b; Tingley et al., 2014). Figure 4.4 presents the general framework of CMA, which decomposes the total treatment effect into the ACME and ADE. The reason for using CMA instead of estimating general causal effects is that while a causal effect indicates whether the treatment has a causal impact, it cannot explain how or why this influence occurs, which CMA can do (Imai et al., 2010a). This distinction is also depicted in Figure 4.4, which explains the difference between general causal effects and CMA. When ACME and ADE act in opposite directions, the values of ADE may mask ACME effects, leading

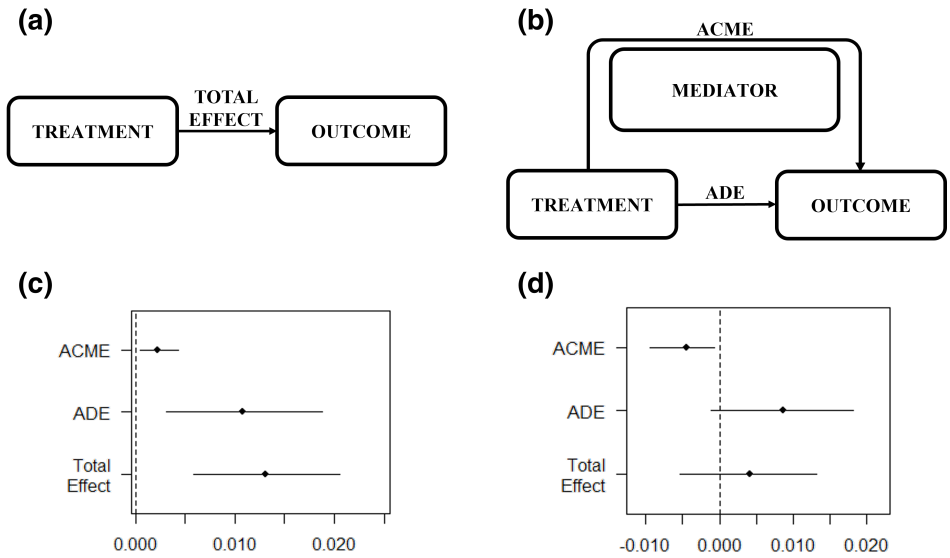

Figure 4.4. The diagram illustrates the framework of the causal mediation analysis (CMA). (a) Overall estimation of the causal total effect. (b) Workflow of the CMA, aimed to identify the causal pathways or mediators through which the treatment influences variations in an outcome. The average causal mediation effect (ACME) reflects the alteration in the outcome as a result of the change in the mediator caused by the intervention, while the average direct effect (ADE) reflects the variation in the outcome that occurs for any other reason but the mediator. In this Doctoral Thesis, the effects were averaged over both trial arms, after verifying that the no-interaction can be assumed. ACME and ADE jointly form the total effect. Two examples of the effects estimated for an outcome and a mediator are represented in (c) and (d), along with their 95th percentile confidence intervals. In (c), all the effects are in the same direction, and none of them include the value of 0, which indicates that these effects are statistically significant different from 0. In contrast, in (d), the ACME shows an opposite direction to the ADE and Total Effect, and only the ACME does not include 0, indicating that is the only effect statistically significant different from 0, which would be masked if just the total effect is observed.

to the conclusion that there are not causal effects if the total effect is analyzed isolated, thus highlighting the usefulness of CMA (Imai et al., 2010a).

The current Doctoral Thesis used the mediation analysis approach originally proposed by Baron and Kenny (Baron and Kenny, 1986) within the framework of linear structural equation modeling, which was later expanded into a general approach by Imai et al. (2010a) to enable the measurement of causal relations and the inclusion of discrete variables. To facilitate clarity and consistency with the original work, the statistical framework of this analysis is explained here using the same notation. Specifically, T_i represents the binary treatment variable and M_i denotes the observed mediator, which is expected to be influenced by the treatment arm. Thus, the mediator variable can take two potential values, $M_i(1)$ or $M_i(0)$, but only one value is observed for each subject, represented as $M_i(T_i)$. In the potential outcomes framework, the outcome variable is influenced by both the treatment and mediator, being denoted by $Y_i(t, m)$. Therefore, the indirect effect for each subject is then defined accordingly (Imai et al., 2010a):

$$\delta_i(t) = Y_i(t, M_i(1)) - Y_i(t, M_i(0)), \tag{4.26}$$

for $t = 0, 1$. According to the original authors (Imai et al., 2010a), the significance of the previous equation can be attributed to the counterfactual inquiry concerning the alteration in the outcome that would result from a change in the mediator value. Specifically, this refers to the shift from the mediator value under one treatment condition (e.g., $M_i(0)$) to that of the other treatment condition ($M_i(1)$), while keeping the treatment value constant at t. If the mediator is not impacted by the treatment, $M_i(1)$ will be equal to $M_i(0)$, leading to an indirect effect of 0 (Imai et al., 2010a). Likewise, the direct effect can be defined as follows (Imai et al., 2010a):

$$\zeta_i(t) = Y_i(1, M_i(t)) - Y_i(0, M_i(t)). \tag{4.27}$$

Therefore, the total effect is computed as the union of both effects (Imai et al., 2010a):

$$\tau_i = Y_i(1, M_i(1)) - Y_i(0, M_i(0)) = \frac{1}{2} \sum_{t=0}^{1} \delta_i(t) + \zeta_i(t). \tag{4.28}$$

Assuming what is known as the no-interaction assumption, it is supposed that the direct and indirect effects are not dependent on the treatment arm. Therefore, for each child, $\delta_i = \delta_i(1) = \delta_i(0)$, $\zeta_i = \zeta_i(1) = \zeta_i(0)$ and $\tau_i = \delta_i + \zeta_i$. Conse-

quently, we can compute the ACME, ADE and total effects ($\bar{\tau}$) across all subjects by using the following formulae (Imai et al., 2010a):

$$ACME(t) = \bar{\delta}(t) = E(Y_i(t, M_i(1)) - Y_i(t, M_i(0))), \tag{4.29}$$

$$ADE(t) = \bar{\zeta}(t) = E(Y_i(1, M_i(t)) - Y_i(0, M_i(t))), \tag{4.30}$$

and

$$\bar{\tau}(t) = E(Y_i(1, M_i(1)) - Y_i(0, M_i(0))) = \frac{1}{2}\sum_{t=0}^{1} \bar{\delta}(t) + \bar{\zeta}(t). \tag{4.31}$$

Again, under no-interaction assumption, we obtain $ACME = \bar{\delta} = \bar{\delta}(1) = \bar{\delta}(0)$ and $ADE = \bar{\zeta} = \bar{\zeta}(1) = \bar{\zeta}(0)$. The joint of ACME and ADE then leads to obtaining the average total effect ($\bar{\tau} = \bar{\delta} + \bar{\zeta}$) (Imai et al., 2010a).

A crucial assumption for valid interpretations of CMA is known as sequential ignorability, which assumes that the treatment assignment is independent of the mediators and outcomes (satisfied in randomized controlled trials, such as the CHAT study), and that all baseline confounding variables have been accounted for. To partially deal with confounding variables, the software implementation developed by Tingley et al. (2014) allows the inclusion of baseline covariates in the CMA computation. However, there may be remaining unobserved confounding variables, being it difficult to totally fulfill with this assumption (Imai et al., 2010a). To address this, Imai et al. (2010a,b) also proposed a sensitivity analysis methodology that can evaluate the robustness of the conclusion extracted in case that the sequential ignorability assumption is violated, and they strongly recommended performing this analysis when using CMA (Imai et al., 2010a,b).

Therefore, in Martín-Montero et al. (2022), we utilized CMA to assess whether modifications in HRV parameters were causally linked to the treatment, exploiting the design of the CHAT study, which provided an appropriate framework for CMA evaluation. In that work, the intervention was represented by any of the two randomized treatment arms assigned to the children (early AT or WWSC). We considered changes in RP for each of the six frequency bands, including both classical and OSA-specific frequency bands, as outcomes. Those outcomes were defined as the difference in RP between baseline and follow-up ($\Delta RP_{band} = RP_{band\_follow-up} - RP_{band\_baseline}$). We included five continuous and two binary mediators, all of which were related to the severity of the disease to varying degrees. The continuous mediators comprised changes in AHI, obstructive AHI (OAHI), oxygen desaturation index (ODI), the minimum satura-

tion (minsat) during night, and the total arousal index (TAI). To assess disease resolution, we incorporated two binary mediators: OSA resolution, which was equal to 1 for those subjects where follow-up OAHI $\leq$ 2 e/h and OAI $\leq$ 1 e/h were present, and 0 otherwise, and OSA_{O+C} resolution, which took into account the presence of central apneas, and was defined as follow-up AHI $\leq$ 2 e/h and apnea index (AI) $\leq$ 1 e/h.

The R package developed by Tingley et al. (2014) was utilized to conduct all the computations. Initially, we assessed whether the effects on HRV were reliant on the treatment arms, obtaining that it did not occur. Thus, we assumed no-interaction, providing the averaged effect regardless treatment arm (Imai et al., 2010a,b). To control for as many baseline confounders as possible, we incorporated well-known confounders such as age, race, sex, BMI z-score, average HR, and tonsil size in the analysis. Finally, we tested the potential infringement of the sequential ignorability assumption by carrying out a sensitivity analysis using this software (Imai et al., 2010a,b).

4.7 Statistical analysis

The present Doctoral Thesis used several techniques to describe and assess the results obtained from signal processing approaches. These methods include statistical tests, performance metrics for pediatric OSA diagnosis or sleep stage classification, measures of agreement, and validation approaches.

4.7.1 Statistical tests

Statistical hypothesis tests are techniques used in statistical inference to determine if it is possible to extract conclusions about a population based on a sample of data (Jobson, 2012). In this research, we initially evaluated the normality and homoscedasticity of the distributions under study through the use of the Lilliefors (Lilliefors, 1967), and Levene (Levene, 1960) tests, respectively. We obtained that not all the variables passed normality and homoscedasticity tests. Accordingly, non-parametric tests were performed through the realization of this Doctoral Thesis to identify statistically significant differences among the groups being studied (Jobson, 2012). Thus, in the search of the pediatric OSA-related frequency bands in Martín-Montero et al. (2021b), Mann-Whitney U test was used to evaluate the differences between the averaged PSDns of each two OSA severity groups. Regarding the assessment of differences in the HRV parameters,

the Mann-Whitney U test was computed when comparing two different severity groups, the Wilcoxon signed rank test for intra-group comparisons, and the Kruskal-Wallis test to assess differences between more than two groups. The significance level was adjusted for the number of subjects, and different p-values were used across studies. In Martín-Montero et al. (2021a,b), where there were larger sample size (1738 subjects), the significance level was set at $p < 0.01$. In Martín-Montero et al. (2023, 2022), the threshold of significance was set at $p < 0.05$. However, in Martín-Montero et al. (2023), although the database was slightly shorter than in the two first studies of the compendium, a segment-level approach was followed, thus increasing considerably the sample size used to assess differences. When it occurs, the p-value would not be enough for a comprehensive evaluation of the differences (Sullivan and Feinn, 2012). Accordingly, our conclusions were based on the effect size computed using the non-parametric Cohen's d measure, with differences interpreted as small ($0.2 \leq d < 0.5$), medium ($0.5 \leq d < 0.8$), or large ($d \geq 0.8$) effect size (Cohen, 1998; Sullivan and Feinn, 2012). Bonferroni correction was applied in Martín-Montero et al. (2023, 2021a,b), while false discovery rate was used in Martín-Montero et al. (2022) when multiple comparisons were present.

To obtain complementary information and aid in the physiological interpretation of the results, additional statistical techniques were utilized. Specifically, Spearman's partial correlation coefficient (ρ_S) was chosen to evaluate the relationship between the HRV features and polysomnographic indices such as AHI, OAHI, OAI, ODI, minsat, TAI, the number of awakenings during the night (#Awakenings), the wake after sleep onset (WASO) time, and the percentage of sleep spent in different stages (%N1, %N2, %N3, and %REM). Using ρ_S also allows for the control of possible confounding factors, such as age or the baseline confounders included in the CMA. Furthermore, boxplot representations were used as a graphical approach to illustrate the distribution of the HRV features computed through the different studies. When using boxplots, the box of the plot represents 50% of the data, with the lower portion of the box indicating the first quartile and the upper portion indicating the third quartile. Thus, boxplots are frequently used in data analysis, as they offer insights into the range and skewness of the features being represented.

4.7.2 Diagnostic performance metrics

Throughout the realization of this Doctoral Thesis, the diagnostic capabilities of individual HRV features, as well as the diagnostic yield of several models obtained from the different feature-engineering approaches were evaluated. Diagnostic performance was assessed using statistical measures that take into account the number of successfully or erroneously classified subjects. In a binary classification task, a confusion matrix can be constructed based on the following components (Flemons and Littner, 2003):

- True positives (TP), measured as the sum of positive subjects that have been successfully classified as positive (e.g., children with OSA classified as OSA positives).
- True negatives (TN), measured as the sum of negative subjects that have been successfully classified as negative (e.g., children without OSA classified as OSA negatives).
- False positives (FP), measured as the sum of negative subjects that have been erroneously classified as positive (e.g., children without OSA classified as OSA positives).
- False negatives (FN), measured as the sum of positive subjects that have been erroneously classified as negatives (e.g., children with OSA classified as OSA negatives).

Hence, by obtaining those values, the subsequent statistics parameters were selected as the main diagnostic performance metrics used to evaluate the diagnostic efficiency in the binary classification tasks using the three AHI thresholds aforementioned (1, 5 and 10 e/h) (Flemons and Littner, 2003):

- Sensitivity (Se). Percentage of subjects with the disease (AHI $\geq$ threshold) successfully classified:

$$Se = \frac{TP}{TP + FN} \cdot 100 \tag{4.32}$$

- Specificity (Sp). Percentage of subjects without the disease (AHI $<$ threshold) successfully classified:

$$Sp = \frac{TN}{TN + FP} \cdot 100 \tag{4.33}$$

- Accuracy (Acc). Percentage of subjects successfully classified. For a binary classification problem, it is computed as follows:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \cdot 100 \tag{4.34}$$

- Area under the Receiver-Operating Characteristics curve (AUC). The effectiveness of diagnostic tests can be evaluated using the receiver operating characteristic (ROC) curve, which is a 2D parametric curve representing the trade-off between Se and $1 - Sp$. The ROC curve is constructed by varying the decision threshold of the test, and the upper left corner of the curve represents an ideal classification result, where $Se = 1$ and $1 - Sp = 0$ (Fawcett, 2006; Zweig and Campbell, 1993). AUC is a measure used to summarize the ROC curve, computed as the area comprised between the curve and the abscissa axis. Accordingly, AUC values close to 1 indicate high diagnostic performance (Fawcett, 2006; Zweig and Campbell, 1993). In this doctoral Thesis, AUC has been measured as a diagnostic performance metric for the machine-learning models developed, and a ROC curve analysis has been performed to evaluate the individual diagnostic performance of some of the HRV features. This was done by selecting as optimum threshold the feature values that provides a ROC curve closer to the ideal classification point (in the training set) (Fawcett, 2006; Zweig and Campbell, 1993).

Along with those parameters, in Martín-Montero et al. (2023) we decided to include some additional metrics to complement the evaluation of the clinical applicability of the LSBoost and AdaBoost models:

- Multiclass accuracy Acc_k. The binary Acc definition can also be extended for a multiclass classification problem as the sum of instances rightly classified in each class, computed based on the main diagonal of the confusion matrix as:

$$Acc_K = \frac{\sum_{i=1}^{K} n_{i,i}}{N} \cdot 100 \tag{4.35}$$

  where K is the number of classes, and N is the sample size. For the LSBoost model, we computed the multiclass accuracy for four classes (Acc_4), where each class was one of the four OSA severity levels defined. Regarding the AdaBoost model, we computed the three-class accuracy (Acc_3), with each class being one out of the three sleep stages considered.

- Positive predictive value (PPV), also known as precision. Percentage of

subjects successfully classified among all the subjects that has been classified as positives:

$$PPV = \frac{TP}{TP + FP} \cdot 100 \tag{4.36}$$

- Negative predictive value (NPV). Percentage of subjects successfully classified among all the subjects that has been classified as negatives:

$$NPV = \frac{TN}{TN + FN} \cdot 100 \tag{4.37}$$

- Positive likelihood ratio (LR^+). Proportion of subjects from the positive group successfully classified in relation to the proportion of subjects from the negative group erroneously classified:

$$LR^+ = \frac{Se}{1 - Sp} \tag{4.38}$$

- Negative likelihood ratio (LR^-). Proportion of subjects from the positive group erroneously classified in relation to the proportion of subjects from the negative group successfully classified:

$$LR^- = \frac{1 - Se}{Sp} \tag{4.39}$$

- $F1 - score$. This diagnostic performance metric, especially useful when there exist imbalance between classes, is computed as the harmonic mean of precision (i.e., PPV), and recall (i.e., Se):

$$F1 - score = 2 \cdot \frac{PPV \cdot Se}{PPV + Se} = \frac{2TP}{2TP + FP + FN} \tag{4.40}$$

4.7.3 Measures of agreement

Agreement measures are computed to determine the level of consistency between the predicted and true labels of a sample set. These measures are useful in evaluating the precision of classification models and comparing the performance of different classifiers. Thus, besides to the diagnostic performance metrics, an additional agreement measure has been included in this Doctoral Thesis:

- Cohen's kappa coefficient (k). This statistic is a metric computed to evaluate the agreement between the classes observed and the predicted ones, while

taking into account the possibility of agreement occurring by chance. This measure is computed as (Cohen, 1960):

$$k = \frac{p_o - p_e}{1 - p_e} \tag{4.41}$$

where p_o represents the observed concordance between predicted and actual classes, and p_e represents the likelihood of agreement occurring by chance. The range of k values is between -1 and 1, where a k value of -1 indicates total absence of agreement, k=1 means full agreement, and k=0 means that the agreement occurred by chance (Cohen, 1960). In Martín-Montero et al. (2023), the k statistic was used to evaluate the overall performance of the LSBoost model in pediatric OSA diagnosis, as well as for the AdaBoost model in sleep stage classification. Moreover, the optimization of the MLP models in Martín-Montero et al. (2021a) used this measure of agreement by selecting the combination of N_H/N_{AB} and λ/ν that maximizes k in each case.

4.7.4 Validation approaches

Validation approaches are techniques frequently used to improve the reliability and generalization of research findings. It is a common practice to divide the available datasets into subsets to fit the required optimization steps and reduce the risk of overfitting (Witten et al., 2011). To meet these requirements, two validation techniques were used in this Doctoral Thesis. The hold-out method was used to divide the available databases into training-test sets (Martín-Montero et al., 2021a,b) or training-validation-test sets (Martín-Montero et al., 2023). Furthermore, in Martín-Montero et al. (2021a, 2022), bootstrapping techniques were used for different purposes.

- **Hold-out validation**. In order to ensure accurate validation of a model, it is recommended to use different sets to optimize the various stages of the methodology followed (Witten et al., 2011). The simplest form of the hold-out validation strategy involves fitting the model parameters using a training set and estimating the performance of the model using an independent test set (the hold-out group) (Bishop, 2006; Witten et al., 2011). If the model fitting process requires more than one optimization phase, the dataset should be split into as many sets as needed (Witten et al., 2011). Typically, when it occurs, three subsets are used: the training set for adjusting

the model parameters, the validation set for fitting the hyperparameters, and the test set for independently assessing diagnostic yield. When there is insufficient data to use a different set for each phase, hold-out should be combined with additional validation techniques such as bootstrapping to ensure robustness and generalization of the obtained results. For instance, in Martín-Montero et al. (2021b), the private database was used as a training set, while the public database was the test set to increase the generalization of the results. In Martín-Montero et al. (2021a), the same division was applied for comparison purposes, but since the MLP models needed to optimize hyperparameters, bootstrapping was also used. Finally, in Martín-Montero et al. (2023), the dataset was divided into training-validation-test sets to optimize the LSBoost and AdaBoost models. Concretely, for the non-randomized group, 75% of the subjects were allocated to the training set (comprising 567 recordings), 12.5% were assigned to the validation set (94 recordings), and the remaining 12.5% were designated to the test set (94 recordings). In the baseline and follow-up groups, among the 404 subjects undergoing a follow-up study, 50% were distributed to the training set (consisting of 404 recordings, 202 from baseline and 202 from follow-up), 25% were allocated to the validation set (101 recordings from baseline and 101 from follow-up), and the final 25% were included in the test set (comprising 101 recordings from baseline and 101 from follow-up). It is important to note that for each child with both baseline and follow-up recordings, both recordings were systematically incorporated into the same group to prevent any biases. The remaining 47 recordings from the baseline group without follow-up studies were included in the training set.

- **Bootstrapping**. This validation technique involves generating M bootstrap replicates on which to apply the proposed method (Efron and Tibshirani, 1994). A bootstrap replicate refers to a sample that is created by randomly resampling the original set with replacement. This means that some observations may be selected several times, while others may not be selected at all (Witten et al., 2011). Each bootstrap replicate has the same length than the original set, and the process is repeated M times, corresponding to the number of replicates created. As a result, M estimates will be generated for each statistical metric. In Martín-Montero et al. (2021a), bootstrapping was used for two different purposes. Firstly, in the feature-selection stage, the FCBF algorithm was implemented on 1000 bootstrap replicates of the train-

ing set, choosing as relevant those features selected over 500 times. Additionally, the optimization of the MLP hyperparameters was carried out on a different 1000 bootstrap replicates from the training group. Finally, in Martín-Montero et al. (2022) the confidence intervals of the CMA were established using 2000 non-parametric bootstrap iterations.

In this chapter, the methodology applied throughout the realization of this Doctoral Thesis have been presented. In the following chapter, the most relevant results obtained with this methodology will be presented.

Chapter 5

Results

The purpose of this chapter is to report the main findings of the Doctoral Thesis. The chapter is organized based on the various approaches that were applied as follows: 5.1 Spectral analysis and evaluation of the novel pediatric OSA-related frequency bands, 5.2 Bispectral analysis, 5.3 Causal mediation analysis, and 5.4 HRV segments characterization. As a result, the chapter is directly linked to the publications that compose the Doctoral Thesis compendium.

5.1 Spectral analysis: novel OSA-specific frequency bands

In Martín-Montero et al. (2021b), we conducted a search for frequency bands that would be able to reflect the specific alterations that OSA cause in the ANS. As detailed in Section 4.2.2, the search led to three novel pediatric OSA-specific frequency bands (see Figure 4.3): BW1 (0.001-0.005 Hz), BW2 (0.028-0.074 Hz), and BWRes (0.04 Hz around maximum value within HF). This frequency bands were established in the UofC database, that was settled as training group in this study.

Table 5.1 displays the obtained *RPs* for each severity group (median [interquartile range]) in both the training (UofC database) and test set (nonrandomized group from the CHAT database) for the classic and OSA-specific frequency ranges. The p-values resulting from the Kruskal-Wallis test have also been included. It can be observed that RP_{BW2}, RP_{LF} and LF/HF increased with the severity (p-value < 0.01 after Bonferroni correction), while RP_{BW1} decreased in both sets. Additionally, RP_{BWRes} and RP_{HF} decreased with increasing OSA sever-

Table 5.1. Features for the four OSA severity groups (median [interquartile range]) in the training (UofC database) and test (nonrandomized group from CHAT database) sets. This table has been adapted from Martín-Montero et al. (2021b).

Features	No-OSA	mild OSA	moderate OSA	severe OSA	*p*-value
Training set					
RP_{VLF}	0.370 [0.174]	0.359 [0.163]	0.381 [0.179]	0.371 [0.164]	0.675
RP_{LF}	0.225 [0.060]	0.224 [0.075]	0.235 [0.081]	0.244 [0.090]	**«0.01**
RP_{HF}	0.317 [0.179]	0.340 [0.195]	0.300 [0.218]	0.275 [0.213]	**<0.01**
LF/HF	0.706 [0.510]	0.697 [0.594]	0.814 [0.791]	0.892 [0.985]	**«0.01**
RP_{BW1}	0.083 [0.055]	0.082 [0.050]	0.083 [0.047]	0.071 [0.049]	**<0.01**
RP_{BW2}	0.169 [0.054]	0.175 [0.068]	0.185 [0.086]	0.213 [0.107]	**«0.01**
RP_{BWRes}	0.119 [0.110]	0.121 [0.121]	0.110 [0.115]	0.087 [0.098]	**«0.01**
Test set					
RP_{VLF}	0.337 [0.140]	0.332 [0.155]	0.282 [0.149]	0.342 [0.186]	0.200
RP_{LF}	0.218 [0.060]	0.227 [0.063]	0.222 [0.090]	0.259 [0.110]	**«0.01**
RP_{HF}	0.368 [0.167]	0.363 [0.184]	0.388 [0.198]	0.307 [0.217]	0.015
LF/HF	0.610 [0.407]	0.649 [0.462]	0.597 [0.539]	0.818 [0.886]	**<0.01**
RP_{BW1}	0.081 [0.044]	0.078 [0.039]	0.063 [0.045]	0.061 [0.043]	**<0.01**
RP_{BW2}	0.148 [0.055]	0.161 [0.062]	0.165 [0.078]	0.209 [0.113]	**«0.01**
RP_{BWRes}	0.132 [0.108]	0.123 [0.107]	0.134 [0.143]	0.103 [0.093]	0.004[a]

RP: relative power, OSA: obstructive sleep apnea, VLF: very low frequency, LF: low frequency, HF: high frequency.
p-values $< 10^{-4}$ after Bonferroni correction appears as $<< 0.01$.
[a]Non-significant after applying Bonferroni correction.
p-values statistically significant (< 0.01 after Bonferroni correction) appears in bold.

ity, but only reaching statistically significant differences in the training set for the Kruskal-Wallis test. However, if statistically significant differences are evaluated between each pair of severity groups, both databases reach the same results, showing statistically significant differences when comparing severe OSA group against the no-OSA and mild OSA groups. RP_{VLF} did not exhibit any statistically significant differences.

To help giving a physiological interpretation to the novel frequency bands, Table 5.2 presents the results in the test set of the Spearman's partial correlation analysis performed, which controlled for the influence of age, between the *RPs* and various polysomnographic indices. The analysis revealed that RP_{BW2} had the strongest correlations with OSA-related indices (positively correlated with AHI, OAHI, OAI and ODI). RP_{BW1} showed positive correlations with macro sleep disruption variables (#Awakenings and WASO) and negative correlations with AHI and OAHI. RP_{BWRes} had statistically significant negative correlations with the same variables than RP_{HF} (OAI and WASO), as well as with TAI. Of note, although the correlations observed were not notably strong, among all the indices exhibiting statistically significant values, our proposed features consistently demonstrated the highest $|\rho_S|$.

Table 5.2. Evaluation of the partial correlations in the test set (nonrandomized group from CHAT database) between features and the polysomnographic indices considered. This table has been adapted from Martín-Montero et al. (2021b).

Classic bands								
PSG index	RP_{VLF}		RP_{LF}		RP_{HF}		LF/HF	
	ρ_S	p-value	ρ_S	p-value	ρ_S	p-value	ρ_S	p-value
AHI	-0.031	0.391	**0.150**	**<0.01**	-0.075	0.040	0.118	0.001[a]
OAHI	-0.073	0.043	0.088	0.015	-0.012	0.737	0.046	0.207
OAI	-0.035	0.333	0.067	0.066	-0.031	0.392	0.052	0.154
ODI	0.039	0.289	**0.194**	**«0.01**	**-0.161**	**«0.01**	**0.195**	**«0.01**
#Awakenings	**0.133**	**<0.01**	0.036	0.324	-0.115	0.014	0.086	0.018
WASO	0.071	0.049	0.112	0.002[a]	**-0.146**	**<0.01**	**0.145**	**<0.01**
%N1	0.003	0.930	0.063	0.084	-0.040	0.266	0.058	0.111
%N2	-0.085	0.019	-0.076	0.038	0.098	0.007[a]	-0.112	0.002[a]
%N3	0.068	0.060	0.074	0.043	-0.089	0.014	0.101	0.005[a]
%REM	0.041	0.262	-0.083	0.022	0.030	0.404	-0.047	0.197
TAI	0.031	0.389	**0.128**	**<0.01**	-0.098	0.007[a]	**0.126**	**<0.01**
Novel bands								
PSG index	RP_{BW1}		RP_{BW2}				RP_{BWRes}	
	ρ_S	p-value	ρ_S	p-value			ρ_S	p-value
AHI	**-0.132**	**<0.01**	**0.233**	**«0.01**			-0.101	0.005[a]
OAHI	**-0.157**	**<0.01**	**0.164**	**«0.01**			-0.033	0.368
OAI	-0.096	0.008[a]	**0.149**	**<0.01**			-0.049	0.180
ODI	-0.033	0.358	**0.220**	**«0.01**			**-0.192**	**«0.01**
#Awakenings	**0.174**	**«0.01**	0.069	0.059			-0.096	0.008[a]
WASO	**0.186**	**«0.01**	0.054	0.141			**-0.195**	**«0.01**
%N1	0.001	0.969	0.087	0.017			-0.063	0.083
%N2	-0.073	0.045	-0.092	0.011			0.083	0.023
%N3	0.058	0.111	0.052	0.155			-0.066	0.069
%REM	0.048	0.187	-0.049	0.175			0.041	0.262
TAI	-0.059	0.105	**0.220**	**«0.01**			**-0.123**	**<0.01**

PSG: polysomnography, RP: relative power, VLF: very low frequency, LF: low frequency, HF: high frequency, AHI: apnea-hypopnea index, OAHI: obstructive AHI, OAI: obstructive apnea index, ODI: oxygen desaturation index, WASO: wake after sleep onset, %N1: sleep time in N1 stage, %N2: sleep time in N2 stage, %N3: sleep time in N3 stage, %REM: sleep time in REM stage, TAI: total arousal index.

[a]Non-significant after applying Bonferroni correction.

Statistically significant correlations (p-value < 0.01 after applying Bonferroni correction) appears in bold.

Following the HRV characterization, we wanted to compare the utility of the novel frequency bands in pediatric OSA context against the parameters commonly used. Thus, the diagnostic performance results in the test set have been collected in Table 5.3 for each measure computed individually and for the LDA models constructed using the classic and new OSA-specific frequency bands (with and without ABW1 and ABW2). RP_{BW2} achieved the highest overall performance in all three OSA severity cutoffs evaluated, with the highest AUC at

Table 5.3. Diagnostic performance for each feature individually, as well as for the LDA models in the test set (nonrandomized group from CHAT database). This table has been adapted from Martín-Montero et al. (2021b).

Feature/model	AHI cutoff	Se	Sp	Acc	AUC
RP_{VLF}	**1 e/h**	68.9	31.6	56.3	0.518
	5 e/h	33.0	65.0	60.2	0.456
	10 e/h	40.6	64.2	62.1	0.495
RP_{LF}	**1 e/h**	43.5	62.9	50.1	0.557
	5 e/h	52.7	58.4	57.6	0.590
	10 e/h	59.4	58.4	58.5	0.666
RP_{HF}	**1 e/h**	35.5	71.9	47.8	0.523
	5 e/h	39.3	68.1	63.8	0.540
	10 e/h	43.5	76.7	73.7	0.605
LF/HF	**1 e/h**	37.7	70.3	48.7	0.540
	5 e/h	45.5	66.8	63.7	0.567
	10 e/h	49.3	70.8	68.8	0.643
RP_{BW1}	**1 e/h**	66.3	45.3	**59.2**	0.559
	5 e/h	65.2	54.0	55.6	0.621
	10 e/h	69.6	52.3	53.9	0.624
RP_{BW1}	**1 e/h**	32.7	78.1	48.1	0.591
	5 e/h	45.5	82.0	**76.6**	0.670
	10 e/h	58.0	78.2	76.4	0.751
RP_{BWRes}	**1 e/h**	45.5	56.6	49.3	0.532
	5 e/h	44.6	64.0	61.2	0.571
	10 e/h	49.3	64.0	62.6	0.628
$\text{LDA}_{Classic}$	**1 e/h**	25.7	81.3	44.5	0.559
	5 e/h	46.4	72.2	68.4	0.633
	10 e/h	50.7	75.3	73.1	0.685
$\text{LDA}_{Specific}$ (with ABW1 and ABW2)	**1 e/h**	42.5	72.3	52.6	0.592
	5 e/h	50.0	80.9	76.4	0.688
	10 e/h	63.8	84.7	**82.8**	**0.796**
$\text{LDA}_{Specific}$ (without ABW1 and ABW2)	**1 e/h**	37.7	80.1	52.0	**0.597**
	5 e/h	48.2	80.8	76.0	**0.696**
	10 e/h	62.8	84.3	82.3	0.774

RP: relative power, VLF: very low frequency, LF: low frequency, HF: high frequency, AHI: apnea-hypopnea index, LDA: linear discriminant analysis, Se: sensitivity (%), Sp: specificity (%), Acc = accuracy (%), AUC: area under receiver-operating characteristics curve.
For each AHI cutoff, the highest Acc and AUC has been highlighted in bold.

all the cutoffs, and the highest Acc, Se and Sp at the 5 and 10 e/h thresholds. Regarding the LDA models, the OSA-specific frequency bands outperformed the classic ones in all diagnostic metrics, except Sp in the 1 e/h threshold, but showing a strongly unbalanced Se/Sp pair. When comparing the LDA models with and without ABW1 and ABW2, we just observed a slight decline in AUC in the two lowest severity thresholds if those bands were excluded. This fact, as well as the absence of statistically significant correlations with PSG indices observed

in Martín-Montero et al. (2021b) for those bands, led us to concluded that ABW1 and ABW2 had no clinical utility, considering them as spurious and discarding those ranges in further research.

5.2 Bispectral analysis

After defining and characterizing the novel HRV frequency bands specific to pediatric OSA, the following step involved examining phase coupling, Gaussianity, and nonlinearity of pediatric HRV in the presence of OSA. For this purpose, bispectral analysis was conducted in Martín-Montero et al. (2021a). To ensure a fair comparison, the same training/test set division as in Martín-Montero et al. (2021b) was utilized, with the UofC database as the training set and the non-randomized group of the CHAT database as the test set. Figure 5.1 displays the averaged bispectral matrix in the frequency range of 0-0.4 Hz for each considered OSA severity group. In this figure, for each child the bispectral matrix has been calculated, which contains the bispectral power values for each f_1, f_2 pair. Subsequently, each subject is assigned to the corresponding OSA severity group, and the bispectral power values are averaged, giving rise to the four subfigures represented in Figure 5.1. This figure illustrates that, in the no-OSA group, the bispectral power was primarily concentrated below 0.02 Hz, but as OSA worsened, it expanded to higher frequencies. As described in Section 4.2.2.2, six sub-regions were defined within this bispectral matrix: three based on the classic HRV frequency ranges and three based on the HRV OSA-specific frequency bands. Detailed 3D zoomed-in views of each averaged bispectral region, segmented by severity groups, can be found in Figures A1-A6 from Martín-Montero et al. (2021a). In the classic regions, a total of 42 features were computed, while for the OSA-specific bispectral matrices, 38 features were calculated. The FCBF algorithm was then applied to select two optimal subsets: one from each subgroup of regions. The results are presented in Figure 5.2, which shows that the optimal subset for the classic regions ($BISP_{Classic}$) comprised VLF_f2m, LF_BE_2, and HF_PE, whereas the optimal subset for the OSA-specific regions ($BISP_{Specific}$) consisted of $BW2\_RP_{Diag}$, $BW2\_BE_1$, $BWRes\_Bmin$, and $BWRes\_BE_3$.

Figures 5.3 and 5.4 depict the boxplot distributions for each OSA severity group of the features chosen in the two optimal subsets, $BISP_{Classic}$ and $BISP_{Specific}$, respectively. The corresponding p-values from the Kruskal-Wallis test are also provided. In relation to the features within the $BISP_{Classic}$ subset, a discernible pattern was observed. VLF_f2m exhibited an increasing trend,

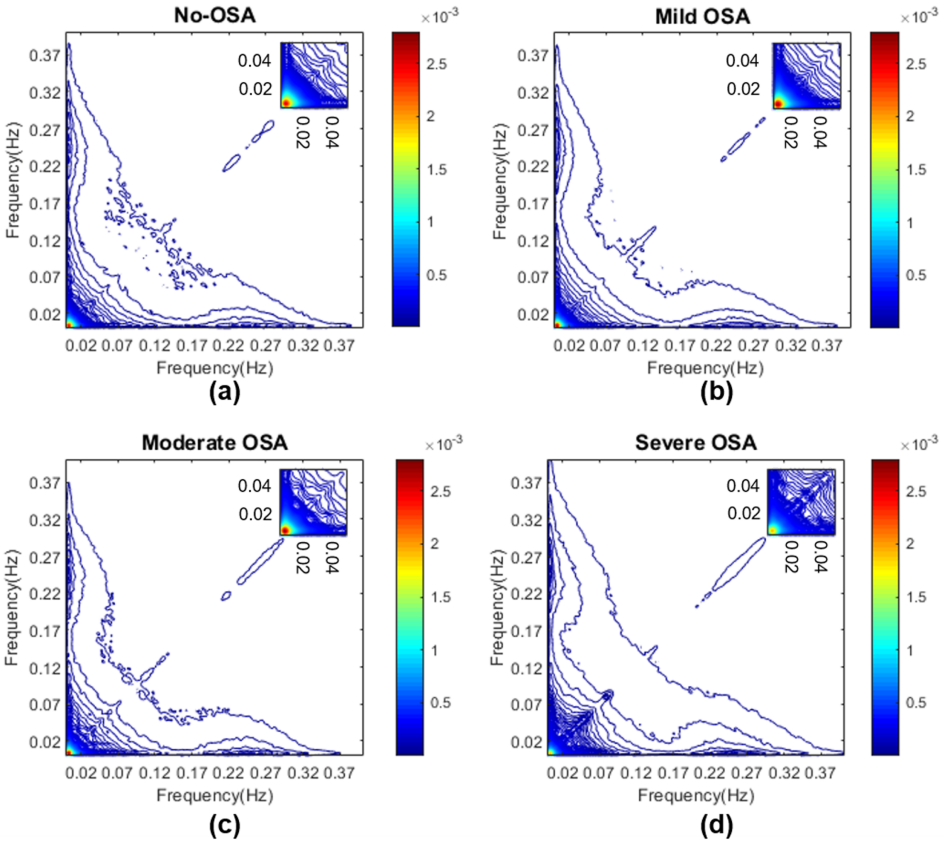

Figure 5.1. Bispectral matrix in the range 0-0.4 Hz averaged for each severity group in the training set. In the upper right corner, a zoomed representation between 0-0.05 Hz have been depicted, to better visualize the coupling focus. (a) No-OSA group (AHI < 1 e/h), (b) mild OSA group ($1 \leq AHI < 5$ e/h), (b) moderate OSA group ($5 \leq AHI < 10$ e/h), (b) severe OSA group ($AHI \geq 10$ e/h). This figure has been derived from Martín-Montero et al. (2021a).

whereas LF_BE_2 and HF_PE demonstrated a decreasing trend as OSA severity worsened. On the other hand, for the features comprising the $BISP_{Specific}$ subset, $BW2\_RP_{Diag}$ and $BWRes\_BE_3$ exhibited increments, particularly in the first parameter, while $BW2\_BE_1$ experienced a decrease across the OSA severity groups. Among all the features selected by the FCBF algorithm, only $BWRes\_Bmin$ did not show statistically significant differences between the OSA groups.

Subsequently, the features from both optimal subsets were utilized as input features to construct six MLP models for binary classification, two models for each AHI cutoff. Specifically, three MLP models were constructed using the

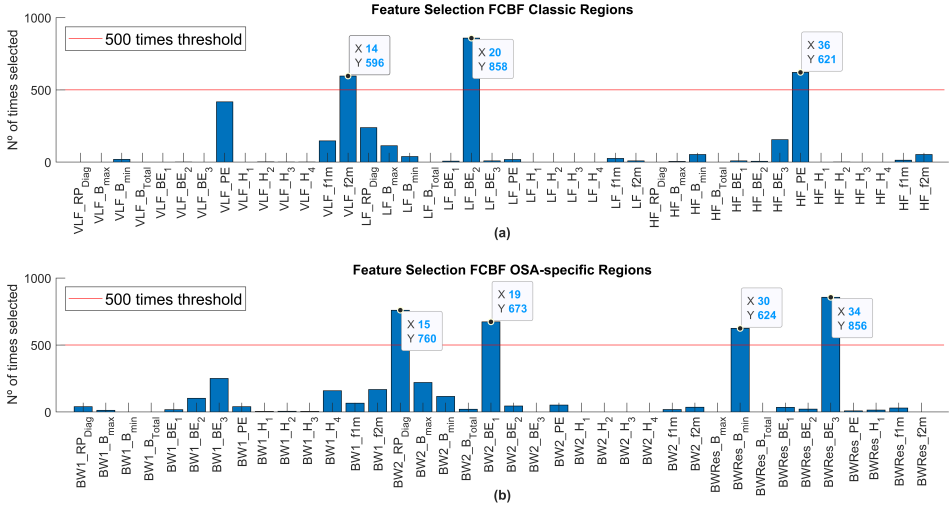

Figure 5.2. Results of the feature-selection stage applying FCBF algorithm in the training set using 1000 bootstrap replicates. The decision threshold (500 times) has been depicted as a red line. (a) Features selected from the classic regions, (b) Features selected from the OSA-related regions. This figure has been taken from Martín-Montero et al. (2021a).

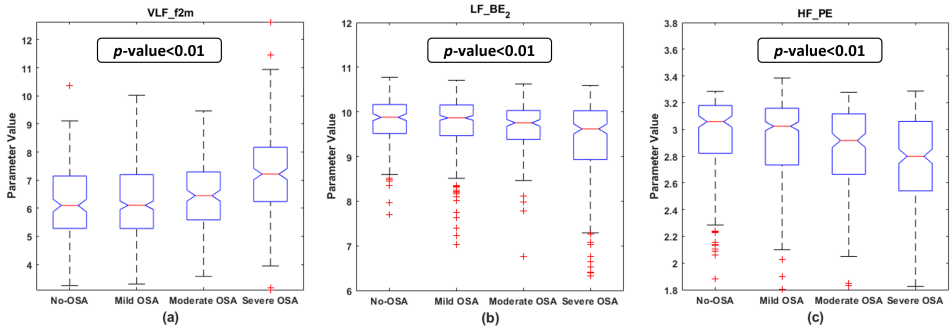

Figure 5.3. Boxplots representation of the features that were selected in the optimum $BISP_{Classic}$ subset, along with the p-value obtained from the corresponding Kruskal-Wallis test. (a) VLF_f2m boxplot distribution and p-value, (b) LF_BE_2 boxplot distribution and p-value, (c) HF_PE boxplot distribution and p-value. This figure has been taken from Martín-Montero et al. (2021a).

$BISP_{Classic}$ features, namely $MLP1_{Classic}$, $MLP5_{Classic}$, and $MLP10_{Classic}$ models, corresponding to AHI cutoffs of 1, 5, and 10 e/h, respectively. Similarly, the remaining three models, $MLP1_{Specific}$, $MLP5_{Specific}$, and $MLP10_{Specific}$, were formed using the $BISP_{Specific}$ features. As outlined in Section 4.4.1.2, the hyperparameters of each MLP model were optimized through 1000 bootstrap replicates from the training set. The NH values tested ranged from 2 to 20 in steps of 1, and

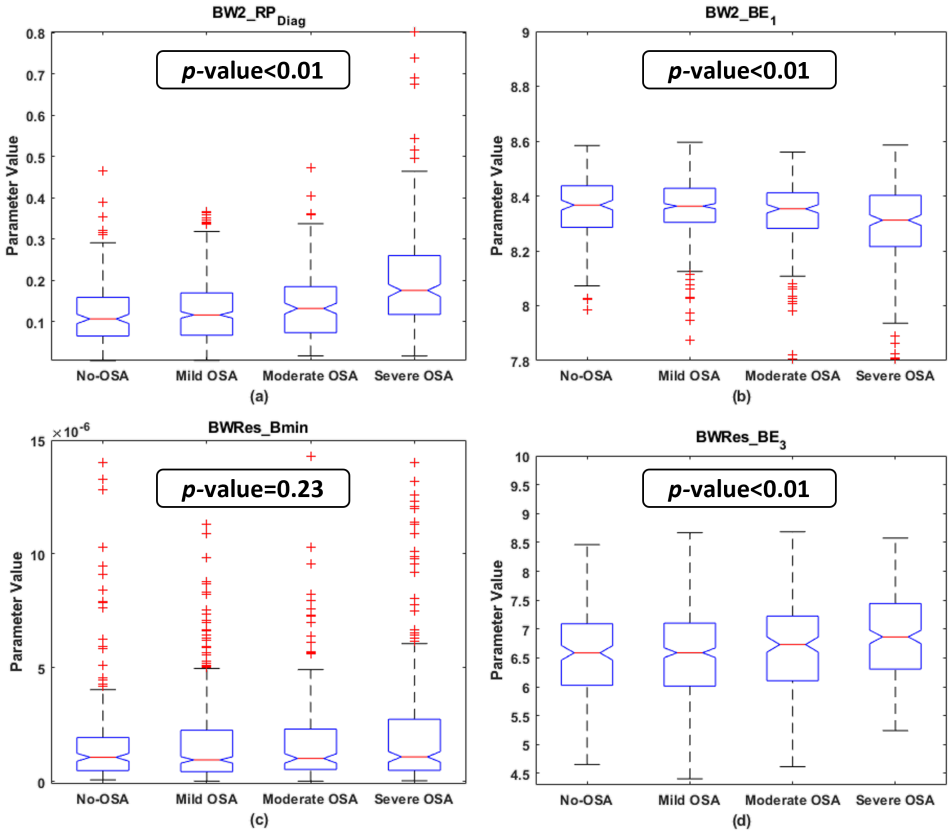

Figure 5.4. Boxplots representation of the features that were selected in the optimum $BISP_{Specific}$ subset, along with the p-value obtained from the corresponding Kruskal-Wallis test. (a) $BW2\_RP_{Diag}$ boxplot distribution and p-value, (b) $BW2\_BE_1$ boxplot distribution and p-value, (c) $BWRes\_Bmin$ boxplot distribution and p-value, (d) $BWRes\_BE_3$ boxplot distribution and p-value. This figure has been taken from Martín-Montero et al. (2021a).

from 22 to 50 in steps of 2, while the λ values ranged from 0.5 to 10 in steps of 0.5. The combination of hyperparameters resulting in the highest k value was found to be $N_H = 2$ and $\lambda = 5$ in four out of the six models ($MLP1_{Classic}$, $MLP5_{Classic}$, $MLP5_{Specific}$, and $MLP10_{Specific}$), $N_H = 34$ and $\lambda = 5$ for the $MLP10_{Classic}$ model, and $N_H = 38$ and $\lambda = 5$ in the $MLP1_{Specific}$ model. Consequently, these specific hyperparameters were utilized to assess the diagnostic performance of all the developed MLP models.

Table 5.4 presents the diagnostic performance reached by each of the developed MLP models, along with the individual diagnostic yield for each severity thresholds of the features included in the optimal subsets. Regarding the individ-

Table 5.4. Evaluation of the diagnostic performance for binary classification by each feature selected individually, as well as each MLP optimized model in the test set (nonrandomized group from CHAT database). This table has been adapted from Martín-Montero et al. (2021a).

Threshold: AHI = 1 e/h				
Feature/Model	**Se**	**Sp**	**Acc**	**AUC**
VLF_f2m	44.5	72.3	53.9	0.605
LF_BE_2	42.1	72.7	52.4	0.581
HF_PE	42.9	63.3	49.8	0.550
$BW2\_RP_{Diag}$	50.9	64.8	55.6	**0.629**
$BW2\_BE_1$	47.1	59.4	51.3	0.559
$BWRes\_Bmin$	40.5	57.4	46.2	0.482
$BWRes\_BE_3$	41.5	57.4	46.9	0.513
$MLP1_{Classic}$	52.3	59.4	54.7	0.600
$MLP1_{Specific}$	76.3	38.3	**63.4**	0.627
Threshold: AHI = 5 e/h				
Feature/Model	**Se**	**Sp**	**Acc**	**AUC**
VLF_f2m	62.5	72.2	70.8	0.749
LF_BE_2	56.3	74.4	71.7	0.670
HF_PE	45.5	72.1	68.2	0.628
$BW2\_RP_{Diag}$	60.7	77.7	75.2	0.747
$BW2\_BE_1$	56.3	70.1	68.0	0.671
$BWRes\_Bmin$	58.9	45.3	47.3	0.567
$BWRes\_BE_3$	47.3	58.4	56.8	0.569
$MLP5_{Classic}$	50.9	86.2	**81.0**	0.774
$MLP5_{Specific}$	62.5	84.2	**81.0**	**0.791**
Threshold: AHI = 10 e/h				
Feature/Model	**Se**	**Sp**	**Acc**	**AUC**
VLF_f2m	63.8	76.7	75.6	0.784
LF_BE_2	58.0	81.5	79.4	0.740
HF_PE	53.6	72.1	70.4	0.663
$BW2\_RP_{Diag}$	68.1	76.0	75.3	0.789
$BW2\_BE_1$	47.8	76.0	73.4	0.692
$BWRes\_Bmin$	56.5	50.6	51.1	0.557
$BWRes\_BE_3$	55.1	59.4	59.0	0.614
$MLP10_{Classic}$	43.5	96.5	**91.7**	**0.847**
$MLP10_{Specific}$	66.7	91.6	89.3	0.841

VLF: very low frequency, LF: low frequency, HF: high frequency, MLP: multi-layer perceptron, AHI: apnea-hypopnea index, Se: sensitivity (%), Sp: specificity (%), Acc: accuracy (%), AUC: area under receiver-operating characteristic curve.
For each AHI cutoff, the highest Acc and AUC has been highlighted in bold.

ual features, $BW2\_RP_{Diag}$ demonstrated the best overall performance. It achieved the highest Acc and AUC in the 1 e/h, the highest Acc in the 5 e/h, and the highest AUC in the 10 e/h thresholds. Notably, $BW2\_RP_{Diag}$ exhibited a more balanced Se/Sp pair than LF_BE_2, which had a higher Acc in the 10 e/h thresh-

old. Additionally, all the MLP models outperformed the individual diagnostic ability of the features. The MLP models based on OSA-specific region features showed the highest diagnostic ability in the 1 and 5 e/h thresholds, while the $MLP10_{Classic}$ model outperformed the rest in the 10 e/h cutoff, albeit with a significantly unbalanced Se/Sp pair.

5.3 Causal mediation analysis

In Martín-Montero et al. (2022), our aim was to assess whether changes in HRV parameters following OSA treatment could reliably identify variations in OSA severity indices or reflect OSA resolution. To accomplish this, we conducted a CMA, as described in Section 4.6, by averaging mediation effects across treatment arms. Through this analysis, we were able to evaluate the influence of the chosen mediator variables (OSA severity indices and resolution) on the features selected as outcomes (extracted from the HRV), as well as the extent of these impact (represented by ACME values). Additionally, it allowed us to identify the remaining effects unassociated with the mediators (termed as ADE). The results of the CMA are presented in Table 5.5, which shows the averaged values of ACME and ADE, along with their corresponding 95th percentile confidence intervals for each mediator and outcome considered in the study. The associated p-values can be consulted in Supplementary Table S11 from Martín-Montero et al. (2022). Notably, among all the evaluated parameters, only ΔRP_{BW2} exhibited statistically significant ACMEs for all the mediators, and these effects were mediated in the negative direction. Moreover, ΔRP_{BW2} was the only parameter that showed statistically significant ACME with OSA resolution, while obtaining non-statistically significant ADE effects with $\Delta Minsat$ and any of the resolution mediators. This highlights the importance of these mediators in the context of BW2. Regarding the OSA_{O+C} resolution mediator, apart from BW2 it also mediated changes in $\Delta(LF/HF)$. All the remaining parameters obtained ACME effects to some extent in at least one of the mediators considered, except for ΔRP_{VLF}, which did not achieve any statistically significant ACME.

After performing the CMA, we proceeded to examine whether changes mediated by the resolution mediators could differentiate the HRV activity of children who experienced OSA resolution compared to those who did not, following randomization. Accordingly, we evaluated differences at follow-up between children with resolved and unresolved OSA using both resolution criteria, focusing on parameters where the change in the mediator resulted in a statistically signifi-

Table 5.5. Results from the CMA evaluating treatment effects on variations in HRV features (follow-up - baseline) through several mediators. This table has been derived from Martín-Montero et al. (2022).

Mediator		ΔRP_{VLF}	ΔRP_{LF}	ΔRP_{LF}	$\Delta LF/HF$	ΔRP_{BW1}	ΔRP_{BW2}	ΔRP_{BWRes}
ΔAHI	ACME	-0.001	**-0.007** *	0.005	**-0.033**	0.002	**-0.011***	0.003
	(95% CI)	(-0.008 to 0.006)	**(-0.011 to -0.002)**	(-0.001 to 0.011)	**(-0.062 to -0.007)**	(-0.001 to 0.005)	**(-0.017 to -0.007)**	(-0.001 to 0.008)
	ADE	0.047 *	0.018 *	-0.066 *	0.204 *	0.011 *	0.015 *	-0.045 *
	(95% CI)	(0.024 to 0.069)	(0.008 to 0.027)	(-0.086 to -0.047)	(0.131 to 0.280)	(0.003 to 0.018)	(0.006 to 0.024)	(-0.060 to -0.031)
ΔOAHI	ACME	-0.001	**-0.006**	0.004	**-0.029**	**0.003**	**-0.011***	0.002
	(95% CI)	(-0.007 to 0.005)	**(-0.010 to -0.002)**	(-0.002 to 0.010)	**(-0.059 to -0.004)**	**(0.001 to 0.006)**	**(-0.016 to -0.006)**	(-0.002 to 0.010)
	ADE	0.047 *	0.017	-0.065 *	0.200 *	0.010	0.015 *	-0.044 *
	(95% CI)	(0.025 to 0.069)	(0.008 to 0.027)	(-0.083 to -0.045)	(0.125 to 0.280)	(0.003 to 0.018)	(0.006 to 0.023)	(-0.059 to -0.030)
OSA_{O+C} resol	ACME	-0.001	-0.001	0.002	**-0.015**	0.001	**-0.003**	0.001
	(95% CI)	(-0.006 to 0.003)	(-0.004 to 0.001)	(-0.002 to 0.007)	**(-0.040 to -0.003)**	(-0.001 to 0.003)	**(-0.005 to -0.001)**	(-0.003 to 0.004)
	ADE	0.046 *	0.012	-0.063 *	0.193 *	0.012 *	0.006	-0.043 *
	(95% CI)	(0.027 to 0.068)	(0.003 to 0.022)	(-0.082 to -0.045)	(0.115 to 0.263)	(0.005 to 0.019)	(-0.003 to 0.016)	(-0.056 to -0.029)
OSA resol	ACME	-0.002	-0.001	0.001	-0.033	0.002	**-0.005**	-0.001
	(95% CI)	(-0.012 to 0.008)	(-0.005 to 0.004)	(-0.007 to 0.010)	(-0.072 to 0.008)	(-0.002 to 0.005)	**(-0.010 to -0.001)**	(-0.007 to 0.005)
	ADE	0.048 *	0.012	-0.062 *	0.199 *	0.011	0.009	-0.041
	(95% CI)	(0.026 to 0.068)	(0.002 to 0.022)	(-0.082 to -0.043)	(0.120 to 0.280)	(0.003 to 0.019)	(-0.001 to 0.018)	(-0.055 to -0.026)
ΔODI	ACME	-0.001	**-0.005**	**0.005**	**-0.027**	0.002	**-0.008** *	0.003
	(95% CI)	(-0.006 to 0.005)	**(-0.009 to -0.002)**	**(0.001 to 0.010)**	**(-0.051 to -0.008)**	(-0.001 to 0.004)	**(-0.012 to -0.004)**	(0.001 to 0.007)
	ADE	0.046 *	0.016 *	-0.066 *	0.198 *	0.011	0.012	-0.045 *
	(95% CI)	(0.025 to 0.070)	(0.007 to 0.025)	(-0.084 to -0.048)	(0.127 to 0.271)	(0.004 to 0.019)	(0.003 to 0.020)	(-0.059 to -0.031)
ΔMinsat	ACME	-0.002	-0.001	**0.004**	-0.010	0.001	**-0.002**	0.003
	(95% CI)	(-0.006 to 0.002)	(-0.005 to 0.001)	**(0.001 to 0.008)**	(-0.027 to 0.002)	(-0.001 to 0.002)	**(-0.005 to -0.001)**	(-0.001 to 0.007)
	ADE	0.048 *	0.013	-0.064 *	0.180 *	0.012	0.006	-0.045 *
	(95% CI)	(0.027 to 0.067)	(0.003 to 0.022)	(-0.083 to -0.047)	(0.110 to 0.256)	(0.005 to 0.020)	(-0.003 to 0.016)	(-0.059 to -0.030)
ΔTAI	ACME	-0.001	**-0.008** *	**0.007**	**-0.028**	**0.002**	**-0.011***	**0.005**
	(95% CI)	(-0.006 to 0.005)	**(-0.012 to -0.004)**	**(0.002 to 0.013)**	**(-0.056 to -0.004)**	**(0.001 to 0.005)**	**(-0.016 to -0.006)**	**(0.001 to 0.009)**
	ADE	0.046 *	0.019 *	-0.067 *	0.199 *	0.011	0.015 *	-0.047 *
	(95% CI)	(0.026 to 0.067)	(0.010 to 0.028)	(-0.086 to -0.049)	(0.129 to 0.268)	(0.003 to 0.018)	(0.006 to 0.023)	(-0.061 to 0.033)

RP: relative power, VLF: very low frequency, LF: low frequency, HF: high frequency, AHI: apnea-hypopnea index, OAHI: obstructive AHI, OSA_{O+C} resol: obstructive and central apnea resolution, OSA resol: obstructive sleep apnea resolution, ODI: Oxygen desaturation index 3%, Minsat: minimum saturation level, TAI: total arousal index, ACME: average causal mediation effect, ADE: average direct effect, CI: confidence interval.
Statistically significant ACME effects ($p < 0.05$) appears in bold.
*p<0.001.

cant ACME in each case. The results are presented in Figure 5.5, which displays the boxplot distributions of RP_{BW2} at follow-up of each group for both mediators, as well as LF/HF for the OSA_{O+C} resolution mediator. As can be seen, RP_{BW2} was the only feature that exhibited a discernible distinction (p-value < 0.05) in HRV activity between children who experienced resolution of the disease and those who did not.

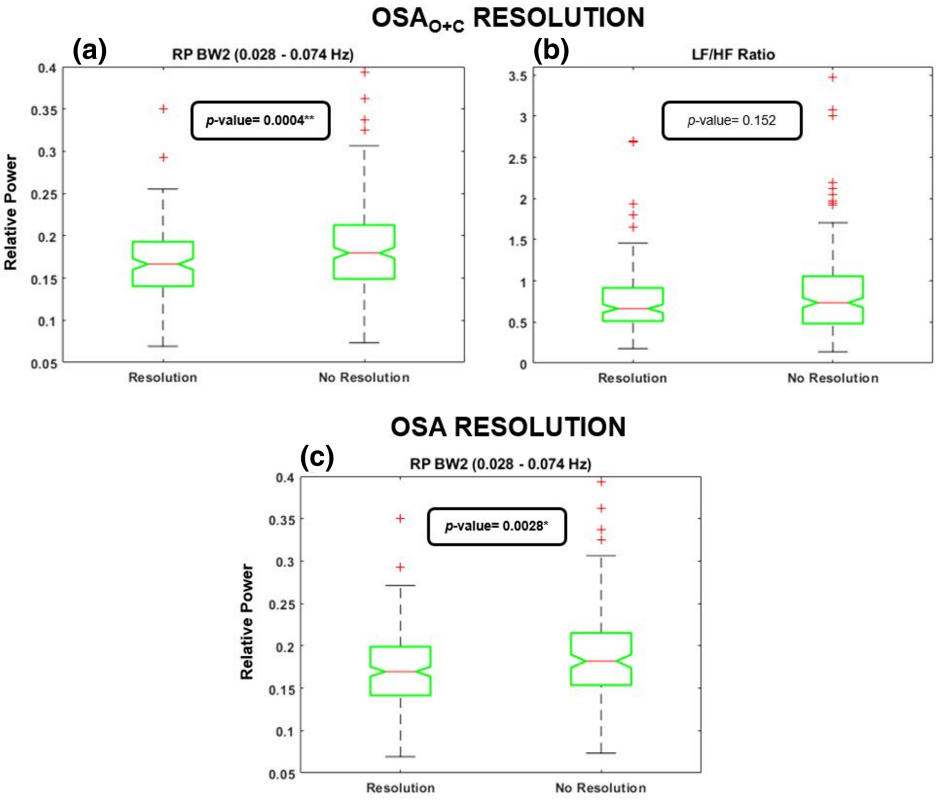

Figure 5.5. Differences at follow-up between resolved and unresolved children when evaluating both resolution mediators. Only the boxplots distribution of the parameters that showed ACME effects with those mediators were evaluated. The p-values resulting from the Mann-Whitney U tests are also provided. (a) Differences in RP_{BW2} using OSA_{O+C} resolution criteria.(b) Differences in LF/HF ratio using OSA_{O+C} resolution criteria. (c) Differences in RP_{BW2} using OSA resolution criteria. This figure has been taken from Martín-Montero et al. (2022).

5.4 HRV segments characterization

The outcomes of the previous investigations primarily focused on examining the behavior of HRV throughout the entire night. In the last step, we wanted to assess how sleep stages and the presence of apneic events influence overnight HRV in the context of pediatric sleep apnea. To this end, in Martín-Montero et al. (2023), HRV 10-min segments were characterized and analyzed depending on the sleep stage (wake,W; NREM; and REM) as well as the number of apneic events contained within each segment (less than 1 event per segment, e/s; 1 to 5 e/s; 5 to 10 e/s; and above 10 e/s). Figures 5.6-5.8 depict the boxplot distributions in the training set for each segment type, representing the features extracted from the temporal domain, classic frequency bands, and OSA-related frequency bands, respectively. The corresponding p-values for the Mann-Whitney U test conducted between each pair of segments can be found in Supplemental Table 1 from Martín-Montero et al. (2023). However, as explained in Section 4.7, given the large number of segments included in the study, differences are better quantified using Cohen's d measure. Therefore, the Cohen's d values for each comparison are presented in Table 5.6.

In general terms, when comparing different severity segments within the same sleep stage, NREM segments exhibited a higher number of comparisons

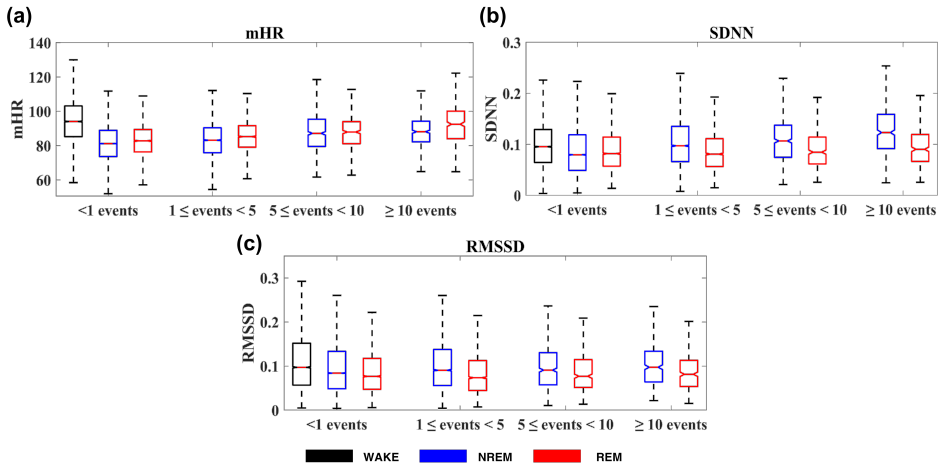

Figure 5.6. Boxplot distribution in the training set of the features in the temporal domain for every type of segment considered. (a) mHR boxplot distributions, (b) $SDNN$ boxplot distributions, (c) $RMSSD$ boxplot distributions. This figure has been derived from Martín-Montero et al. (2023).

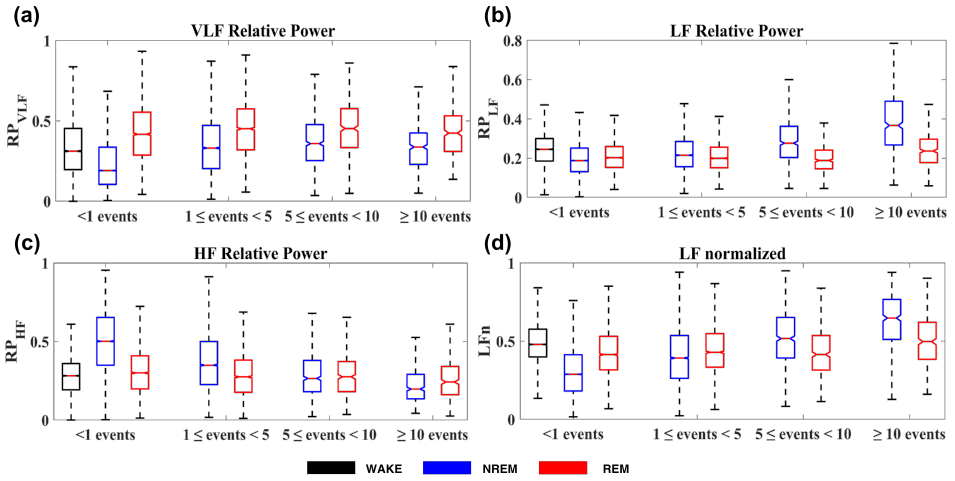

Figure 5.7. Boxplot distribution in the training set of the features in the frequency domain from the classic frequency bands for every type of segment considered. (a) RP_{VLF} boxplot distributions, (b) RP_{LF} boxplot distributions, (c) RP_{HF} boxplot distributions, (d) LF_n boxplot distributions. This figure has been derived from Martín-Montero et al. (2023).

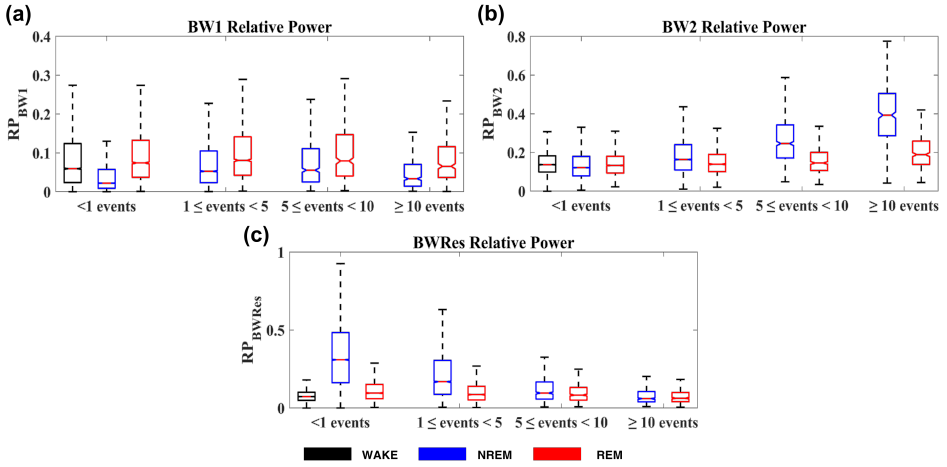

Figure 5.8. Boxplot distribution in the training set of the features in the frequency domain from the OSA-specific frequency bands for every type of segment considered. (a) RP_{BW1} boxplot distributions, (b) RP_{BW2} boxplot distributions, (c) RP_{BWRes} boxplot distributions. This figure has been derived from Martín-Montero et al. (2023).

with substantial effect sizes ($d \geq 0.5$), while the effects were reduced during REM. This fact is also reflected in the boxplots (Figures 5.6-5.8), where REM trends followed similar patterns to NREM but less pronounced. Among the intra-sleep

NREM stage comparisons, RP_{BW2} showed the highest effect sizes in five out of six comparisons, indicating an increase in RP_{BW2} with the presence of events (see Figure 5.8.b). This increase was also observed for mHR and $SDNN$ in the temporal domain (see Figure 5.6), as well as in LFn (see Figure 5.7.b). Conversely, RP_{BWRes} exhibited the opposite trend in these comparisons, showing medium or large effect sizes, as well as RP_{HF} to a lesser extent. In the case of intra-sleep REM stage, Table 5.6 reflects that several parameters had negligible or small effect sizes ($d \leq 0.2$), with mHR and RP_{BW2} exhibiting the highest effect sizes in these comparisons. Moving on to inter-sleep stage differences, in the absence of apneic events, only RP_{BW2} showed negligible effect sizes between stages, while the largest differences were observed in RP_{BWRes} (in W vs NREM and NREM vs REM comparisons) and mHR (in W vs REM comparison). When including apneic events, RP_{BWRes} obtained the highest differentiation between NREM and REM for segments containing 1 to 5 e/s, while RP_{BW2} regained relevance when comparing segments with 5 to 10 e/s, as well as those with above 10 e/s, consistently demonstrating the largest effect size.

After completing the feature-extraction stage, the ten extracted features were used as input for the AdaBoost and LSBoost models to assess the clinical usefulness of HRV characterization. The validation set was selected to optimize the hyperparameters for both models, as explained in sections 4.4.2 and 4.4.3. The hyperparameters combinations that yielded the highest Cohen's k value were as follows: $\lambda = 0.3$ and $N_{LSB} = 300$ for the LSBoost model, and $\nu = 0.1$ and $N_{AB} = 3000$ for the AdaBoost model. Following the optimization of the models, their clinical applicability was evaluated. The LSBoost model was assessed for diagnosing pediatric OSA, while the AdaBoost model was evaluated for sleep stage classification. Confusion matrices in the test set for both tasks, along with the corresponding multiclass Acc and Cohen's k values are presented in Figure 5.9. The performance of the LSBoost model was evaluated on a per-subject basis, while the performance of the AdaBoost model was computed per-segment. The darkness of the cells along the main diagonal in the confusion matrices indicates the performance of each model, with darker colors indicating better performance. The diagnostic metrics achieved by the LSBoost model in the test set for the per-subject binary classification task, considering each AHI cutoff, are presented in Table 5.7. Additionally, the diagnostic performance results of the AdaBoost model for sleep stage classification in the test set, considering individual sleep stages, are summarized in Table 5.8. Finally, to assess the role of each feature in both tasks, the results of the relative importance evaluation are displayed in Figure 5.10.

Table 5.6. Cohen's d values measuring the effect size linked to the comparison of the differences performed in the training set (CHAT database) between each pair of type of segments included in the study. This table has been derived from Martín-Montero et al. (2023).

Intra-Stages	**NREM Segments**					
Feature	<1 **vs** $1 \leq$ **events** <5	<1 **vs** $5 \leq$ **events** <10	<1 **vs** ≥ 10 **events**	$1 \leq$ **events** <5 **vs** $5 \leq$ **events** <10	$1 \leq$ **events** <5 **vs** ≥ 10 **events**	$5 \leq$ **events** <10 **vs** ≥ 10 **events**
mHR	**0.147**	**0.539****	**0.634****	**0.412***	**0.516****	0.100
$SDNN$	**0.318***	**0.456***	**0.745****	**0.140**	**0.433***	**0.312***
$RMSSD$	**0.076**	0.065	**0.123**	0.011	0.050	0.068
RP_{VLF}	**0.610***	**0.754***	**0.579***	**0.137**	0.035	**0.201**
RP_{LF}	**0.325**	**1.007*****	**2.038*****	**0.628****	**1.553*****	**0.733****
RP_{HF}	**0.633****	**1.042*****	**1.335*****	**0.452***	**0.773****	**0.413***
LF_n	**0.578****	**1.233*****	**1.912*****	**0.584****	**1.196*****	**0.625****
RP_{BW1}	**0.440***	**0.483***	**0.072**	0.028	**0.337**	**0.391**
RP_{BW2}	**0.546****	**1.531*****	**3.106*****	**0.779****	**2.017*****	**0.963*****
RP_{BWRes}	**0.608****	**1.023*****	**1.272*****	**0.541****	**0.859*****	**0.565****
Intra-Stages	**REM Segments**					
Feature	<1 **vs** $1 \leq$ **events** <5	<1 **vs** $5 \leq$ **events** <10	<1 **vs** ≥ 10 **events**	$1 \leq$ **events** <5 **vs** $5 \leq$ **events** <10	$1 \leq$ **events** <5 **vs** ≥ 10 **events**	$5 \leq$ **events** <10 **vs** ≥ 10 **events**
mHR	**0.242***	**0.500****	**0.904*****	**0.263***	**0.671****	**0.398***
$SDNN$	0.036	0.043	**0.148**	0.082	**0.190**	0.115
$RMSSD$	0.070	0.000	0.028	0.071	0.101	0.031
RP_{VLF}	**0.142**	**0.194**	0.064	0.054	0.080	0.142
RP_{LF}	0.046	**0.177**	**0.423***	0.134	**0.473***	**0.587****
RP_{HF}	**0.156**	**0.202***	**0.388***	0.049	**0.243***	0.213*
LF_n	**0.097**	0.043	**0.503****	0.054	**0.392***	**0.459***
RP_{BW1}	**0.067**	0.084	0.134	0.017	**0.203***	0.224*
RP_{BW2}	**0.105**	**0.220***	**0.875*****	0.113	**0.751****	**0.573****
RP_{BWRes}	**0.152**	**0.243**	**0.511****	0.103	**0.401**	**0.379**
Inter-Stages	<1 **event Segments**			$1 \leq$ **events** <5 **Segments**	$5 \leq$ **events** <10 **Segments**	≥ 10 **events Segments**
Feature	**W vs NREM**	**W vs REM**	**NREM vs REM**	**NREM vs REM**	**NREM vs REM**	**NREM vs REM**
mHR	**0.958*****	**0.755****	**0.115**	**0.190**	0.007	**0.298**
$SDNN$	**0.223***	**0.204***	**0.023**	**0.343***	**0.452***	**0.709****
$RMSSD$	**0.126**	**0.279***	**0.144**	**0.299***	**0.243***	**0.300**
RP_{VLF}	**0.547****	**0.441***	**1.074*****	**0.597****	**0.569****	**0.647****
RP_{LF}	**0.417***	**0.261***	**0.158**	**0.210***	**0.927*****	**1.096*****
RP_{HF}	**1.233*****	**0.286**	**0.954*****	**0.480***	0.060	**0.176**
LF_n	**1.168*****	**0.453***	**0.728****	**0.209***	**0.465***	**0.700****
RP_{BW1}	**0.589****	**0.008**	**0.739****	**0.334***	**0.313***	**0.539****
RP_{BW2}	**0.062**	0.021	**0.080**	**0.351***	**1.005*****	**1.420*****
RP_{BWRes}	**1.460*****	**0.514****	**1.133*****	**0.764****	**0.361***	0.072

NREM: non-rapid eye movement, REM: rapid eye movement, mHR: mean heart rate, SDNN: standard deviation of normal-to-normal interval time series, RMSSD: root mean square of successive differences of NN intervals, RP: relative power, VLF: very low frequency, LF: low frequency, HF: high frequency, W: wake.
Statistically significant differences ($p < 0.05$) appears in bold
*Small effect ($0.2 \leq d < 0.5$), **medium effect ($0.5 \leq d < 0.8$), *** large effect ($d \geq 0.8$).

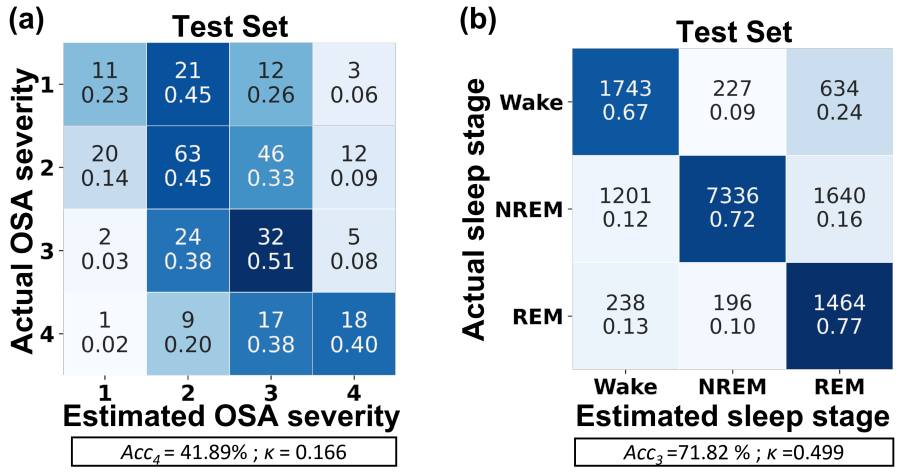

Figure 5.9. Confusion matrices in the test set for the tasks performed to assess the clinical applicability of the HRV segments characterization. (a) Confusion matrix for the LSBoost model in the per-subject pediatric OSA diagnosis task. 1: no-OSA group (AHI $<$ 1 e/h), 2: mild OSA (1 $\leq$ AHI $<$ 5 e/h), 3: moderate OSA (5 $\leq$ AHI $<$ 10 e/h), 4: severe OSA (AHI $\geq$ 10 e/h). (b) Confusion matrix for the AdaBoost model in the per-segment sleep stage classification task. This figure has been adapted from Martín-Montero et al. (2023).

Regarding the diagnosis of pediatric OSA, Figure 5.9.a shows that the LSBoost model tended to overestimate the AHI for the no-OSA group, achieving the worst proportion of correctly classified subjects, while moderate OSA subjects obtained the highest performance. The binary classification results presented in Table 5.7 demonstrate an overall improvement in diagnostic performance (higher AUC) with increasing severity cutoffs. These outcomes were primarily influenced by RP_{BW2}, which accounted for the largest proportion of feature importance (72.01%, as shown in Figure 5.10.a). Following RP_{BW2}, $SDNN$ and RP_{LF} were the next most influential features, accounting for 7.09% and 6.08% of the feature importance, respectively.

Finally, for sleep stage classification task, Figure 5.9.b shows that, in general, REM segments obtained the highest proportion of well-classified segments. Nevertheless, when examining the individual sleep stage classification results (Table 5.8), REM sleep stage had the lowest precision (39.7%) despite a high recall (77.13%). On the other hand, NREM achieved the highest precision (94.55%) and $F1 - score$ (0.818) among the considered sleep stages. Regarding feature importance in this task, Figure 5.10.b highlights that RP_{BWRes} was the feature with the greatest relative importance, accounting for 20.04% of the overall importance.

Table 5.7. Diagnostic yield achieved in the test set (subgroup from CHAT database) by the optimized LSBoost model for binary classification using the three AHI thresholds (1, 5 and 10 events/hour). This table has been derived from Martín-Montero et al. (2023).

Optimized LSBoost Model: $\lambda = 0.3; N_{LSB} = 300$										
	AHI	**Se**	**Sp**	**Acc**	**PPV**	**NPV**	**LR$^+$**	**LR$^-$**	**AUC**	**F1-score**
Test Set	**1 e/h**	90.76	23.40	80.07	86.26	32.35	1.18	0.39	0.651	0.885
	5 e/h	66.67	61.17	63.18	49.66	76.16	1.72	0.54	0.677	0.569
	10 e/h	40.00	92.03	84.12	47.37	89.53	5.02	0.65	0.742	0.434

AHI: apnea-hypopnea index, Se: sensitivity (%), Sp: specificity (%), Acc: accuracy (%), PPV: positive predictive value (%), NPV: negative predictive value (%), LR$^+$: positive likelihood ratio, LR$^-$: negative likelihood ratio, AUC: area under receiver-operating characteristic curve.

Table 5.8. Diagnostic yield achieved in the test set (subgroup from CHAT database) by the optimized AdaBoost model for individual sleep stage classification in the three stages considered (W, NREM and REM). This table has been derived from Martín-Montero et al. (2023).

Optimized AdaBoost Model: $\nu = 0.1; N_{AB} = 3000$				
	Sleep Stage	**Precision (%)**	**Recall(%)**	**F1-score**
Test Set	**W**	54.78	66.94	0.603
	NREM	94.55	72.08	0.818
	REM	39.17	77.13	0.516

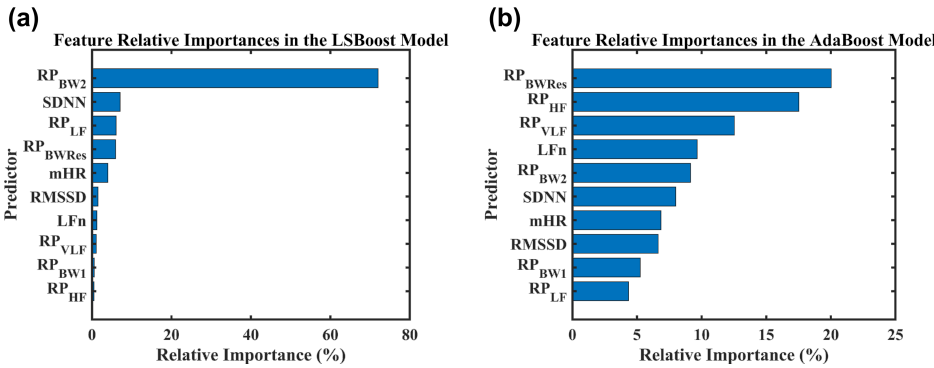

Figure 5.10. Evaluation of the relative feature importance for the models applied. (a) Relative importance of the features selected by the tree base classifiers used in the LSBoost model. (b) Relative importance of the features used as input in the AdaBoost model, obtained through LIME technique. This figure has been taken from Martín-Montero et al. (2023).

Chapter 6

Discussion

The present Doctoral Thesis is intended to characterize nocturnal HRV and assess ANS alterations associated with pediatric OSA and help in its diagnosis. To this end, various HRV signal processing techniques have been used, including temporal, spectral, and bispectral domain analyses. This methodology has enabled the evaluation of ANS alterations specific to pediatric OSA, unveiling novel OSA-specific frequency bands that provide valuable insights into HRV patterns throughout the night and during specific sleep stages. Additionally, the clinical utility of each approach has been demonstrated, along with the effects of pediatric OSA treatment on HRV in relation to changes in OSA severity and resolution using CMA. Thus, the main findings of this research are extensively discussed in this chapter. Furthermore, the clinical applicability of the different approaches for diagnosing pediatric OSA is compared, along with a comparison with state-of-the-art research works in this field. Lastly, the main limitations of the Doctoral Thesis are indicated.

6.1 Characterization of nocturnal HRV in children

As previously stated, spectral analysis (Martín-Montero et al., 2021b), bispectral analysis (Martín-Montero et al., 2021a), CMA (Martín-Montero et al., 2022), and HRV segmentation (Martín-Montero et al., 2023) approaches were followed to characterize alterations in the ANS that occurs due to OSA in the HRV dynamics.

6.1.1 Spectral analysis: novel OSA-specific spectral bands

The HRV spectral analysis conducted in Martín-Montero et al. (2021b) allowed to identify and assess three novel frequency bands of interest, aiming to enhance understanding of cardiovascular dynamics in pediatric OSA presence. These frequency bands exhibited significant correlations with respiratory events and indices of micro and macro sleep disruption, and demonstrated superior diagnostic ability compared to the conventional classic HRV spectral bands. These findings highlights the value of incorporating these novel frequency bands in the investigation of HRV in the field of pediatric OSA.

The first of these frequency bands, referred to as BW1 (0.001-0.005 Hz), encompasses a narrower fragment within the VLF range (0-0.04 Hz). While RP_{VLF} did not exhibit any statistically significant differences between OSA severity groups, RP_{BW1} reached significant differences in both the training and test sets. Additionally, RP_{BW1} exhibited superior individual diagnostic yield compared to RP_{VLF}. Moreover, this novel frequency band was statistically significant correlated with #Awakenings and WASO. Collectively, these findings suggest that sleep fragmentation, a direct consequence of pediatric OSA, primarily manifests within the frequency range covered by BW1, thereby capturing OSA-related alterations more effectively than VLF.

The second frequency band of interest identified was BW2 (0.028-0.074 Hz), which falls between the classic VLF and LF ranges. RP_{BW2} exhibited the strongest correlations with OSA respiratory indices and TAI, and reached the highest individual diagnostic performance in overall terms, particularly for the 5 and 10 e/h AHI thresholds. Previous research analyzing airflow signals in adults also reported a similar range of interest for OSA (approximately 0.025-0.050 Hz) (Gutiérrez-Tobal et al., 2013, 2012). This similarity may be attributed to the presence of cardio-respiratory coordination during OSA (Riedl et al., 2014). However, BW2 is broader than the frequency band observed in adults. These slight dissimilarities may be explained by variations in the disease between adults and children, as well as a shorter duration of cardiac events compared to respiratory events. As explained in Section 1.3, in pediatric OSA, an apneic event is scored if it lasts at least 2 respiratory cycles (approximately 6 seconds) (Berry et al., 2020). Therefore, the frequency range covered by BW2, which captures recurrence patterns between 13 and 35 seconds, aligns with the annotation of apneic events, and suggests a duration range of cardiac patterns as a result of pediatric OSA. Furthermore, the implications of pediatric OSA on the ANS can be inferred from

changes in BW2 activity. It is well-known that apneic events lead to the activation of the SNS, often reported as an increase in LF band (Aljadeff et al., 1997; Gozal et al., 2013; Walter et al., 2016). LF is generally associated with both SNS and PNS activity, with the PNS showing a faster response time (Nisbet et al., 2014). Accordingly, we propose that alterations in these frequency domains caused by OSA (which consistently elevates SNS activity) are specifically reflected in the BW2 frequency range. Based on these findings, we propose that OSA-related cardiac recurrent patterns due to alterations of the ANS are specifically captured by BW2, being particularly relevant for spectral analysis of HRV in the context of pediatric OSA.

The last OSA-related spectral band identified was BWRes, an individually adaptive frequency band that encompasses a 0.04 Hz range centered around the respiratory peak, which represents the highest value within the HF range. The rationale behind exploring this frequency band was to capture the respiratory modulation of HRV activity while mitigating the age-related effects that can influence activity within the classic HF range. As expected, we observed that the higher the OSA severity, the lower the magnitude of RP_{BWRes}. This band exhibited negative correlations with TAI, WASO and, interestingly, only ODI among all the respiratory indices evaluated, with these correlations being stronger than those observed for RP_{HF}. This suggests that normal respiratory activity is disrupted by awakening and micro-awakening periods, which may occur as a result of oxyhemoglobin desaturation (American Thoracic Society, 1996; Berry et al., 2012), thus leading to a decrease in HRV activity within BWRes. Besides, the AUC for RP_{BWRes} outperformed RP_{HF} in binary classification across all three AHI cutoffs, supporting the use of adaptive BWRes instead of HF when assessing HRV in children with OSA.

6.1.2 Bispectral analysis

To overcome the limitations of spectral analysis and examine the nonlinear, non-Gaussian, and non-stationary characteristics of HRV associated with OSA, we followed a bispectral approach in Martín-Montero et al. (2021a). This approach involved defining and extracting features from bispectral regions of HRV, which were based on both the traditional HRV frequency ranges and the OSA-specific frequency ranges identified in our earlier study. The feature-selection process conducted for both types of regions demonstrated the complementary nature of certain features, resulting in the identification of two optimal subsets of features.

The overall exploratory analysis revealed interesting findings, as depicted in

Figure 5.1. In the absence of OSA, there was a concentration of bispectral power below 0.02 Hz, indicating a certain coupling focus in the no-OSA group. However, with increasing OSA severity, there was a shift of bispectral power towards higher frequencies, leading to a significantly lower coupling focus in the severe OSA group. This dispersion of bispectral power to higher frequencies with OSA severity suggests an increase in HRV Gaussianity and linearity at VLF, likely due to OSA-related apneic events (Chua et al., 2010). This finding was further supported by the selected feature VLF_f2m, which exhibited increased values with OSA severity (see Figure 5.3), indicating the spreading of coupling focus from the VLF region (0-0.04 Hz range) to higher frequencies, specifically the BW2 region (0.028-0.074 Hz range), which is known to reflect alterations associated with apneic events. Additionally, a novel bispectral parameter developed during this doctoral thesis, RP_{Diag}, was also selected within BW2 region, and its values increased with OSA severity (see Figure 5.4). As explained in Section 4.2.2.2, RP_{Diag} measures the phase coupling between harmonic components of HRV. Increasing nonlinear interactions among the harmonics of children affected by OSA would be suggested by the higher $BW2\_RP_{Diag}$ values with OSA severity. Accordingly, the raising in this feature with OSA suggest that apneic events result in the emergence of less random and more periodic harmonics in HRV activity within this region. Notably, $BW2\_RP_{Diag}$ exhibited the highest individual diagnostic performance among the selected features. Therefore, these findings highlight the utility of the BW2 bispectral region, particularly in conjunction with the novel parameter RP_{Diag}, when studying HRV in pediatric OSA.

The three bispectral entropy features demonstrated their utility and were selected in both optimal subsets. Specifically, $BW2\_BE1$, LF_BE2, and $BWRes\_BE3$ were chosen, with the first two exhibiting decreases with OSA severity, while the latter increased with the disease. The decrease in $BW2\_BE1$ and LF_BE2 indicates a reduction irregularity within both frequency ranges in the presence of apneic events. This can be attributed to the previously mentioned increase in less random harmonics within BW2 region (partially overlapping with the LF region) as OSA severity worsens. It appears that, as BW2 region is more effective in reflecting the effects of apneic events, $BE1$ is sufficient to characterize irregularity changes caused by OSA in this region. However, in LF, the inclusion of quadratic amplitude ($BE2$) is necessary to capture these effects. Regarding BWRes, BE3 increased with OSA, reflecting an increase of irregularity. This may occur because in the absence of OSA, bispectral power in BWRes is concentrated around the respiratory peak. However, with the increasing presence of OSA effects, respiration

becomes more irregular, leading to a decrease in coupling around the respiratory peak and its redistribution to other frequencies. Consequently, HRV irregularity is amplified in the bispectral respiratory region, while irregularity in bispectral regions associated with apneic duration decreases as a result of OSA. Along with bispectral entropies, PE was also chosen in the HF region, decreasing as OSA severity increases. This indicates that the disease-related alterations result in a reduction in the irregularity of HRV phase. These findings highlights the significance of including entropy features when conducting HRV bispectral analysis to evaluate changes in irregularity caused by pediatric OSA.

The last of the features selected by the FCBF algorithm was $BWRes\_Bmin$. Although it did not exhibit a clear trend in its evolution across OSA severity groups (see Figure 5.4.c), its selection indicates that it provides complementary and non-redundant information to the other selected features in the optimum features subset specific to OSA. It is worth noting that none of the features computed in BW1 were chosen by the algorithm, suggesting that the analysis of BW2 and BWRes regions may be sufficient to characterize HRV patterns in the presence of pediatric OSA through bispectral analysis.

6.1.3 Causal mediation analysis

The computation of the CMA methodology in Martín-Montero et al. (2022) enabled us to infer that the changes observed in HRV following OSA intervention were causally linked to the effects of the treatment. This analysis established connections between alterations in HRV features and changes in polysomnographic indices directly related to OSA severity and its resolution. All the mediators considered were found to causally impact at least one of the evaluated outcomes. Among the various HRV features extracted, ΔRP_{BW2} was the only parameter that reached statistically significant ACMEs for each mediator included. It exhibited the highest level of statistical significance for most of the mediators, underscoring its importance in reflecting treatment effects on HRV activity.

The evaluation of disease resolution was conducted using two criteria: the conventional approach, as in the original CHAT study, and a stricter approach that also considered central apneas. When considering OSA resolution as a mediator (based on OAI and OAHI), the CMA revealed that this mediator only had a causal effect on ΔRP_{BW2}. When central apneas were included to establish OSA_{O+C} resolution as a mediator, two parameters reached statistically significant ACMEs: ΔRP_{BW2} and $\Delta LF/HF$. However, when evaluating the differences

in HRV activity between children who resolved OSA and those who did not for those features, it was found that resolution significantly affected only BW2. Although $\Delta LF/HF$ showed a statistically significant ACME with OSA_{O+C} resolution, there were no differences at follow-up between children with and without OSA_{O+C} resolution in LF/HF activity. It may be due to the higher ADE reached in the opposite direction, pointing to the presence of other factors that mask the effects of resolution on this feature. Therefore, BW2 activity was the only feature that allowed for differentiation between resolved and unresolved OSA children. In the light of these results, we propose BW2 as a potential biomarker for pediatric OSA resolution.

6.1.4 HRV segments characterization

In Martín-Montero et al. (2023), we conducted a characterization of nocturnal HRV segments, considering both sleep stages and the presence of apneic events. The applied methodology allowed us to uncover that, when comparing segments with varying degrees of apneic event presence, the differences were more pronounced in NREM sleep compared to REM sleep. It is known that during NREM sleep there is a decrease in SNS activity and an increase in PNS activity, which is reversed during REM sleep and wakefulness (Qin et al., 2021). Also, increments of sympathetic activation occurs in children with OSA (Baharav et al., 1999; Qin et al., 2021). The results revealed that features related to SNS activity (RP_{BW2}, mHR, $SDNN$, and RP_{LF}) exhibited clear increases with the presence of apneic events within NREM sleep (see Figures 5.6-5.8). Nevertheless, these effects were reduced during REM sleep, with RP_{BW2} and mHR being the only parameters showing considerable effect sizes in this sleep stage in three out of the six conducted comparisons. Interestingly, the reduction in differences between severity segments was observed during REM sleep, despite apneic events being more likely to occur in clusters during this stage (Qin et al., 2021). Thus, it appears that the inherent basal SNS excitation during REM sleep may mask the ANS alterations that occur as a result of apneic events. Therefore, considering that RP_{BW2} showed considerable effect size in all comparisons between severity segments during NREM sleep, and had the highest relative importance in the XAI technique of feature importance analysis for diagnosing pediatric OSA (see Figure 5.10), we emphasize the evaluation of BW2 activity as a reliable way to assess specific SNS activation due to pediatric OSA, especially when apneic events occur during NREM sleep.

The usefulness of RP_{BW2} assessment in characterizing HRV changes in the presence of apneic events appears to diminish in their absence, yielding negligible effects between sleep stages for non-apneic segments. Notwithstanding, another OSA-specific frequency range defined in this Doctoral Thesis, RP_{BWRes}, was the only parameter exhibiting medium to large effects in all sleep stage comparisons in the absence of apnea events. This can be attributed to the significant differences in respiratory patterns across sleep stages (Isler et al., 2016; Sazonova et al., 2006; Schechtman and Harper, 1992; Terrill et al., 2012), as RP_{BWRes} aims to reflect respiratory modulation in HRV. Respiratory control transitions from voluntary and automatic control during wakefulness to fully automatic control during sleep. It leads to regular respiratory patterns during NREM sleep and irregular respiration with brief breathing pauses occurring in healthy children during REM sleep (Marcus et al., 2008). Consistent with the effect size analysis, the application of LIME XAI technique to evaluate feature importance revealed that RP_{BWRes} attained the highest relative importance in the AdaBoost model for sleep stage classification. These findings highlight the significance of respiration in distinguishing between sleep stages, with these alterations effectively captured through HRV spectral analysis in the RP_{BWRes} parameter. The influence of respiration is further supported by the analysis of RP_{HF} and LF_n (influenced by activity in HF), which also reflects respiratory activity and reached high relative importance as well.

The increased activity in HF and BWRes in NREM compared to wake and REM sleep in non-apneic segments evidenced an excitation of PNS in NREM compared to the other two stages evaluated. However, when these segments accounted with apneic events, it occurs a reduction in PNS activity (decreased RP_{BWRes} and RP_{HF}), reaching levels similar to REM in $5 \leq$ apneic events < 10 e/s and ≥ 10 e/s segments. As a result of this decline, both features showed small or negligible effects sizes when differentiating NREM from REM. On the other hand, the increased presence of apneic events leads to RP_{BW2} allowing to differentiate activity of severity segments between NREM and REM, and reaching large effect sizes between NREM and REM for the $5 \leq$ apneic events < 10 e/s and ≥ 10 e/s segments. This result highlights the significance of BW2 to differentiate NREM and REM in presence of apneic events. Therefore, through these findings we have shown that both BWRes and BW2 novel OSA-specific frequency bands play crucial roles in differentiating sleep stages through HRV analysis in pediatric OSA context.

6.2 Clinical utility of HRV characterization

In the previous section, the clinical relevance of HRV characterization in reflecting the effects of OSA treatment through the CMA analysis performed in Martín-Montero et al. (2022) was discussed, highlighting the potential of analyzing BW2 as a biomarker for OSA resolution. Additionally, the three remaining research works in the compendium of publications have examined the clinical applicability of HRV characterization for diagnosing pediatric OSA, and also to automatically classify sleep stages in Martín-Montero et al. (2023).

6.2.1 Diagnostic performance

This Doctoral Thesis conducted various approaches to evaluate the ability of HRV characterization in diagnosing pediatric OSA. In Martín-Montero et al. (2021b), we compared the diagnostic ability of overnight HRV spectral analysis using classic frequency bands against OSA-specific frequency bands, performing binary classification based on the three commonly used AHI severity thresholds. In Martín-Montero et al. (2021a), we used the FCBF algorithm to develop two optimal subsets of HRV features. This study followed a similar approach to the first one, even using the same database and training/test division but focusing on bispectral analysis. Lastly, in Martín-Montero et al. (2023), we assessed the clinical utility of HRV segment characterization in diagnosing pediatric OSA by optimizing an LSBoost regression model to estimate AHI, also performing binary classification in the three AHI severity cutoff later. First of all, it is important to note that in both Martín-Montero et al. (2021b) and Martín-Montero et al. (2021a), the machine-learning models constructed by combining HRV features (LDA and MLP, respectively) outperformed the individual diagnostic performance of each feature. This highlights the usefulness of the feature-engineering approaches to integrate HRV information and aid in the diagnosis of pediatric OSA. For comparison purposes, Table 6.1 presents the results obtained by the machine-learning models developed in the three studies.

Regarding intra-study comparisons, the LDA model constructed with HRV features from the novel OSA-specific frequency band outperformed the diagnostic yield of the model using classic HRV features across all three AHI cutoffs, only being surpassed in terms of Sp in the 1 e/h threshold but showing a stronger unbalanced in the Se/Sp pair. Concerning bispectral feature models, consistent with the spectral analysis, the OSA-related region features models generally outper-

Table 6.1. Diagnostic yield achieved by the different machine-learning models through the studies from the present Doctoral Thesis in the three AHI severity thresholds. All results were obtained on the test set corresponding to each study, i.e., nonrandomized group from CHAT database in Martín-Montero et al. (2021b) and Martín-Montero et al. (2021a), and a subgroup formed by the different groups from CHAT database in Martín-Montero et al. (2023).

Threshold: AHI = 1 e/h										
Study	**Model**	**Se**	**Sp**	**Acc**	**PPV**	**NPV**	**LR$^+$**	**LR$^-$**	**AUC**	**F1-score**
Martín-Montero et al. (2021b)	**LDA$_{Classic}$**	25.70	81.30	44.50	72.90	35.90	1.37	0.91	0.559	0.326
	LDA$_{Specific}$	37.70	80.10	52.00	78.80	39.70	1.89	0.778	0.597	0.437
Martín-Montero et al. (2021a)	**MLP1$_{Classic}$**	52.30	59.40	54.70	71.58	38.87	1.29	0.80	0.600	0.535
	MLP1$_{Specific}$	76.30	38.30	63.40	70.74	45.16	1.24	0.62	0.627	0.693
Martín-Montero et al. (2023)	**LSBoost**	90.76	23.40	80.07	86.26	32.35	1.18	0.39	0.651	0.885
Threshold: AHI = 5 e/h										
Study	**Model**	**Se**	**Sp**	**Acc**	**PPV**	**NPV**	**LR$^+$**	**LR$^-$**	**AUC**	**F1-score**
Martín-Montero et al. (2021b)	**LDA$_{Classic}$**	46.40	72.20	68.40	22.50	88.60	1.67	0.74	0.633	0.553
	LDA$_{Specific}$	48.20	80.80	76.00	30.30	90.00	2.51	0.64	0.696	0.590
Martín-Montero et al. (2021a)	**MLP5$_{Classic}$**	50.90	86.20	81.00	39.04	91.00	3.69	0.570	0.774	0.625
	MLP5$_{Specific}$	62.50	84.20	81.00	40.70	92.82	3.96	0.45	0.791	0.706
Martín-Montero et al. (2023)	**LSBoost**	66.67	61.17	63.18	49.66	76.16	1.72	0.54	0.677	0.569
Threshold: AHI = 10 e/h										
Study	**Model**	**Se**	**Sp**	**Acc**	**PPV**	**NPV**	**LR$^+$**	**LR$^-$**	**AUC**	**F1-score**
Martín-Montero et al. (2021b)	**LDA$_{Classic}$**	50.70	75.30	73.10	17.10	93.80	2.05	0.65	0.685	0.599
	LDA$_{Specific}$	62.30	84.30	82.30	28.50	95.70	3.97	0.45	0.774	0.709
Martín-Montero et al. (2021a)	**MLP10$_{Classic}$**	43.50	96.50	91.70	55.56	94.45	12.43	0.59	0.847	0.590
	MLP10$_{Specific}$	66.70	91.60	89.30	44.23	96.48	7.94	0.36	0.841	0.764
Martín-Montero et al. (2023)	**LSBoost**	40.00	92.03	84.12	47.37	89.53	5.02	0.65	0.742	0.434

AHI: apnea-hypopnea index, Se: sensitivity (%), Sp: specificity (%), Acc: accuracy (%), PPV: positive predictive value (%), NPV: negative predictive value (%), LR$^+$: positive likelihood ratio, LR$^-$: negative likelihood ratio, AUC: area under receiver-operating characteristic curve, LDA: linear discriminant analysis, MLP: multi-layer perceptron, LSBoost: least-squares boosting.

formed the classic ones. At the AHI = 1 e/h cutoff, $MLP1_{Specific}$ achieved higher Acc and AUC than $MLP1_{Classic}$, albeit with an imbalanced Se/Sp pair. At the AHI = 5 e/h threshold, both models achieved the same Acc, but $MLP5_{Specific}$ reached higher AUC and a more balanced Se/Sp pair. In the highest severity threshold, the $MLP10_{Classic}$ model achieved the topmost Acc and AUC values among the models in the study. However, $MLP10_{Specific}$ achieved a more balanced Se/Sp pair at the expense of a slight decrease in Acc and AUC (see Table 6.1), also outperforming $MLP10_{Classic}$ model in terms of NPV, $F1-score$ and LR^-. These findings further support the notion that OSA-specific frequency bands in HRV analysis provide more specific information about pediatric OSA compared to the

classic features utilized so far. Finally, in the last study, the diagnostic yield of the LSBoost model improved with increasing OSA severity but yielded moderate overall diagnostic performance. Therefore, the optimal approach to leverage the LSBoost model would involve implementing an automated methodology to exclude the presence of OSA in its most severe form, as it achieved a Sp of 92.03% and NPV of 89.53% at the AHI 10 e/h cutoff, along with detecting the presence of OSA in its mildest form, with a Se of 90.76% and PPV of 86.26% in the lowest severity threshold.

Although the first two studies followed a whole-night HRV approach for pediatric OSA diagnosis, and the last study focused on segment-level analysis, all three studies conducted per-subject binary classification using the three AHI severity thresholds, allowing for comparison diagnostic performance results across the studies (see Table 6.1). The first two studies involved the UofC dataset as the training set and the nonrandomized group from CHAT database as the test set. When considering Acc and AUC as overall performance metrics, it is evident that the MLP models developed in the bispectral approach outperformed the LDA models using spectral analysis, consistently achieving the highest diagnostic yield throughout the entire Doctoral Thesis at the severity thresholds of 5 and 10 e/h. To address the possibility that this improvement could be attributed to the utilization of more sophisticated machine-learning techniques, Appendix B in Martín-Montero et al. (2021a) demonstrated that even with LDA models, the bispectral approach still outperformed spectral analysis. Hence, we propose that this improvement may also be attributed to the ability of bispectral analysis to capture the nonlinear dynamics present in overnight HRV in the presence of OSA, which are lost in spectral analysis. As mentioned, the overall diagnostic performance of the LSBoost model at the 5 and 10 e/h AHI cutoffs was lower than that of the MLP models. Notwithstanding, the LSBoost model achieved the highest diagnostic yield across the studies at the 1 e/h severity threshold in terms of Se, Acc, PPV, AUC, and $F1-score$, thus reinforcing the notion of the clinical utility of the LSBoost model constructed at the segment level to recognize OSA even in its mildest severity degree.

6.2.2 Automatic classification of sleep stages

In addition to evaluating the clinical usefulness of HRV segment characterization for pediatric OSA diagnosis in Martín-Montero et al. (2023), this study also aimed to assess the ability of this methodology to classify sleep stages as W, NREM,

or REM sleep using an AdaBoost model. Initially, it is evident that the performance metrics for sleep stage classification surpass those of the OSA diagnostic approach (see Figure 5.9), reaching significantly higher multiclass classification metrics in terms of Cohen's k (0.499 versus 0.166) and accuracy (Acc_3 = 71.82% versus Acc_4 = 41.89%). Among the sleep stages considered,the AdaBoost model reached the highest overall classification yield in NREM sleep (see Table 5.8), achieving a recall of 72.08% and precision of 94.55%. These findings are coherent with the results obtained during the characterization stage, where NREM sleep exhibited the greatest differences between severity segments.

In the light of the results achieved in Martín-Montero et al. (2023), we suggested a potential two-stage clinical application using a single-channel approach in the context of pediatric OSA and based on a segment-level HRV characterization. The first stage would focus on detecting NREM segments through HRV analysis, given its superior performance metrics. Subsequently, the second stage would involve estimating AHI based on these NREM segments, as those were the segments that allowed for the best differentiation.

6.3 Comparison with previous research works

To delve into the significance, novelty, and robustness of the findings obtained in this Doctoral Thesis, this section offers a comprehensive comparison with previous studies that have examined HRV in the context of pediatric OSA. Firstly, we present a comparison of our results with previous studies that have characterized HRV in pediatric OSA. Next, we compare our findings with prior research that has investigated the effects of OSA treatment on HRV. Lastly, we provide a comparison of the performance achieved in several research works that have employed cardiovascular signals for diagnosing pediatric OSA and classifying sleep stages.

6.3.1 Characterization of HRV in pediatric OSA context

To the best of our knowledge, the research conducted by us in Martín-Montero et al. (2021b) represented the first work to identify HRV frequency bands specifically associated with pediatric OSA. Similarly, the application of bispectral HRV analysis in the context of pediatric OSA, as carried out in Martín-Montero et al. (2021a), was a novel approach. However, it is important to acknowledge the similarities between our work and previous studies that characterized overnight HRV

using conventional parameters (i.e., temporal measures, such as mHR, SDNN or RMSSD, and spectral activity, measured as the relative power in the classic frequency ranges). The main findings of these previous studies have been collected in Table 6.2. The existing literature consistently reports an increase in LF activity and LF/HF ratio, along with a decrease in HF activity as the severity of OSA increases (Baharav et al., 1999; Horne et al., 2018; Liao et al., 2010a,b; Van Eyck et al., 2016; Walter et al., 2013). Moreover, prior studies have analyzed the effects of OSA on the ANS, showing incremented SNS activation due to intermittent hypoxia and hypercapnia (Narkiewicz and Somers, 1997; Somers et al., 1998), as well as altered autonomic reactivity during wakefulness in children, indicating that the effects of OSA extend after the sleep period (O'Brien and Gozal, 2005). These findings are coherent with the spectral analysis conducted and presented in Table 5.1, providing further support to the existing evidence that pediatric OSA disrupts HRV by elevating SNS activity and reducing PNS activity, leading to defective cardiac autonomic regulation. Furthermore, the lack of differences observed in RP_{VLF} (as shown in Table 5.1) was also reported in prior studies evaluating this frequency band (Gil et al., 2009; Kwok et al., 2011).

The segment-level characterization of HRV considering both sleep stages and the presence of apneic events was performed for the first time in the context of pediatric OSA in Martín-Montero et al. (2023). Nevertheless, it is possible to make comparisons with previous works that analyzed the behavior of classic HRV features across sleep stages. Kontos et al. (2020) examined HRV segments during sleep stages in healthy children and adolescents, obtaining comparable results to our findings in the absence of apneic events. They reported similar trends for mHR, which is actually the inverse of mean NN (the metric they evaluated), RP_{LF}, and RP_{HF} compared to our results in Figures 5.6 and 5.7 for non-apneic segments. Some previous studies comparing HRV segments between children with and without OSA excluded segments containing apneic events from their analysis, focusing solely on non-apneic segments. In this regard, three studies (Liao et al., 2010b; Vlahandonis et al., 2014; Walter et al., 2013) agreed in reporting a decrease in RP_{HF} from NREM to REM in all severity groups evaluated (healthy, primary snorers, mild OSA, and moderate-severe OSA children). We observed a similar decrease in RP_{HF} for non-apneic segments in our results (see Figure 5.7). Likewise, the study by Nisbet et al. (2013) evaluated HRV classical features across sleep stages, comparing different OSA severity groups while excluding apneic events from their analysis. They informed no differences in RP_{LF} between groups, suggesting that the disease may not cause autonomic dysfunction in pre-

Table 6.2. Methodological summary of the state-of-the-art research works focused on the characterization of HRV in pediatric OSA context, along with their relevant findings.

Study	Aim	Methodology	Sleep stage	Relevant findings
Baharav et al. (1999)	OSA effects on HRV	Spectral analysis	W, LS, SWS, REM	Higher LF and LF/HF with OSA in all sleep stages
Horne et al. (2018)	OSA effects on HRV	Spectral analysis	W, N1, N2, N3, REM	In NREM for OSA children, higher HR versus controls, and lower HF and higher LF versus normal weight PS
Liao et al. (2010a)	OSA effects on HRV	Spectral and temporal analysis	———	SDB decreased HF, RMSSD, SDNN and increased LF/HF and HR
Liao et al. (2010b)	OSA effects on HRV	Spectral and temporal analysis	W, NREM, REM	Lower HF in SDB vs controls in REM. Overall decrease in HF and RMSSD from NREM to REM
Van Eyck et al. (2016)	OSA effects on HRV	Spectral and temporal analysis	———	Increase in LF/HF and mHR with OSA severity
Walter et al. (2013)	OSA effects on HRV	Spectral analysis	W, LS, SWS, REM	In REM, all SDB severities showed lower LF and HF than controls. Overall decrease in LF and HF, and increase in LF/HF from NREM to REM
Narkiewicz and Somers (1997)	OSA effects on the ANS	ANS evaluation during hypoxia and hypercapnia	W, NREM, REM	Increased SNS due to intermittent hypoxia and hypercapnia, more marked during REM
Somers et al. (1998)	OSA effects on the ANS	ANS evaluation during hypoxia and hypercapnia	———	Sympathoexcitation due to hypoxia and hypercapnia, more marked in apnea subjects
O'Brien and Gozal (2005)	OSA effects on the ANS	Pulse arterial tonometry analysis in SDB children	Wakefulness	SDB children presented increased SNS activity during wakefulness, reflecting altered ANS regulation
Gil et al. (2009)	OSA effects on HRV	Spectral analysis	———	Increased sympathetic activity during DAP events. Absence of differences in VLF
Kwok et al. (2011)	OSA effects on HRV	Spectral and temporal analysis	Wakefulness	No differences between OSA and controls in the frequency domain. Only pNN50 lower in OSA
Kontos et al. (2020)	HRV segments analysis	Spectral and temporal analysis	N1, N2, N3, REM	In healthy children, decreased mNN and HF, and increased LF from NREM to REM
Vlahandonis et al. (2014)	OSA effects on HRV	Spectral analysis	LS, SWS, REM	In all groups, increase in LF and LF/HF, and decrease in HF from NREM to REM. Also, reduced LF, HF and increase LF/HF four years later
Nisbet et al. (2013)	OSA effects on HRV	Spectral and temporal analysis	W, LS, SWS, REM	Absence of differences in LF between groups, suggesting no autonomic dysfunction
Wu et al. (2022)	OSA effects on HRV	Spectral and temporal analysis	W, N1, N2, N3, REM	In all sleep stages, higher mHR and LF/HF for OSA children
Martín-Montero et al. (2021a)	OSA effects on HRV	OSA-specific spectral analysis	———	Three novel OSA-specific frequency bands (BW1. BW2 and BWRes) provided improved characterization of pediatric OSA
Martín-Montero et al. (2021b)	OSA effects on HRV	Bispectral analysis	———	Two bispectral features subsets with complementary information about alterations in non-Gaussianity, nonlinearity and irregularity HRV behavior due to OSA
Martín-Montero et al. (2023)	OSA effects on HRV segments	Spectral and temporal analysis	W, NREM, REM	Basal sympathetic activity during REM mask sympathetic activation during apneic events, which are better differentiated during NREM

OSA: Obstructive sleep apnea, W: wake, LS: light sleep, SWS: slow wave sleep, REM: rapid eye movement, LF: low frequency, HF: high frequency, HR: heart rate, PS: primary snorers, SDB: sleep-disordered breathing, RMSSD: root mean square of beat-to-beat differences, SDNN: standard deviation of NN intervals, NREM: non-rapid eye movement, ANS: autonomic nervous system, DAP: decreases in the amplitude fluctuations of PPG, pNN50: percentage of differences between adjacent NN intervals that are over50 msec, VLF: very low frequency.

school children. Notwithstanding, we found strong evidence of autonomic dysfunction in the presence of OSA, with apneic events modulating HRV activity. Therefore, the exclusion of events from their analysis may explain the lack of differences, thus supporting our approach in Martín-Montero et al. (2023), where we evaluated the appearance of apneic events to assess alterations in HRV throughout the night. On the other hand, there are other studies in the literature that included apneic events but combined segments with and without them in their analysis to evaluate HRV activity across sleep stages. Despite this difference, some similarities have been found. Baharav et al. (1999) examined HRV across sleep stages in healthy and OSA children, reporting greater RP_{LF} and LF/HF in all sleep stages, aligning with our results for these features as the presence of apneic events increased. Horne et al. (2018) informed elevated HR in overweight OSA children compared to healthy children, as well as reduced RP_{HF} in normal weight primary snorers within NREM. Similarly, they found greater RP_{LF} in OSA children compared to overweight primary snorers. Considering that the presence of apneic events in the HRV segments analyzed for healthy children and primary snorers is significantly low, these results align with the trends observed in our study for those parameters as severity of the segments increased. Finally, in a recent study conducted by Wu et al. (2022) they obtained similar results to ours when studying HRV segments across sleep stages in healthy and OSA children, reporting increased mHR and LF/HF for OSA children in all sleep stages.

6.3.2 Evaluation of pediatric OSA treatment effects on HRV

The design of the CHAT sleep study provided a suitable framework for numerous researchers to investigate the effects of pediatric OSA treatment in a broad set of biomedical signals, cardiovascular measures, and so on. However, prior to its conception, several studies also examined the impact of pediatric OSA treatment on HRV signal, primarily focusing on AT. Table 6.3 presents the main findings achieved in these previous research works. In a prospective trial, Şaylan et al. (2011) assessed HRV differences between 15 OSA children who underwent AT and 15 healthy children, revealing altered HRV in OSA children before AT that remained unchanged after the surgical procedure. In contrast to these findings, Kaditis et al. (2011) reported a reduction in RP_{LF}, mainly representing a decrease in sympathetic tone, when comparing HRV activity before and after AT in 21 OSA children. In disagreement with the previous two studies, Muzumdar et al. (2011), who analyzed 18 OSA children, observed an increase in RP_{LF} and RP_{HF}

Table 6.3. Methodological summary of the state-of-the-art research works focused on the evaluation of the effects of pediatric OSAS treatment on HRV, along with their relevant findings.

Study	Aim	Methodology	Relevant findings
Şaylan et al. (2011)	AT effects on HRV	Spectral analysis	All spectral featured in OSA children altered before surgical process. Those parameters remained altered following AT
Kaditis et al. (2011)	AT effects on HRV	Spectral analysis and BNP levels	Increased mRR intervals and decreased BNP levels after AT
Muzumdar et al. (2011)	AT effects on HRV	Spectral and temporal analysis	Decrease in HR and LF/HF following AT, reflecting a decline in the proportion of sympathetic activity with OSA improvement
Kirk et al. (2020)	Noninvasive ventilation effects on HRV	Spectral and temporal analysis	12-months noninvasive ventilation resulted in decrease in HR and no changes in HRV features
Isaiah et al. (2020)	OSA treatment effects on HRV	Spectral and temporal analysis	Changes founded in classic HRV features were not causally attributable to treatment effects
Martín-Montero et al. (2022)	OSA treatment effects on HRV	Spectral and temporal analysis	OSA treatment affects HRV through changes in severity and OSA resolution, mainly in OSA-specific BW2 frequency band

AT: adenotonsillectomy, OSA: obstructive sleep apnea, BNP: brain natriuretic peptide blood, mRR: mean of RR intervals, HR: heart rate, LF/HF: low frequency/high frequency ratio.

after AT, being greater in HF as also observed a decrease in LF/HF ratio. A more recent study conducted by Kirk et al. (2020) in 12 children with obesity also reported changes in LF/HF ratio following treatment with noninvasive ventilation, but only when assessing HRV around arousal events, with no overall changes in HRV observed 12 months after treatment application. As can be observed, there are discrepancies in the existing literature regarding the effects of pediatric OSA treatment on HRV activity. Among all the previously reported results in classic HRV parameters, our findings are only consistent with the observed increase in LF activity after surgical treatment reported by Muzumdar et al. (2011) (see Supplemental Figure S4 from Martín-Montero et al. (2023)). Nevertheless, it is important to note that there are several methodological differences between those studies and our research conducted in Martín-Montero et al. (2023). Notably, our study analyzed a significantly larger population from a multicenter database, thereby enhancing the robustness and generalizability of our findings.

Finally, it is worth comparing our study with the previous work conducted by Isaiah et al. (2020), as they performed a similar CMA to assess the effects of pediatric OSA treatment on classical HRV features using the CHAT database. Interestingly, they reported no treatment effects on classical HRV features, regard-

less of whether early AT or WWSC was performed. Even though we observed that the main treatment effects are reflected through the novel OSA-specific frequency band BW2, we also reported treatment influence in the same HRV classic parameters that Isaiah et al. (2020) evaluated. Upon comparing both approaches, we believe that these differences may be attributed to their interpretation of the results using the proportion mediated parameter instead of the ACME. The proportion mediated parameter allows for the assessment of the influence of ACME on the averaged total effect (Imai et al., 2010a). However, despite the potential utility of proportion mediated, Figure 4.4 shows that if CMA results in ACME and ADE in opposite directions, it could lead to non-significant proportion mediated parameter, even if the ACME is statistically significant. Thus, evaluating the proportion mediated parameter in isolation can mask ACME. Consequently, in Martín-Montero et al. (2023), we re-evaluated the results reported by Isaiah et al. (2020) for the classical HRV frequency measures, focusing on ACME effects. Our re-evaluation revealed that treatment also influences classical HRV features through some of the mediators considered (see Table 5.5).

6.3.3 Comparison of machine-learning approaches

As mentioned in Section 1.5.3, some studies have explored the information derived from the cardiovascular system to understand the impact of pediatric OSA on the ANS following different approaches. Table 6.4 presents a summary of the highest diagnostic performance achieved in these studies, as well as the diagnostic yield obtained by our LDA model utilizing OSA-specific HRV features from Martín-Montero et al. (2021b), the two MLP models constructed with HRV bispectral features from Martín-Montero et al. (2021a), and the diagnostic performance of the LSBoost model developed in Martín-Montero et al. (2023).

Shouldice et al. (2004) performed apneic event detection to classify HRV 1-minute segments as apneic or non-apneic, followed by the assignation of OSA condition to a children if over 12.5 apneic segments/hour were detected. When classifying 25 children from the test set using an AHI cutoff of 1 e/h, they reported Acc = 84%, Se = 85.7%, PPV = 85.7%, Sp = 81.8%, and NPV = 81.8%. Comparing this performance with our highest results obtained at the 1 e/h threshold, it can be observed that our LSBoost model outperformed these outcomes in terms of Se (90.76%) and PPV (86.26%). This reinforces the notion that our LSBoost model would be clinically valuable for detecting OSA presence in its lowest severity degree. Three studies conducted by the same research group (Gil et al.,

Table 6.4. Highest diagnostic performance achieved in the state-of-the-art research works that have focused on the automated diagnosis of pediatric OSA using cardiovascular signals.

Study	Model	#Total/#Test	AHI	Se	Sp	Acc	PPV	NPV	AUC
Shouldice et al. (2004)	QDA (HRV)	50/25	1	85.70	81.80	84.00	85.70	81.80	-
Gil et al. (2009)	LDA (PPG + HRV)	21/21	<5 control >18 OSA	87.50	71.40	80.00	-	-	-
Gil et al. (2010)	LDA (PPG + PTTV)	21/21	<5 control >18 OSA	75.00	85.70	80.00	-	-	-
Lázaro et al. (2014)	LDA (PPG + PRV)	21/21	<5 control >18 OSA	100	71.40	86.70	-	-	-
Dehkordi et al. (2016)	LASSO (PRV)	146/146	5	76.00	68.00	71.00	-	-	0.780
Martín-Montero et al. (2021b)	$\mathbf{LDA}_{Specific}$	1738/757	1	37.70	80.10	52.00	78.80	39.70	0.597
			5	48.20	80.80	76.00	30.30	90.00	0.696
			10	62.80	84.30	82.30	28.50	95.70	0.774
Martín-Montero et al. (2021a)	$\mathbf{MLP1}_{Classic}$	1738/757	1	52.30	59.40	54.70	71.58	38.87	0.600
	$\mathbf{MLP5}_{Classic}$		5	50.90	86.20	81.00	39.04	91.00	0.774
	$\mathbf{MLP10}_{Classic}$		10	43.50	96.50	91.70	55.56	94.45	0.847
Martín-Montero et al. (2021a)	$\mathbf{MLP1}_{Specific}$	1738/757	1	76.30	38.30	63.40	70.74	45.16	0.627
	$\mathbf{MLP5}_{Specific}$		5	62.50	84.20	81.00	40.70	92.82	0.791
	$\mathbf{MLP10}_{Specific}$		10	66.70	91.60	89.30	44.23	96.48	0.841
Martín-Montero et al. (2023)	LSBoost	1018/296	1	90.76	23.40	80.07	86.26	32.35	0.651
			5	66.67	61.17	63.18	49.66	76.16	0.677
			10	40.00	92.30	84.12	47.37	89.53	0.742

AHI: apnea-hypopnea index, Se: sensitivity (%), Sp: specificity (%), Acc: accuracy (%), PPV: positive predictive value (%), NPV: negative predictive value (%), AUC: area under receiver-operating characteristic curve, QDA: quadratic discriminant analysis, HRV: heart rate variability, LDA: linear discriminant analysis, PPG: photoplethysmography, PTTV: pulse transit time variability, PRV: pulse rate variability, LASSO: least absolute shrinkage operating characteristic curves.

2010, 2009; Lázaro et al., 2014) also focused on detecting apneic event presence to consequently report a per-subject classification extracting HRV features from decreases in amplitude fluctuations of PPG. These studies included 21 subjects, 10 OSA children (AHI $>$ 18 e/h) and 11 controls (AHI $<$ 5 e/h). When considering HRV information, they reported *Acc* ranging from 73.3% to 80%, *Se* from 62.5% to 87.5%, and *Sp* from 71.45% to 85.7%. Despite the lower *Se* identified in our research, Table 6.4 shows that the MLP models constructed with bispectral features achieved higher *Acc* and, more interesting, these models are particularly effective in ruling out children without severe OSA. It is evidenced by the higher *Sp* achieved at a more stringent severity threshold (10 e/h). This capability has the potential to minimize the number of individuals referred for further evaluation, thereby reducing waiting lists. Lázaro et al. (2014) also noted that when considering PRV information rather than HRV, combined with decreased amplitude fluctuation events, the *Se* and *Acc* increased to 100% and 86.7%, respectively (see Table 6.4). Notwithstanding, there exist a marked difference between all these previous works and our research, mainly in the child population included, but also in the criterion followed to establish OSA severity, hindering further com-

prehensive comparisons.

Dehkordi et al. (2016). conducted a study in which they characterized the PRV signal using temporal and frequency approaches, as well as detrended fluctuation analysis, for diagnosing OSA in a population of 146 children. By combining this information, they directly reported per-subject classification using a cutoff of 5 e/h for OSA severity, reaching an Acc of 71%, Se of 76%, Sp of 68%, and AUC of 78%. From Table 6.4, it can be observed that our highest performance at this cutoff (MLP5$_{Specific}$ model) surpassed their Acc, AUC, and Sp, thereby obtaining better diagnostic yield in overall terms, and making it more suitable for clinical implementation. Lastly, in the comparison with previous studies that performed automated pediatric OSA diagnosis through cardiovascular signal analysis, Cohen and de Chazal (2015) conducted an automated classification algorithm focusing exclusively on analyzing HRV. Nevertheless, their work only provided classification results for apneic event detection without further reporting per-subject classification. Consequently, comparing their results with the OSA diagnosis performance obtained in this Doctoral Thesis is not meaningful.

Regarding pediatric OSA diagnosis, the results compared so far in this section have mainly focused on studies that used cardiovascular signals exclusively. As it was stated in the introduction section, the systematic review performed by Gutiérrez-Tobal et al. (2022) pointed out that the majority of research works that focused on machine-learning approaches to diagnose pediatric OSA were based on SpO2 signals. In this systematic review, that did not include any HRV study, a meta-analysis was performed to gather Se and Sp metrics from 19 research works that met their inclusion criteria (Gutiérrez-Tobal et al., 2022). While caution must be exercised when comparing results obtained from different biological sources, a comparison can still be made between these findings and the performance metrics reported in this Doctoral Thesis. Gutiérrez-Tobal et al. (2022) provided diagnostic performance metrics for pediatric OSA at three AHI cutoffs: 1, 5, and 10 e/h, reporting Se values of 84.9%, 71.4%, and 65.2% for the respective cutoffs, while Sp values were reported as 49.9%, 83.2%, and 93.1%. Comparison with our models, as shown in Table 6.4, reveals that our LSBoost model achieved higher Se at the 1 e/h cutoff, and the LDA$_{Specific}$ and MLP1$_{Classic}$ models achieved higher Sp. For the 5 and 10 e/h cutoff, both MLP models performed within the same range of Sp as the meta-analysis, as also did our LSBoost models for the 10 e/h cutoff. Furthermore, $MLP10_{Specific}$ also achieved similar Se in 10 e/h. Therefore, the overall similar diagnostic yield achieved in this Doctoral Thesis, particularly with the $MLP10_{Specific}$ model, underscores the clinical utility of the HRV charac-

terization conducted to help in pediatric OSA diagnosis.

Regarding sleep stage classification, this task in children based on cardiac measures has been rarely explored. To the best of our knowledge, the AdaBoost model developed in Martín-Montero et al. (2023) was the first instance where HRV features were utilized for sleep stage classification in the context of pediatric OSA. However, there have been some prior studies that focused on sleep stage classification in healthy children. Haddad et al. (1987) examined 9 infants aged 1 to 4 months to differentiate between REM and quiet sleep using cardiorespiratory measures. They found that alterations in respiratory cycle time had the highest classification ability, achieving a *Se* of 93% for quiet sleep and nearly 99% for REM. Notwithstanding, they did not specifically report the performance of cardiac measures alone. In a study by Harper et al. (1987), cardiorespiratory measures were used to classify sleep stages in 25 infants over 6 months of age. They classified 1-minute segments into wake, quiet sleep, and REM, reporting an overall accuracy (Acc_3) of 82% using cardiac measures alone, which increased to 84.8% when combining seven cardiac and respiratory measures. Lastly, Lewicke et al. (2008) developed various machine-learning models for the automatic classification of sleep versus wake using 30-second HRV segments in a cohort of 190 infants. They reported an *Acc* of around 78%, which improved to 85-87% when excluding 30% of segments that hindered classification. Unfortunately, due to the differences in cohort characteristics, sample size, and the specific sleep stages considered, it is virtually impossible to directly compare these findings with our results regarding sleep stage classification.

6.4 Limitations of the study

In the course of this Doctoral Thesis, we have demonstrated the usefulness of the approaches followed to characterize alterations in the ANS in presence of pediatric OSA using HRV analysis. Notwithstanding, there exist certain limitations that deserve to be mentioned.

The first of these limitations pertains to the size of the database, specifically concerning the distribution of the population across different severity groups of pediatric OSA. Although the research conducted involved a considerably large population, ranging from 404 to 1738 children, there is an imbalance in the number of children included in each severity group. Similarly, when considering different types of HRV 10-minute segments, despite the great number of segments from each type, there was a predominance of NREM and non-apneic segments.

Tests carried out for balancing the dataset did not change the performance of the models. However, to enhance the robustness and generalizability of our findings, future studies should aim to increase the sample size, ensuring a more balanced distribution of severity groups, or including a greater representation of REM, wake, and apneic segments.

Likewise, the studies conducted did not assess the population stratification by factors such as age, sex, BMI, or other relevant variables, which could have unveiled OSA subgroups where HRV analysis could be more specific. Nonetheless, the definition of the OSA-specific frequency band BWRes considered the age-related variations that may affect HRV and respiratory peaks. Additionally, all potential covariates that could influence HRV activity were controlled for to the best extent possible, considering them in partial correlations and incorporating them into the CMA.

Other limitation of this study concerns to the diagnostic ability achieved by the developed models. It is important to note that, in certain instances of this research, the objective was not to maximize diagnostic performance, but rather to compare the clinical usefulness of different characterization approaches employed. Nevertheless, when comparing with state-of-the-art studies using HRV signals, our models demonstrated similar or higher diagnostic performance. Specifically, the MLP models exhibited high diagnostic yield in the 5 e/h and 10 e/h AHI cutoffs, and our work was the only one conducting binary classification in the 10 e/h AHI cutoff using cardiovascular signals. However, it is worth noting that the diagnostic yield obtained in this Doctoral Thesis is surpassed by considering other biomedical signals obtained during PSG. Furthermore, the outcomes achieved, particularly in the lower severity thresholds, are not yet sufficient for widespread diagnostic application in real clinical settings. Therefore, future studies should focus on implementing more sophisticated predictive models aimed at further enhancing diagnostic performance.

Concerning the applicability of our novel pediatric HRV OSA-specific frequency bands to an adult population, we have not conducted testing for adult suitability. Nevertheless, when comparing the outcomes of our study with a previous adult-focused investigation (Gutiérrez-Tobal et al., 2015), it becomes apparent that dissimilarities in the occurrence and duration of apneic events, coupled with variations in respiratory patterns, hinder the direct implementation of our pediatric HRV OSA-specific bands in adults. Alternatively, should there be an interest in employing OSA-specific HRV frequency bands for adults, future studies could replicate our methodology, defining novel adults OSA-specific HRV fre-

quency bands that would be able to adapt to these differences between populations.

There are also a limitation related to the feature-extraction techniques used in this Doctoral Thesis. In recent years, nonlinear techniques based on the chaos theory have been evaluated in the context of pediatric OSA. These techniques have been found to provide additional insights into the dynamics of other biomedical signals, such as SpO_2 or AF (Álvarez et al., 2017, 2013; Barroso-García et al., 2020; Crespo et al., 2018; Garde et al., 2014; Hornero et al., 2017; Jiménez-García et al., 2020). While the feature-extraction techniques used in the time and spectral domain in this study have demonstrated their utility, the nonlinear behavior of HRV in the presence of pediatric OSA through nonlinear parameters has not been evaluated. Therefore, future studies should consider the inclusion of nonlinear metrics, such as central tendency measure, Lempel-Ziv complexity, or sample entropy, to assess their usefulness in pediatric OSA and their complementarity with the HRV features examined in this Doctoral Thesis.

Also aligning with the search of more complex predictive models to evaluate clinical utility of HRV, in recent years there has been growing interest in using deep-learning techniques to tackle healthcare challenges, including the automatic diagnosis of pediatric OSA, due to their promising results and potential utility (Gutiérrez-Tobal et al., 2022). However, in this Doctoral Thesis, we focused solely on assessing machine-learning approaches to evaluate diagnostic performance, which can be considered a limitation of this study. It is important to note that one of the main drawbacks of deep-learning techniques is their lack of interpretability, with the perception of 'black box' being even more pronounced compared to traditional machine-learning methodologies. Since one of the main aims of this thesis was to investigate the ANS alterations associated with pediatric OSA, using a deep-learning approach would have hindered our ability to gain insights into these alterations. Nevertheless, for future research aiming to enhance the diagnostic yield of pediatric OSA, the inclusion of deep-learning techniques might be considered, coupled with more powerful XAI techniques to improve our understanding of ANS alterations in pediatric OSA.

Finally, there are also limitations associated with the acquisition of ECG signals and the diagnosis of OSA. In this study, all children underwent laboratory-based PSG to establish OSA diagnosis and severity. While it is considered the gold standard for OSA diagnosis, it has certain drawbacks. It involves the use of a large number of sensors and requires children to spend a night in a laboratory, away from their home and parents. This intrusive nature of the procedure can

lead to changes in sleep architecture and may not fully represent the usual sleep pattern of the children (Álvarez et al., 2017; Kheirandish-Gozal, 2010). Home-based respiratory polygraphy has been shown to be a feasible alternative for diagnosing pediatric OSA (Alonso-Álvarez et al., 2015; Chiner et al., 2020). Therefore, to ensure the reproducibility of the findings in this Doctoral Thesis, future studies should include the analysis of at-home OSA databases in children.

Chapter 7

Conclusions

The four studies conducted during the course of this Doctoral Thesis share a common focus: the comprehensive characterization of ANS alterations resulting from pediatric OSA from HRV analysis. To achieve this objective, automatic signal processing methodologies based in novel HRV features were designed, implemented, and assessed. Specifically, OSA-related HRV features in the spectral and bispectral domain were proposed to investigate cardiac behavior in the presence of OSA and assess the effects of OSA treatment. Additionally, the clinical utility of these new HRV features was evaluated by using machine-learning models for the automatic diagnosis of pediatric OSA and the classification of sleep stages. The findings demonstrated that the newly proposed HRV features reflect ANS alterations associated with pediatric OSA, with these features being useful to evaluate the effects of OSA treatment and to aid in the diagnosis of the disease.

Section 7.1 of this chapter highlights the primary contributions made in the research. Subsequently, Section 7.2 presents the conclusions drawn from the findings of this Doctoral Thesis. Lastly, in Section 7.3, future research lines are stated.

7.1 Contributions

The primary contributions offered by the compendium of publications presented in this Doctoral Thesis are detailed below:

1) **Definition of three novel pediatric OSA-specific spectral bands: BW1 (0.001-0.005 Hz), BW2 (0.028 - 0.074 Hz) and BWRes (0.04 Hz around maximum value within HF).** These frequency ranges have been identified and thoroughly characterized, demonstrating their potential usefulness in as-

sessing HRV alterations and treatment outcomes. These new approaches also outperformed conventional methods in diagnosing pediatric OSA.

2) **Characterization of HRV through bispectral features in the context of pediatric OSA.** To the best of our knowledge, this is the first research to use bispectral HRV analysis for the characterization and diagnosis of pediatric OSA. It enabled the extraction of two distinct feature subsets, one from the regions bounded by the classic HRV frequency ranges, and the other one from the OSA-specific bispectral regions. Those feature subsets provided complementary and non-redundant information about the changes in the non-Gaussianity, nonlinearity, and irregularity patterns of HRV in the presence of OSA. Additionally, a novel bispectral parameter, RP_{Diag}, has been introduced.

3) **Establishment of causality relationships between changes in HRV metrics and OSA treatment effects.** As far as we are concerned, this is the first study to attribute changes in HRV features directly to treatment effects, as measured by alterations in polysomnographic indices or the resolution of the disease. The findings from this analysis have led to the identification of BW2 as a potential biomarker of pediatric OSA resolution.

4) **HRV segment characterization considering sleep stages and presence of apneic events.** To the best of our knowledge, this is the first study to conduct a segment-level analysis of HRV behavior taking into account sleep stages and apneic events in pediatric OSA context. The evolution of novel OSA-specific HRV parameters through sleep stages have also been introduced in this research.

5) **Evaluation of novel machine-learning models to diagnose pediatric OSA.** Although the machine-learning methodologies applied in this Doctoral Thesis are not new, the models developed, as well as its application in the context of pediatric OSA are a novelty. These models utilize various types of novel HRV features, including spectral features for binary classification (LDA), bispectral features for binary classification (two MLP models), and classic as well as OSA-specific HRV features from 10-minute segments for regression (LSBoost).

6) **Evaluation of a machine-learning model to classify sleep stages using HRV features in pediatric OSA population.** Although it was previously

evaluated in healthy population, to the best of our awareness, this is the first time that a machine-learning model (AdaBoost) has been developed solely based on HRV features for the purpose of automatically classifying sleep stages in pediatric OSA context.

7) **Evaluation of XAI techniques through feature importance metrics in the ensemble-learning model developed.** To the best of our knowledge, this is also the first time that feature importance has been assessed in models aimed at diagnosing pediatric OSA (relative importance with the LSBoost model), and classifying sleep stages (LIME with the AdaBoost model). This evaluation has revealed that two of the novel HRV spectral features (RP_{BW2} and RP_{BWRes}) hold the highest importance in these two tasks, increasing the explainability of the ensemble-learning models developed.

7.2 Conclusions

In the previous sections, we have performed a thorough evaluation and discussion regarding the outcomes reached in this Doctoral Thesis. Thus, it facilitates the extraction of the principal conclusions, which are detailed here:

1) Novel pediatric OSA-specific HRV frequency band BW1 (0.001-0.005 Hz) is associated with macro sleep disruptions resulting from OSA, measured as WASO time and the number of awakenings during night. The treatment of this condition induces changes within this spectral range, which can be attributed to alterations in the OAHI.

2) Novel pediatric OSA-specific HRV frequency band BW2 (0.028-0.074 Hz) is strongly associated with the duration of apneic events. Treatment interventions for OSA result in changes within this frequency range, primarily driven by alterations in OSA severity, with decreases in the activity of this band reflecting a decrease of the severity. Furthermore, it was the only spectral band where changes were causally mediated by OSA resolution and allows to differentiate HRV activity between children with and without OSA resolution. Accordingly, it has been proposed as a potential biomarker of pediatric OSA resolution. Moreover, when considering only the HRV signal, this OSA-specific frequency band has proven to be highly valuable in automated diagnosis of pediatric OSA.

3) Novel individual pediatric OSA-specific HRV frequency band BWRes (0.04

Hz around individual respiratory peak) is related to oxygen desaturations and micro sleep disruptions. Treatment interventions for OSA induce changes within this frequency band, driven by variations in arousal indices. Moreover, this OSA-specific frequency band has proven to be highly useful in automatically classifying sleep stages in pediatric OSA using the HRV signal.

4) The findings from the bispectral analysis revealed that pediatric OSA is associated with increased linearity and Gaussianity of the HRV signal at very low frequencies. Additionally, the occurrence of apneic events induces a shift in the coupling pattern of overnight HRV, transitioning from frequencies below 0.02 Hz to higher frequencies, related to the duration of those events. Furthermore, the presence of OSA leads to a decrease in HRV irregularity at frequency components associated with BW2 and LF frequency ranges, while there is an increase in irregularity in frequency components associated with respiration.

5) Pediatric OSA leads to elevated sympathetic activity, which is evident in HRV patterns during the night. It is anticipated that sympathetic activity would be higher during the REM sleep stage due to the increased occurrence of apneic events in this stage. Surprisingly, the underlying sympathetic excitation resulting from apneic events appears to be masked by the basal sympathetic activation observed during REM sleep. Therefore, it becomes easier to discern these OSA-related alterations in HRV when analyzing data from NREM sleep stage. Consequently, we emphasize the significance of considering the presence of apneic events together with sleep stages when assessing HRV changes in the context of pediatric OSA.

6) The HRV analysis methods developed in this Doctoral Thesis have demonstrated their utility in aiding to the diagnosis of pediatric OSA. The MLP models, constructed using optimal bispectral feature subsets, achieved the highest performance for the AHI severity thresholds of 5 ($Acc = 81\%$, $AUC = 0.791$) and 10 e/h ($Acc = 91.70\%$, $AUC = 0.847$). The LSBoost model, on the other hand, exhibited the highest performance for the AHI severity cutoff of 1 e/h ($Acc = 80.07\%$, $AUC = 0.651$), emphasizing the effectiveness of segment-level characterization in detecting the presence of OSA, even in its mildest form.

Based on the aforementioned considerations, it can be concluded that the

characterization of HRV using the novel indices developed during the realization of the present Doctoral Thesis, and computed across the new pediatric OSA-specific frequency bands, enable a more specific assessment of the ANS alterations with OSA. In addition, the inclusion of these indices in pediatric OSA studies have the potential to contribute to the simplification of the diagnosis of the disease.

7.3 Future research lines

Throughout the completion of this Doctoral Thesis, several issues have emerged that warrant further investigation in order to enhance our understanding and complementing the work presented here. Next, the following areas are identified as potential directions for future research to complement and expand upon our current findings:

1) The evaluation of the automatic signal processing methods stratifying by subgroups of children with diverse clinical and demographic characteristics, such as age, BMI, and gender. This approach will enable the identification of specific subgroups where optimal results can be attained.

2) To enhance the reliability and generalizability of our results, it is suggested to validate the developed methodology using ECG signals acquired through portable devices in an at-home approach. This validation will enable us to assess the performance of our methodology in a real-world setting, outside of a laboratory environment, thereby increasing the applicability and practicality of our findings.

3) Integration of the HRV signal with feature-extraction techniques from the EDR. The EDR signal offers a non-invasive means of monitoring respiration activity directly from the ECG, without the need for dedicated respiratory signals. By combining the information extracted from HRV with that obtained from the EDR, we can assess the complementary nature of these signals and potentially develop a unified approach for diagnosing pediatric OSA. This is particularly advantageous as both signals are derived from the ECG, simplifying the diagnostic process.

4) Another important future area for investigation is the assessment of the clinical utility of the HRV characterization conducted in this Doctoral Thesis for evaluating cardiovascular risks. This evaluation holds significance due

to the established association between cardiovascular risk and OSA. By examining the HRV features proposed in this research, we can gain insights into the cardiovascular effects of OSA and explore potential strategies to mitigate these effects.

5) The marked differences in the presence and duration of apneic events, along with the differences in the respiratory patterns between children and adult make inefficiently the application in adult population of the pediatric OSA-specific frequency bands developed during the realization of this Doctoral Thesis. However, the replication of the methodology followed to the identification of our novel bands of interest using an adult cohort would allow to unravel adult HRV OSA-specific frequency bands to gain insights into the characterization of the alterations in the ANS due to OSA in adults.

6) The inclusion of non-linear parameters would be valuable to explore in future research. These types of features have proven their usefulness in pediatric OSA context when applied to other biological signals. Accordingly, the implementation of features such as central tendency measure, Lempel-Ziv complexity, or sample entropy, would complement the feature-extraction techniques implemented in this research, while providing new insights into the ANS behavior in the context of pediatric OSA.

7) Finally, the utilization of novel deep-learning techniques, which are becoming increasingly important in the field of healthcare, can enhance the diagnostic capabilities of HRV. Deep-learning methods are typically applied to raw data, and their integration with advanced XAI techniques could be crucial for obtaining deeper insights into the impact of OSA on cardiac behavior. The incorporation of these approaches also would improve and maximize the diagnostic performance reached for pediatric OSA.

In conclusion, this Doctoral Thesis focused on providing a comprehensive understanding of the ANS behavior in the context of pediatric OSA through HRV analysis. The applied methodology enabled us to ascertain that the assessment of HRV signals using novel pediatric OSA-specific spectral bands and features presented in this research offers a more specific evaluation of ANS alterations. Consequently, it has the potential to streamline the diagnosis of OSA in children.

Chapter 8

Papers included in the compendium of publications

Martín-Montero et al. (2021b): Heart rate variability spectrum characteristics in children with sleep apnea.

DOI: 10.1038/s41390-020-01138-2

Adrián Martín-Montero, Gonzalo C. Gutiérrez-Tobal, Leila Kheirandish-Gozal, Jorge Jiménez-García, Daniel Álvarez, Félix del Campo, David Gozal, and Roberto Hornero. *Pediatric Research*, vol. 89 (7), p. 1771-1779, 2021. Impact factor in 2021: 3.953, Q1 in "PEDIATRICS" (JCR-WOS).

Abstract: Background: Classic spectral analysis of heart rate variability (HRV) in pediatric sleep apnea-hypopnea syndrome (SAHS) traditionally evaluates the very low frequency (VLF: 0-0.04 Hz), low frequency (LF: 0.04–0.15 Hz), and high frequency (HF: 0.15–0.40 Hz) bands. However, specific SAHS-related frequency bands have not been explored. Methods: One thousand seven hundred and thirty-eight HRV overnight recordings from two pediatric databases (0–13 years) were evaluated. The first one (981 children) served as training set to define new HRV pediatric SAHS-related frequency bands. The associated relative power (RP) were computed in the test set, the Childhood Adenotonsillectomy Trial database (CHAT, 757 children). Their relationships with polysomnographic variables and diagnostic ability were assessed. Results: Two new specific spectral bands of pediatric SAHS within 0–0.15 Hz were related to duration of apneic events, number of awakenings, and wakefulness after sleep onset (WASO), while an adaptive individual-specific new band from HF was related to oxyhemoglobin desaturations, arousals, and WASO. Furthermore, these new spectral bands showed improved diagnostic ability than classic HRV. Conclusions: Novel spectral bands provide improved characterization of pediatric SAHS. These findings may pioneer a better understanding of the effects of SAHS on cardiac function and potentially serve as detection biomarkers.

Martín-Montero et al. (2021a): Bispectral Analysis of Heart Rate Variability to Characterize and Help Diagnose Pediatric Sleep Apnea.

DOI: 10.3390/e23081016

Adrián Martín-Montero, Gonzalo C. Gutiérrez-Tobal, David Gozal, Verónica Barroso-García, Daniel Álvarez, Félix del Campo, Leila Kheirandish-Gozal, and Roberto Hornero. *Entropy*, vol. 23 (8), p. 1016, 2021. Impact factor in 2021: 2.738, Q2 in "PHYSICS, MULTIDISCIPLINARY" (JCR-WOS).

Abstract: Pediatric obstructive sleep apnea (OSA) is a breathing disorder that alters heart rate variability (HRV) dynamics during sleep. HRV in children is commonly assessed through conventional spectral analysis. However, bispectral analysis provides both linearity and stationarity information and has not been applied to the assessment of HRV in pediatric OSA. Here, this work aimed to assess HRV using bispectral analysis in children with OSA for signal characterization and diagnostic purposes in two large pediatric databases (0–13 years). The first database (training set) was composed of 981 overnight ECG recordings obtained during polysomnography. The second database (test set) was a subset of the Childhood Adenotonsillectomy Trial database (757 children). We characterized three bispectral regions based on the classic HRV frequency ranges (very low frequency: 0–0.04 Hz; low frequency: 0.04–0.15 Hz; and high frequency: 0.15–0.40 Hz), as well as three OSA-specific frequency ranges obtained in recent studies (BW1: 0.001–0.005 Hz; BW2: 0.028–0.074 Hz; BWRes: a subject-adaptive respiratory region). In each region, up to 14 bispectral features were computed. The fast correlation-based filter was applied to the features obtained from the classic and OSA-specific regions, showing complementary information regarding OSA alterations in HRV. This information was then used to train multi-layer perceptron (MLP) neural networks aimed at automatically detecting pediatric OSA using three clinically defined severity classifiers. Both classic and OSA-specific MLP models showed high and similar accuracy (Acc) and areas under the receiver operating characteristic curve (AUCs) for moderate (classic regions: Acc = 81.0%, AUC = 0.774; OSA-specific regions: Acc = 81.0%, AUC = 0.791) and severe (classic regions: Acc = 91.7%, AUC = 0.847; OSA-specific regions: Acc = 89.3%, AUC = 0.841) OSA levels. Thus, the current findings highlight the usefulness of bispectral analysis on HRV to characterize and diagnose pediatric OSA.

Martín-Montero et al. (2022): Heart rate variability as a potential biomarker of pediatric obstructive sleep apnea resolution.

DOI: 10.1093/sleep/zsab214

Adrián Martín-Montero, Gonzalo C. Gutiérrez-Tobal, Leila Kheirandish-Gozal, Fernando Vaquerizo-Villar, Daniel Álvarez, Félix del Campo, David Gozal, and Roberto Hornero. *SLEEP*, vol. 45 (2), p. zsab214, 2022. Impact factor in 2022: 5.599, Q1 in "CLINICAL NEUROLOGY" (JCR-WOS).

Abstract: Study Objectives: Pediatric obstructive sleep apnea (OSA) affects cardiac autonomic regulation, altering heart rate variability (HRV). Although changes in classical HRV parameters occur after OSA treatment, they have not been evaluated as reporters of OSA resolution. Specific frequency bands (named BW1, BW2, and BWRes) have been recently identified in OSA. We hypothesized that changes with treatment in these spectral bands can reliably identify changes in OSA severity and reflect OSA resolution. Methods: Four hundred and four OSA children (5–9.9 years) from the prospective Childhood Adenotonsillectomy Trial were included; 206 underwent early adenotonsillectomy (eAT), while 198 underwent watchful waiting with supportive care (WWSC). HRV changes from baseline to follow-up were computed for classical and OSA-related frequency bands. Causal mediation analysis was conducted to evaluate how treatment influences HRV through mediators such as OSA resolution and changes in disease severity. Disease resolution was initially assessed by considering only obstructive events, and was followed by adding central apneas to the analyses. Results: Treatment, regardless of eAT or WWSC, affects HRV activity, mainly in the specific frequency band BW2 (0.028–0.074 Hz). Furthermore, only changes in BW2 were specifically attributable to all OSA resolution mediators. HRV activity in BW2 also showed statistically significant differences between resolved and nonresolved OSA. Conclusions: OSA treatment affects HRV activity in terms of change in severity and disease resolution, especially in OSA-related BW2 frequency band. This band allowed to differentiate HRV activity between children with and without resolution, so we propose BW2 as potential biomarker of pediatric OSA resolution.

Martín-Montero et al. (2023): Pediatric sleep apnea: Characterization of apneic events and sleep stages using heart rate variability.

DOI: 10.1016/j.compbiomed.2023.106549

Adrián Martín-Montero, Pablo Armañac-Julián, Eduardo Gil, Leila Kheirandish-Gozal, Daniel Álvarez, Jesús Lázaro, Raquel Bailón, David Gozal, Pablo Laguna, Roberto Hornero and Gonzalo C. Gutiérrez-Tobal. *Computers in Biology and Medicine*, vol. 154, p. 106549, 2023. Impact factor in 2022 (last year available): 7.669, D1 in "MATHEMATICAL AND COMPUTATIONAL BIOLOGY" (JCR-WOS).

Abstract: Heart rate variability (HRV) is modulated by sleep stages and apneic events. Previous studies in children compared classical HRV parameters during sleep stages between obstructive sleep apnea (OSA) and controls. However, HRV-based characterization incorporating both sleep stages and apneic events has not been conducted. Furthermore, recently proposed novel HRV OSA-specific parameters have not been evaluated. Therefore, the aim of this study was to characterize and compare classic and pediatric OSA-specific HRV parameters while including both sleep stages and apneic events. A total of 1,610 electrocardiograms from the Childhood Adenotonsillectomy Trial (CHAT) database were split into 10-minute segments to extract HRV parameters. Segments were characterized and grouped by sleep stage (wake, W; non-rapid eye movement, NREMS; and REMS) and presence of apneic events (under 1 apneic event per segment, e/s; 1-5 e/s; 5-10 e/s; and over 10 e/s). NREMS showed significant changes in HRV parameters as apneic event frequency increased, which were less marked in REMS. In both NREMS and REMS, power in BW2, a pediatric OSA-specific frequency domain, allowed for the optimal differentiation among segments. Moreover, in the absence of apneic events, another defined band, BWRes, resulted in best differentiation between sleep stages. The clinical usefulness of segment-based HRV characterization was then confirmed by two ensemble-learning models aimed at estimating apnea-hypopnea index and classifying sleep stages, respectively. We surmise that basal sympathetic activity during REMS may mask apneic events-induced sympathetic excitation, thus highlighting the importance of incorporating sleep stages as well as apneic events when evaluating HRV in pediatric OSA.

Appendix A

Scientific achievements

A.1 Publications

A.1.1 Papers indexed in the Journal Citation Reports (JCR)

1. Pablo Armañac-Julián, **Adrián Martín-Montero**, Jesús Lázaro, Roberto Hornero, Pablo Laguna, Leila Kheirandish-Gozal, David Gozal, Eduardo Gil, Raquel Bailón, Gonzalo C. Gutiérrez-Tobal, "Persistent Sleep Disordered Breathing Independently Contributes to Metabolic Syndrome in Prepubertal Children", *Pediatric Pulmonology*, vol. 59 (1), January, 2024, pp. 111-120. DOI: 10.1002/ppul.26720. Impact factor in 2022 (last year available): 3.127, Q2 in "PEDIATRICS" (JCR-WOS).

2. Clara García-Vicente, Gonzalo C. Gutiérrez-Tobal, Jorge Jiménez-García, **Adrián Martín-Montero**, David Gozal, Roberto Hornero, "ECG-based Convolutional Neural Network in Pediatric Obstructive Sleep Apnea Diagnosis", *Computers in Biology and Medicine*, vol. 167, p.107628, December, 2023. DOI: 10.1016/j.compbiomed.2023.107628. Impact factor in 2022 (last year available): 7.669, D1 in "MATHEMATICAL AND COMPUTATIONAL BIOLOGY" (JCR-WOS).

3. **Adrián Martín-Montero**, Pablo Armañac-Julián, Eduardo Gil, Leila Kheirandish-Gozal, Daniel Álvarez, Jesús Lázaro, Raquel Bailón, David Gozal, Pablo Laguna, Roberto Hornero, Gonzalo C. Gutiérrez-Tobal, "Pediatric Sleep Apnea: Characterization of Apneic Events and Sleep Stages Using Heart Rate Variability", *Computers in Biology and Medicine*, vol. 154,

p.106549, March, 2023. DOI: 10.1016/j.compbiomed.2023.106549. Impact factor in 2022 (last year available): 7.669, D1 in "MATHEMATICAL AND COMPUTATIONAL BIOLOGY" (JCR-WOS).

4. **Adrián Martín-Montero**, Gonzalo C. Gutiérrez-Tobal, Leila Kheirandish-Gozal, Fernando Vaquerizo-Villar, Daniel Álvarez, Félix del Campo, David Gozal, Roberto Hornero, "Heart Rate Variability as a Potential Biomarker of Pediatric Obstructive Sleep Apnea Resolution", *Sleep*, vol. 45 (2), pp. zsab214, February, 2022. DOI: 10.1093/sleep/zsab214. Impact factor in 2022: 5.599, Q1 in "CLINICAL NEUROLOGY" (JCR-WOS).

5. Gonzalo C. Gutiérrez-Tobal, Javier Gomez-Pilar, Leila Kheirandish-Gozal, **Adrián Martín-Montero**, Jesús Poza, Daniel Álvarez, Félix del Campo, David Gozal, Roberto Hornero, "Pediatric Sleep Apnea: The Overnight Electroencephalogram as a Pehnotypic Biomarker", *Frontiers in Neuroscience*, vol. 15, pp. 644697, November, 2021. DOI: 10.3389/fnins.2021.644697. Impact factor in 2021: 5.152, Q2 in "NEUROSCIENCES" (JCR-WOS).

6. **Adrián Martín-Montero**, Gonzalo C. Gutiérrez-Tobal, David Gozal, Verónica Barroso-García, Daniel Álvarez, Félix del Campo, Leila Kheirandish-Gozal, Roberto Hornero, "Bispectral Analysis of Heart Rate Variability to Characterize and Help Diagnose Pediatric Sleep Apnea", *Entropy*, vol. 23 (8), pp. 1016, August, 2021. DOI: 10.3390/e23081016. Impact factor in 2021: 2.738, Q2 in "PHYSICS, MULTIDISCIPLINARY" (JCR-WOS).

7. **Adrián Martín-Montero**, Gonzalo C. Gutiérrez-Tobal, Leila Kheirandish-Gozal, Jorge Jiménez-García, Daniel Álvarez, Félix del Campo, David Gozal, Roberto Hornero, "Heart rate variability spectrum characteristics in children with sleep apnea", *Pediatric Research*, vol. 89 (7), pp.1771-1779, May, 2021. DOI: 10.1038/s41390-020-01138-2. Impact factor in 2021: 3.953, Q1 in "PEDIATRICS" (JCR-WOS).

8. Jorge Jiménez-García, Gonzalo C. Gutiérrez-Tobal, María García, Leila Kheirandish-Gozal, **Adrián Martín-Montero**, Daniel Álvarez, Félix del Campo, David Gozal, Roberto Hornero, "Assessment of Airflow and Oximetry Signals to Detect Pediatric Sleep Apnea-Hypopnea Syndrome Using AdaBoost", *Entropy*, vol. 22 (6), pp. 670, June, 2021. DOI:

10.3390/e22060670. Impact factor in 2020: 2.524, Q2 in "PHYSICS, MULTIDISCIPLINARY" (JCR-WOS).

A.1.2 International conferences

1. Pablo Armañac-Julián, **Adrián Martín-Montero**, Jesús Lázaro, Spyridon Kontaxis, Daniel Álvarez, David Gozal, Roberto Hornero, Pablo Laguna, Gonzalo C. Gutiérrez-Tobal, Raquel Bailón, Eduardo Gil, "Characterization of Changes in HRV Metrics During Sleep Apnea Episodes in Pediatric Patients", *12th Conference of The European Study Group on Cardiovascular Oscillations (ESGCO 2022)*, ISBN: 978-1-6654-8512-8, pp. 1-2, Štrbské pleso (Slovaquia), October 9 - October 12, 2022. DOI: 10.1109/ESGCO55423.2022.9931362.

2. **Adrián Martín-Montero**, Pablo Armañac-Julián, Leila Kheirandish-Gozal, Eduardo Gil, David Gozal, Roberto Hornero, Gonzalo C. Gutiérrez-Tobal, "Characterization of Paediatric Heart Rate Variability Segments Based on Sleep Stage and Presence of Apnoeic Events", *26th Congress of The European Sleep Research Society (ESRS 2022)*, ISSN:0962-1105, p. 31, Athens (Greece), September 27 - September 30, 2022. DOI: 10.1111/jsr.13739.

3. Gonzalo C. Gutiérrez-Tobal, Javier Gomez-Pilar, Leila Kheirandish-Gozal, **Adrián Martín-Montero**, Jesús Poza, Daniel Álvarez, Félix del Campo, David Gozal, Roberto Hornero, "Slow EEG Oscillation to Characterize Pediatric Sleep Apnea and Associated Cognitive Impairments", *44th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2022)*, ISBN: 978-1-7281-2782-8, pp. 2957-2960, Glasgow (United Kingdom), July 11 - July 15, 2022. DOI: 10.1109/EMBC48229.2022.9871469.

4. Gonzalo C. Gutiérrez-Tobal, Javier Gomez-Pilar, Leila Kheirandish-Gozal, **Adrián Martín-Montero**, Jesús Poza, Daniel Álvarez, Félix del Campo, David Gozal, Roberto Hornero, "Overnight Slow Waves differ in children suffering from sleep apnea and might characterize their cognitive impairment", *World Congress on Medical Physics & Biomedical Engineering (IUPESM WC2022)*, p. 124, Singapur (Singapur), June 12 - June 17, 2022.

5. Gonzalo C. Gutiérrez-Tobal, Javier Gomez-Pilar, Leila Kheirandish-Gozal, **Adrián Martín-Montero**, Jesús Poza, Daniel Álvarez, Félix del Campo, David Gozal, Roberto Hornero, "Network analysis on overnight EEG spec-

trum to assess relationships between paediatric sleep apnoea and cognition", *XV Mediterranean Conference on Medical and Biological Engineering and Computing (MEDICON 2019)*, ISSN: 1680-0737, pp. 1138-1146, Coimbra (Portugal), September 26 - September 28, 2019 , DOI: 10.1007/978-3-030-31635-8_138.

6. Leila Kheirandish-Gozal, Gonzalo C. Gutiérrez-Tobal, **Adrián Martín-Montero**, Jesús Poza, Daniel Álvarez, Félix del Campo, David Gozal, Roberto Hornero, "Spectral EEG Differences in Children with Obstructive Sleep Apnea", *American Thoracic Society International Conference 2019 (ATS 2019)*, Dallas (United States), May 17 - May 22, 2019, DOI: 10.1164/ajrccm-conference.2019.199.1.

7. Gonzalo C. Gutiérrez-Tobal, Daniel Álvarez, Fernando Vaquerizo-Villar, Verónica Barroso-García, **Adrián Martín-Montero**, Andrea Crespo, Félix del Campo, Roberto Hornero,"Pulse rate variability analysis to enhance oximetry as at-home alternative for sleep apnea diagnosing", *World Congress on Medical Physics & Biomedical Engineering (IUPESM 2018)*, ISBN: 978-981-10-9037-0, pp-213-218, Prague (Czech Republic), June 3 - June 8, 2018, DOI: 10.1007/978-981-10-9038-7_39.

A.1.3 National conferences

1. Clara García-Vicente, Gonzalo C. Gutiérrez-Tobal, Jorge Jiménez-García, **Adrián Martín-Montero**, David Gozal, Roberto Hornero, "ECG-ENET: Red neuronal convolucional explicable para la ayuda en el diagnóstico de la apnea del sueño infantil", *XLI Congreso Anual de La Sociedad Española de Ingeniería Biomédica (CASEIB 2023)*, ISBN: 978-84-17853-76-1, pp. 14-17, Cartagena (Spain), November 22 - November 24, 2023.

2. Pablo Armañac-Julián, **Adrián Martín-Montero**, Jesús Lázaro, Spyridon Kontaxis, Daniel Álvarez, David Gozal, Roberto Hornero, Pablo Laguna, Gonzalo C. Gutiérrez-Tobal, Raquel Bailón, Eduardo Gil, "Caracterización de los Cambios en las Medidas de la HRV durante Episodios de Apnea del Sueño", *XL Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2022)*, ISBN: 978-84-09-45972-8, pp. 448-451, Valladolid (Spain), November 23 - November 25, 2022.

3. Jorge Jiménez-García, Gonzalo C. Gutiérrez-Tobal, María García, Daniel Álvarez, **Adrián Martín-Montero**, Félix del Campo, Leila Kheirandish-Gozal,

David Gozal, Roberto Hornero, "Análisis de flujo aéreo y saturación de oxígeno en sangre mediante transformada wavelet para la detección de la apnea obstructiva del sueño infantil", *XXXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2020)*, ISBN: 978-84-09-25491-0, pp. 315-318, Valladolid (Spain), November 25 - November 27, 2020.

4. Gonzalo C. Gutiérrez-Tobal, Leila Kheirandish-Gozal, **Adrián Martín-Montero**, Jesús Poza, Daniel Álvarez, Félix del Campo, David Gozal, Roberto Hornero, "Diferencias en el espectro del EEG nocturno en niños con apnea del sueño", *XXVII Reunión Anual de la Sociedad Española Del Sueño (SES 2019)*, Vitoria (Spain), April 11 - April 13, 2019.

5. **Adrián Martín-Montero**, Gonzalo C. Gutiérrez-Tobal, Daniel Álvarez, Fernando Vaquerizo-Villar, Verónica Barroso-García, Jorge Jiménez-García, Leila Kheirandish-Gozal, Félix del Campo, David Gozal, Roberto Hornero, "Utilidad de nuevas bandas espectrales en la señal de HRV para ayudar en el diagnóstico de la apnea del sueño infantil", *XXXVII Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2019)*, ISBN: 978-84-09-16707-4, pp. 295-298, Santander (Spain), November 27 - November 29, 2019.

6. Jorge Jiménez-García, Gonzalo C. Gutiérrez-Tobal, María García, Daniel Álvarez, Verónica Barroso-García, Fernando Vaquerizo-Villar, **Adrián Martín-Montero**, Félix del Campo, Leila Kheirandish-Gozal, David Gozal, Roberto Hornero, "Evaluación de la información espectral de las señales de flujo aéreo y saturación de oxígeno en sangre para la ayuda al diagnóstico de la apnea del sueño infantil", *XXXVII Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2019)*, ISBN: 978-84-09-16707-4, pp. 25-28, Santander (Spain), November 27 - November 29, 2019.

7. Gonzalo C. Gutiérrez-Tobal, Daniel Álvarez, Verónica Barroso-García, Fernando Vaquerizo-Villar, **Adrián Martín-Montero**, Andrea Crespo, Félix del Campo, Roberto Hornero, "Aplicación de la entropía espectral a la señal de variabilidad de pulso para incrementar el potencial de la oximetría en el diagnóstico de la apnea del sueño a domicilio", *XXVI Reunión Anual de la Sociedad Española Del Sueño (SES 2018)*, Barcelona (Spain), April 26 - April 28, 2018.

8. **Adrián Martín-Montero**, Gonzalo C. Gutiérrez-Tobal, Jesús Poza, Daniel Álvarez, Fernando Vaquerizo-Villar, Verónica Barroso-García, Saúl J. Ruiz-

Gómez, Leila Kheirandish-Gozal, Félix del Campo, David Gozal, Roberto Hornero, "Caracterización de la apnea del sueño infantil mediante nuevas bandas espectrales del EEG", *XXXVI Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2018)*, ISBN: 978-84-09-06253-9, pp. 249-252, Ciudad Real (Spain), November 21 - November 23, 2018.

9. Saúl J. Ruiz-Gómez, Carlos Gómez, Jesús Poza, Pablo Núñez, Victor Rodríguez-González, **Adrián Martín-Montero**, Aarón Maturana-Candelas, Roberto Hornero, "Estudio del efecto de la conducción de volumen en medidas de conectividad funcional derivadas de la coherencia", *XXXVI Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2018)*, ISBN: 978-84-09-06253-9, pp. 241-244, Ciudad Real (Spain), November 21 - November 23, 2018.

10. Pablo Núñez, Jesús Poza, Carlos Gómez, Saúl J. Ruiz-Gómez, **Adrián Martín-Montero**, Miguel A. Tola-Arribas, Mónica Cano, Roberto Hornero, "Estudio de la conectividad neuronal dinámica en la enfermedad de Alzheimer", *XXXV Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2017)*, ISBN: 978-84-9082-797-0, pp. 341-344, Bilbao (Spain), November 29 - December 1, 2017.

A.2 International internship

Three-month research internship at the *Charité - Universitätsmedizin Berlin*, Germany.

i. **Purpose of the internship**
The main purpose of the three-month research stay was to extend the knowledge in the analysis of cardiovascular signal processing techniques in the context of pediatric OSA through the evaluation of the EDR signal. This signal can be combined with HRV to perform an ECG approach to help in the diagnosis of pediatric OSA. To achieve this goal, the next specific objectives were performed: (1) state-of-the-art revision mainly focused on previous works using EDR signals; (2) development and application of EDR signal processing techniques to extract the respiratory power index (RPI), previously used as an alternative to the AHI in adults; (3) combination of RPI adapted to children with HRV OSA-specific metrics to characterize pediatric OSA; (4) analysis and evaluation of the results obtained.

ii. **Quality indicators of the institutions**
The internship took place at the International Center of Sleep Medicine (ICSM) of *Charité - Universitätsmedizin Berlin*, a prestigious research center specializing in signal processing related to sleep disorders. *Charité - Universitätsmedizin Berlin* is an internationally renowned and multicenter hospital complex with a rich history of over 300 years. It is home to a strong research community consisting of 3700 scientists. Notably, 11 Nobel laureates have been affiliated with the institution. The trainee also collaborated with the Cardiovascular Physics research group at Humboldt University in Berlin during the internship. In 2022, both *Charité - Universitätsmedizin Berlin* and Humboldt University of Berlin ranked between positions 51 and 75 in the Shanghai Ranking. The internship was supervised by Professor Thomas Penzel, the scientific chair of ICSM, who has an impressive publication record with over 250 JCR articles, book chapters, and books in the field of sleep medicine. He has received numerous prestigious awards, including the "Bial Award" for clinical medicine in Portugal in 2001, the "Bill Gruen Award" for Innovations in Sleep Research by the Sleep Research Society in 2008, and the "Somnus Award" by *Schlafmagazin* for outstanding service in sleep medicine in 2012. Professor Penzel has also supervised more than 19 Doctoral Theses to date, further demonstrating his expertise in the field.

A.3 National internship

One-month research internship at the Biomedical Signal Interpretation and Computational Simulation (BSICoS) research group from the University of Zaragoza, Spain.

i. **Purpose of the internship**
The primary objective of this research internship was to delve into cardiovascular signal processing techniques, which is a key area of expertise for the BSICoS research group, the host group for the internship. Thus, the purpose of the internship was directly aligned with the main focus of the present Doctoral Thesis. It was part of an intramural project titled Sleepy-Heart, conducted within the Centro de Investigación Biomédica en Red – Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), with the aim of developing a screening tool for pediatric OSA using HRV to diagnose severely affected children and identify those at risk of developing cardio-

vascular problems associated with OSA. Therefore, the research project topic was also closely connected to this Doctoral Thesis. Moreover, this research internship also allowed the strengthening of the research collaboration between BSICoS and the Biomedical Engineering Group at the University of Valladolid, where this Doctoral Thesis has been developed. As a result of the primary research conducted during the internship, several publications have emerged, including one JCR article (the fourth in the collection of articles presented here), two international conference papers, and one national conference document authored by members of these institutions.

ii. **Quality indicators of the institutions**
The internship was conducted at the Biomedical Signal Interpretation and Computational Simulation (BSICoS) research group from the University of Zaragoza in Spain. The primary focus of the group is to develop methods for biomedical signal processing, with a specific emphasis on personalized interpretation of cardiovascular, respiratory, and autonomic nervous system conditions, as well as their interactions. Their objective is to enhance the impact of information and communication technologies in healthcare and deepen the understanding of the functioning of biological systems through the analysis of noninvasive signals. To reach this goal, they collaborate with clinical teams and research groups that combine expertise from both areas. Since 1990, BSICoS research group has made significant contributions, publishing over 300 JCR articles, which attests to their expertise in the field of biomedical signal processing analysis. They are recognized as an esteemed international institution and one of the leading research groups in biomedical engineering in the country. Professor Pablo Laguna Lasaosa, the scientific chair of BSICoS and the supervisor of the internship, has played a prominent role in advancing the field of biomedical signal interpretation, particularly in the cardiovascular domain. He has co-authored more than 150 research papers, published over 300 international conference papers, and has provided guidance for 18 Ph.D. theses.

A.4 Grants

09/2022: **Erasmus + scholarship to perform a three-month internship** (Berlin, Germany, 01/09/2022 - 01/12/2022), funded by the University of Val-

ladolid with funds from the European Union.

07/2019: **'Ayuda para contratos predoctorales para la formación de doctores (FPI)' grant** (PRE2018-085219), funded by the Ministerio de Ciencia, Innovación y Universidades from the Spanish Government.

A.5 Awards and honors

09/2019: **Young Investigator Competition Award at the XV Mediterranean Conference on Medical and Biological Engineering and Computing (MEDICON 2019)**, for the conference paper entitled "Network analysis on overnight EEG spectrum to assess relationships between paediatric sleep apnoea and cognition", conducted by Gonzalo C. Gutiérrez-Tobal, Javier Gomez-Pilar, Leila Keirandish-Gozal, **Adrián Martín-Montero**, Jesús Poza, Daniel Álvarez, Félix del Campo, David Gozal and Roberto Hornero.

11/2018: **First Jose María Ferrero Corral Award at the XXXVI Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2018)**, awarded by the Sociedad Española de Ingeniería Biomédica (SEIB), for the conference paper entitled "Caracterización de la apnea del sueño infantil mediante nuevas bandas espectrales del EEG", conducted by **Adrián Martín-Montero**, Gonzalo C. Gutiérrez-Tobal, Jesús Poza, Daniel Álvarez, Fernando Vaquerizo-Villar, Verónica Barroso-García, Saúl J. Ruiz-Gómez, Leila Kheirandish-Gozal, Félix del Campo, David Gozal and Roberto Hornero.

11/2017: **First Jose María Ferrero Corral Award at the XXXV Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2017)**, awarded by the Sociedad Española de Ingeniería Biomédica (SEIB), for the conference paper entitled "Estudio de la conectividad neuronal dinámica en la enfermedad de Alzheimer", conducted by Pablo Núñez, Jesús Poza, Carlos Gómez, Saúl J. Ruiz-Gómez, **Adrián Martín-Montero**, Miguel A. Tola-Arribas, Mónica Cano and Roberto Hornero.

Apéndice B

Resumen en castellano

B.1 Introducción

La apnea obstructiva del sueño (AOS) pediátrica es un trastorno respiratorio del sueño que se caracteriza por periodos de interrupción total del flujo aéreo (conocidos como apneas), así como periodos de reducción significativa del flujo aéreo (conocidos como hipopneas) (American Thoracic Society, 1996). Esta enfermedad tiene una alta prevalencia, padeciéndola alrededor del 6 % de la población pediátrica general (Marcus et al., 2012), y se ha asociado con un aumento en los problemas cardiovasculares, como la hipertrofia ventricular izquierda y derecha, alteraciones en la regulación autónoma de la frecuencia cardíaca y cambios en la presión sanguínea (Marcus et al., 2012). Además, la AOS también se ha relacionado con problemas cognitivos, incluyendo déficits en el aprendizaje y en el rendimiento académico (Hunter et al., 2016), lo que resalta la importancia de la detección temprana y el tratamiento para mantener la salud cardiovascular a largo plazo y el potencial académico de los niños. La prueba de referencia diagnóstica para la AOS es la polisomnografía (PSG) nocturna, durante la cual los niños tienen que pasar una noche en un laboratorio del sueño, tratándose de un proceso complejo, costoso y especialmente incómodo para ellos (Tan et al., 2014). Por ello, los inconvenientes ligados a la PSG han motivado la búsqueda de alternativas para diagnosticar la AOS pediátrica y estudiar sus consecuencias (Alonso-Álvarez et al., 2015; Marcus et al., 2012; Tan et al., 2014). Entre todas las alternativas al diagnóstico propuestas, la mayoría se han enfocado al análisis de un conjunto reducido de las señales adquiridas durante la PSG, centrándose principalmente en el análisis de la señal de saturación de oxígeno en la sangre (SpO_2). No obstante,

existe una escasez de evidencia en la literatura científica respecto al uso de señales biomédicas distintas a la SpO_2, como por ejemplo las señales cardiovasculares, como una opción alternativa para el diagnóstico de la AOS (Gutiérrez-Tobal et al., 2022).

En este sentido, en el marco de esta Tesis Doctoral se ha llevado a cabo una caracterización exhaustiva de la variabilidad de la frecuencia cardíaca (*heart rate variability*, HRV) con el objetivo de describir las alteraciones que ocurren en el sistema nervioso autónomo (SNA) debido a la AOS y ayudar a simplificar el diagnóstico de esta enfermedad. Los resultados obtenidos a partir de estas investigaciones han sido publicados en cuatro artículos científicos en revistas indexadas en el *Journal Citation Reports* (JCR) entre los años 2021 y 2023. En estas publicaciones, se emplearon técnicas de procesado de señales biomédicas (extracción de características, selección de características y reconocimiento de patrones) en mayor o menor grado para caracterizar la señal de HRV y obtener información innovadora en el contexto de la AOS pediátrica. En el primer artículo, se realizó un análisis espectral de la señal de HRV para identificar bandas de frecuencia específicas de la AOS (Martín-Montero et al., 2021b). En el segundo artículo, se utilizó el análisis biespectral, una técnica de análisis en el dominio de la frecuencia que permite observar alteraciones en la no Gaussianidad, no linealidad e irregularidad, con el fin de definir regiones basadas en los rangos de frecuencia clásicos y específicos de la AOS en la señal de HRV y caracterizar así las alteraciones asociadas a la AOS (Martín-Montero et al., 2021a). La utilidad de las nuevas características específicas de la AOS para reflejar los efectos del tratamiento de la enfermedad en la HRV se realizó en el tercer artículo mediante el empleo de análisis de mediación causal (Martín-Montero et al., 2022). Por último, en el cuarto artículo se realizó un análisis a nivel de segmento para caracterizar la señal de HRV utilizando métricas convencionales y métricas específicas desarrolladas en el contexto de esta Tesis Doctoral, con el fin de evaluar la influencia en el HRV de las diferentes fases del sueño junto con la presencia de eventos apneicos durante la noche (Martín-Montero et al., 2023).

La evaluación de la utilidad clínica de los diferentes enfoques de caracterización de la HRV se realizó mediante una etapa de reconocimiento de patrones, empleando distintos modelos de *machine-learning* para el diagnóstico de AOS pediátrica (Martín-Montero et al., 2023, 2021a,b), así como para la clasificación automática de fases del sueño (Martín-Montero et al., 2023).

B.2 Hipótesis y objetivos

La presencia de eventos de apnea e hipopnea provoca oscilaciones recurrentes en la señal de la HRV. Estas variaciones cíclicas han sido previamente documentadas en adultos, manifestándose como patrones de bradicardia progresiva seguidos de una taquicardia abrupta al final de los eventos (Guilleminault et al., 1984). En niños también se han observado estos patrones, pero con una mayor variabilidad en función de las características de las pausas respiratorias (Aljadeff et al., 1997; Gozal et al., 2013; O'Driscoll et al., 2009; Vitelli et al., 2016). A pesar de estas alteraciones cardíacas directamente causadas por la enfermedad, el comportamiento periódico de la señal de HRV a causa de la AOS pediátrica ha sido estudiado tradicionalmente utilizando técnicas de análisis espectral dentro de los límites establecidos por las bandas de frecuencia clásicas de la HRV (Malik et al., 1996). Sin embargo, un estudio previo en adultos (Gutiérrez-Tobal et al., 2015) reveló que las alteraciones asociadas a la AOS se reflejan en un rango de frecuencia que se ubica entre las bandas clásicas de frecuencia muy baja (*very low frequency*, VLF) y baja frecuencia (*low frequency*, LF), lo cual sugiere la existencia de bandas de frecuencia específicas de la AOS en la HRV. Por lo tanto, al inicio de esta Tesis Doctoral, se planteó la hipótesis de que *nuevos enfoques de extracción de características podrían permitir una caracterización más específica de los patrones de la HRV nocturna en niños*.

Las alteraciones observadas en el SNA como resultado de la AOS han generado investigaciones sobre el efecto del tratamiento de la enfermedad en las señales cardiovasculares. Varios estudios han informado de una reversión de las alteraciones de la AOS, principalmente manifestada como un retorno del tono simpático a niveles normales (Constantin et al., 2008; El-Hamad et al., 2017; Kaditis et al., 2011; Muzumdar et al., 2011). Dado que estos descubrimientos también se basaron en el análisis convencional de las señales cardiovasculares, se ha planteado la hipótesis de que *las mejoras en la gravedad y resolución de la AOS pediátrica como resultado del tratamiento podrían inducir cambios en las características específicas de la HRV relacionadas con la AOS*.

En los últimos años se han realizado investigaciones sobre métodos automáticos como posibles alternativas al diagnóstico de la AOS mediante la PSG, con la mayoría de dichos estudios centrándose en características derivadas de la señal de SpO_2 (Gutiérrez-Tobal et al., 2022). Sin embargo, debido a la escasa presencia en la literatura de alternativas al diagnóstico de la AOS en niños utilizando señales distintas a la de SpO_2, se plantea la hipótesis de que *las nuevas características de HRV pueden proporcionar información complementaria que, en combinación con mo-*

delos de machine-learning, podrían ser útiles para ayudar en el diagnóstico automático y la determinación de la gravedad de la AOS en la población pediátrica.

Estas consideraciones constituyen la base fundamental de la presente Tesis Doctoral, que se puede sintetizar en la hipótesis global de que *la caracterización de la señal de HRV mediante enfoques novedosos podría revelar una comprensión más profunda del comportamiento nocturno del SNA asociado a la AOS pediátrica, ayudando a simplificar el diagnóstico de la enfermedad*.

En base a esta hipótesis, el objetivo general de la presente Tesis Doctoral es *diseñar, desarrollar y evaluar nuevas técnicas automáticas de procesamiento de la señal de HRV que permitan una descripción exhaustiva de las alteraciones en el comportamiento del SNA durante la noche y facilitar el diagnóstico de la AOS en niños*. Para alcanzar este objetivo principal, se han establecido los siguientes objetivos específicos:

I. Identificar características novedosas de HRV que contengan información relevante y no redundante sobre el funcionamiento del SNA en presencia de eventos apneicos en niños con AOS.

II. Evaluar la capacidad de las características específicas de HRV relacionadas con la AOS para reflejar el impacto del tratamiento en el SNA de los niños.

III. Diseñar e implementar modelos de *machine-learning* para evaluar la utilidad clínica de la caracterización novedosa de HRV, tanto en la detección y evaluación de la gravedad de la enfermedad como en la clasificación de las fases del sueño, utilizando conjuntos óptimos de características de HRV.

B.3 Sujetos y señales

A lo largo de esta Tesis Doctoral, se llevaron a cabo investigaciones utilizando dos bases de datos distintas de electrocardiograma (ECG) pediátrico: una privada y otra pública. La base de datos privada, denominada como base de datos de la Universidad de Chicago (UofC), fue obtenida del *Comer Children's Hospital* de la escuela de medicina de la Universidad de Chicago (Chicago, IL, EE. UU.), mientras que la base de datos pública fue la base de datos del ensayo aleatorizado *Childhood Adenotonsillectomy Trial* (CHAT) (Marcus et al., 2013; Redline et al., 2011). Ambas bases de datos incluyeron registros de ECG de niños de edades comprendidas entre 0 y 13 años, que fueron referidos para una prueba de sueño debido a la presencia de síntomas clínicos de AOS, como ronquidos, apneas, excesiva somnolencia diurna e hipertrofia amigdalar, entre otros. La base de datos

CHAT constaba de 1612 estudios de sueño de múltiples centros, mientras que la base de datos UofC incluía a 981 niños, todos sometidos a una PSG nocturna completa en la unidad pediátrica de sueño de los respectivos hospitales. Los estudios de sueño en ambas bases de datos fueron anotados por especialistas médicos siguiendo las pautas de la Academia Americana de Medicina del Sueño (Berry et al., 2012). El índice de apnea-hipopnea (IAH) se extrajo de estas anotaciones para determinar la gravedad de la AOS en los sujetos pediátricos, evaluado como el número de eventos de apnea e hipopnea por cada hora de sueño (e/h). Por similitud con lo realizado en estudios previos (Church, 2012; Hornero et al., 2017; Tan et al., 2014), se utilizaron tres puntos de corte del IAH (1, 5 y 10 e/h) para categorizar la gravedad de la AOS en cuatro grupos: no-AOS (IAH $<$ 1 e/h), AOS leve (1 $\leq$ IAH $<$ 5 e/h), AOS moderada (5 $\leq$ IAH $<$ 10 e/h) y AOS severa (IAH $\geq$ 10 e/h). Las Tablas C.1 y C.2 recogen los datos demográficos y clínicos de los sujetos incluidos en ambas bases de datos, incluyendo información sobre la población total, edad, sexo, índice de masa corporal y el número de niños pertenecientes a cada grupo de gravedad.

B.4 Métodos

En esta Tesis Doctoral, se aplicaron metodologías de extracción, selección y clasificación de características para analizar las señales de HRV. Estas señales se extrajeron de la señal de ECG monocanal de las diferentes bases de datos, lo cual requirió una etapa inicial de pre-procesamiento. Por tanto, las señales de ECG originales fueron pre-procesadas para eliminar artefactos, como ruido muscular o de movimiento, pérdida de la señal o cambios en la línea de base. A continuación, se

Tabla B.1. Datos demográficos y clínicos de los niños en la base de datos de CHAT.

	Todos	no-AOS	AOS leve	AOS moderada	AOS severa
Sujetos (*n*)	1612 (100 %)	346 (21.5 %)	798 (49.5 %)	253 (15.7 %)	215 (13.3 %)
Edad (años)	7.0 [6.0, 8.0]	7.0 [6.0, 8.3]	6.8 [6.0, 8.0]	7.0 [6.0, 8.0]	6.6 [5.8, 8.0]
Varones (*n*)	774 (48.0 %)	163 (47.1 %)	380 (47.6 %)	123 (48.6 %)	108 (50.2 %)
IMC (kg/m^2)	17.3 [15.6, 21.6]	17.1 [15.5, 19.6]	17.2 [15.6, 20.8]	18.5 [15.4, 23.3]	18.6 [15.9, 23.8]
IAH (e/h)	2.5 [1.2, 6.0]	0.6 [0.4, 0.8]	2.2 [1.6, 3.3]	7.1 [6.0, 8.5]	17.1 [12.7, 26.7]

Los datos se muestran como mediana [percentil 25, percentil 75], o número (porcentaje). IAH= índice de apnea-hipopnea, IMC: índice de masa corporal, AOS: apnea obstructiva del sueño.

Tabla B.2. Datos demográficos y clínicos de los niños en la base de datos UofC.

	Todos	no-AOS	AOS leve	AOS moderada	AOS severa
Sujetos (*n*)	981 (100 %)	173 (17.6 %)	401 (40.9 %)	178 (18.1 %)	229 (23.3 %)
Edad (años)	6 [3, 9]	7 [4, 10]	6 [4, 9]	5 [2, 8]	4 [2, 8]
Varones (*n*)	602 (61.4 %)	107 (61.8 %)	247 (61.6 %)	109 (61.2 %)	139 (60.7 %)
IMC (kg/m^2)	18.0 [16.1, 21.9]	17.6 [15.6, 21.0]	17.9 [16.1, 21.2]	18.6 [16.5, 24.0]	18.3 [16.1, 23.2]
IAH (e/h)	3.8 [1.5, 9.3]	0.5 [0.1, 0.8]	2.5 [1.7, 3.5]	6.8 [5.8, 8.3]	19.1 [13.9, 31.1]

Los datos se muestran como mediana [percentil 25, percentil 75], o número (porcentaje). IAH= índice de apnea-hipopnea, IMC: índice de masa corporal, AOS: apnea obstructiva del sueño.

realizó la extracción del complejo QRS, lo que permitió calcular la HRV a partir de los picos R detectados. La fase de pre-procesado difirió entre los primeros tres artículos (Martín-Montero et al., 2021a,b, 2022) y el último (Martín-Montero et al., 2023). Mientras que en los primeros se realizó un análisis continuo durante toda la noche, en el último artículo se segmentó la señal de ECG en segmentos de 10 minutos antes de calcular la HRV.

Una vez completado el pre-procesado de las señales, se procedió a aplicar las técnicas de procesado de señales biomédicas, las cuales se dividieron en tres etapas fundamentales: extracción de características, selección de características y reconocimiento de patrones. La fase de extracción de características tuvo como objetivo la descripción exhaustiva de la HRV nocturna mediante el uso de técnicas novedosas que pudieran complementar los métodos de análisis tradicionales y proporcionar información adicional sobre las alteraciones que ocurren debido a la AOS. Se emplearon técnicas de diferente naturaleza que permitiesen reflejar el comportamiento de la HRV: técnicas en el dominio temporal (Martín-Montero et al., 2023), para caracterizar la actividad en los diferentes segmentos considerados de la HRV; técnicas en el dominio espectral (Martín-Montero et al., 2023, 2021b, 2022), para describir la actividad frecuencial de la HRV a lo largo de toda la noche y en los segmentos específicos, así como para definir nuevas bandas de frecuencia específicas de la AOS; y técnicas de análisis biespectral (Martín-Montero et al., 2021a), que permiten realizar una evaluación en el dominio de la frecuencia identificando relaciones de fase, así como desviaciones de linealidad y Gaussianidad en los registros de HRV (Chua et al., 2010).

Tras la extracción de características, en el caso del estudio de bispectrum (Martín-Montero et al., 2021a), se obtuvo un conjunto de características que con-

tenía un elevado número de parámetros, lo que podría implicar la presencia de información irrelevante o redundante. Con el fin de abordar esta situación, se realizó una etapa de selección de características para obtener dos subconjuntos óptimos: uno que incluyera características obtenidas en regiones biespectrales definidas según los rangos de frecuencia clásicos de la HRV, y otro que incluyera características obtenidas en regiones definidas en base a los rangos de frecuencia de HRV específicos de la AOS. Para lograrlo, se utilizó el método *fast correlation-based filter* (FCBF) (Yu and Liu, 2004), que permite obtener subconjuntos de características relevantes y no redundantes.

Realizada la extracción y selección de características, los conjuntos de parámetros de HRV obtenidos mediante los diferentes enfoques se utilizaron en la última etapa, la fase de reconocimiento de patrones, con el propósito de evaluar la utilidad clínica de la caracterización de HRV mediante diversos modelos de *machine-learning*. Para evaluar dicha utilidad clínica en el diagnóstico de la AOS pediátrica, se emplearon tres métodos distintos: en Martín-Montero et al. (2021b) se desarrollaron modelos de clasificación binaria de la AOS utilizando análisis lineal discriminante (LDA). En Martín-Montero et al. (2021a), se llevó a cabo la clasificación binaria mediante clasificadores de perceptrón multicapa (*multi-layer perceptron*, MLP) utilizando los dos subconjuntos óptimos de características de HRV obtenidos tras la fase de selección de características. Por último, en Martín-Montero et al. (2023), se emplearon técnicas de *ensemble learning*, utilizando un modelo de regresión *least squares boosting* (LSBoost) para estimar el IAH de los sujetos. Además, también en Martín-Montero et al. (2023) se implementó un modelo de clasificación multiclase *adaptive boosting* (AdaBoost) para distinguir los segmentos de HRV correspondientes a las fases del sueño *rapid eye movement* (REM), *non-rapid eye movement* (NREM) o estado de vigilia. Para profundizar en el entendimiento del proceso de decisiones tomado por los modelos de *ensemble-learning*, en Martín-Montero et al. (2023) se incorporaron también técnicas de *explainable artificial intelligence* (XAI) a los modelos de LSBoost y AdaBoost, centrados en el análisis de la importancia de características.

Por otro lado, se realizó la evaluación de la utilidad clínica de la actividad de la HRV tanto en las bandas clásicas como en las nuevas bandas específicas de la AOS mediante un análisis de mediación causal, para evaluar los efectos del tratamiento de la AOS en los cambios que se producen en esos parámetros de HRV (Martín-Montero et al., 2022).

Por último, se han utilizado diversos métodos de análisis estadístico para evaluar e interpretar los resultados obtenidos mediante las técnicas de procesamien-

to de la señal de HRV desarrolladas en esta Tesis Doctoral. Estos métodos incluyen pruebas estadísticas de hipótesis (como Lilliefors, Levene, Mann-Whitney, Kruskal-Wallis y *d* de Cohen), análisis de correlaciones (mediante correlaciones parciales de Spearman para controlar el posible efecto de covariables de confusión como la edad), visualización de datos (usando *boxplots*), medidas de rendimiento diagnóstico (tales como sensibilidad, especificidad, precisión, área bajo la curva *receiver-operating characteristics*, valor predictivo positivo, valor predictivo negativo, razón de verosimilitud positiva, razón de verosimilitud negativa y *F1-score*), medidas de concordancia (como el índice kappa de Cohen) y estrategias de validación (como *hold-out* y *bootstrapping*).

B.5 Resultados y discusión

Cada uno de los enfoques realizados en esta Tesis Doctoral tenía como objetivo caracterizar el HRV nocturno y evaluar las alteraciones que ocurren en el SNA debido a la AOS en niños, para ayudar en la simplificación de su diagnóstico. Las metodologías aplicadas han permitido caracterizar estas alteraciones, revelando nuevos parámetros que han proporcionado información previamente desconocida tanto durante la noche como en las diferentes fases del sueño. Además, se ha demostrado la utilidad clínica de esta caracterización para evaluar los efectos del tratamiento, así como su aplicabilidad clínica para el diagnóstico de la AOS pediátrica y la clasificación de las fases del sueño.

El análisis espectral llevado a cabo en Martín-Montero et al. (2021b) permitió identificar y evaluar tres bandas de frecuencia, con el objetivo de mejorar la comprensión de las dinámicas cardiovasculares en presencia de AOS en niños. La primera banda, denominada BW1 (0.001 - 0.005 Hz), mostró una asociación con el número de despertares durante la noche y el tiempo despierto después del inicio del sueño, lo que sugiere que la fragmentación del sueño (consecuencia directa de la AOS) se refleja en esta banda frecuencial y puede caracterizarse mediante su análisis. La segunda banda, denominada BW2 (0.028-0.074 Hz), mostró las correlaciones más altas con los índices respiratorios que evalúan la gravedad de la AOS, y también mostró el mejor rendimiento diagnóstico individual entre todos los parámetros analizados. Esta banda de frecuencia captura patrones de recurrencia con una duración de 13 a 35 segundos, lo que podría indicar una duración característica de los patrones cardíacos resultantes de la AOS, siendo adecuada para caracterizar la activación simpática como resultado de la enfermedad. Por último, la tercera banda, llamada BWRes (0.04 Hz alrededor del pico

respiratorio individual), mostró una asociación con la desaturación de oxígeno causada por la AOS, así como con los microdespertares relacionados. Además, se observó que el rendimiento diagnóstico de estas bandas específicas de la AOS fue superior, tanto de forma individual como combinada, en comparación con las bandas clásicas de HRV utilizadas hasta el momento. Estos hallazgos destacan la importancia de considerar estos rangos frecuenciales al analizar la señal de HRV en el contexto de la AOS pediátrica.

En cuanto al análisis biespectral realizado en Martín-Montero et al. (2021a), se encontró que, en ausencia de AOS, existe un enfoque de acoplamiento de potencia biespectral por debajo de 0.02 Hz, el cual se amplía hacia frecuencias más altas a medida que aumenta la severidad de la enfermedad. Este incremento a frecuencias más altas (desde las frecuencias de la región VLF hasta las frecuencias de la región BW2) en presencia de la AOS indica un aumento en la Gaussianidad y linealidad en esas frecuencias como resultado de los eventos apneicos. La caracterización de las regiones biespectrales y la posterior selección de características nos permitieron identificar dos conjuntos óptimos de características: uno basado en las regiones biespectrales definidas según los rangos clásicos de HRV y otro basado en los rangos específicos de la AOS. Estos subconjuntos contenían características complementarias que revelaron alteraciones en la no-Gaussianidad, no-linealidad y la irregularidad de la HRV debido a la AOS. Entre las características seleccionadas, se encontró que una medida novedosa propuesta en este estudio, denominada RP_{Diag}, presentó el mejor rendimiento diagnóstico individual y las correlaciones más altas con los índices polisomnográficos cuando se evaluaba en la región definida según BW2. Esto demuestra la utilidad del análisis biespectral realizado en el contexto de la AOS pediátrica, especialmente en dicha región.

El empleo del análisis de mediación causal realizado en Martín-Montero et al. (2022) permitió inferir que los cambios observados en el HRV después de la intervención para la AOS pediátrica están causalmente relacionados con los efectos del tratamiento. Este análisis estableció vínculos entre las alteraciones en las características de la HRV y las modificaciones en variables polisomnográficas directamente asociadas con la severidad de la AOS y su resolución. Todos los mediadores considerados demostraron tener un impacto causal en al menos uno de los parámetros de la HRV evaluados. De todas las características analizadas, la variación de potencia en la banda BW2 fue el único parámetro que mostró un efecto causal estadísticamente significativo en todos los mediadores, y presentando el nivel más alto de significancia estadística en la mayoría de esos efectos. Además, entre los parámetros que mostraron efectos de causalidad con la resolu-

ción de AOS, solo la potencia en la banda BW2 permitió distinguir la actividad de la HRV entre niños que resolvieron la enfermedad y aquellos que no lo hicieron, proponiéndose por lo tanto como un biomarcador potencial de la resolución de la AOS pediátrica.

Con el fin de completar la caracterización exhaustiva de las alteraciones de la HRV que ocurren en presencia de AOS en niños, en Martín-Montero et al. (2023) se llevó a cabo un análisis a nivel de segmento que examinó tanto la influencia de las fases del sueño como la presencia de eventos apneicos. Este enfoque reveló que, aunque la mayoría de los eventos apneicos suelen ocurrir durante la fase REM, la activación simpática basal inherente a esta etapa del sueño parece inhibir la activación simpática resultante de los eventos típicos de la AOS. Por lo tanto, la activación simpática inducida por la AOS es más discernible durante la fase NREM, destacando la importancia de considerar tanto las fases del sueño como la presencia de eventos apneicos para lograr una caracterización completa del comportamiento de la HRV en presencia de AOS. Además, la evaluación de la importancia de las características reveló que la caracterización en dos de las nuevas bandas espectrales de HRV específicas de la AOS, BW2 y BWRes, tienen la mayor relevancia para el diagnóstico de la AOS pediátrica y la clasificación automática de las fases del sueño, respectivamente. Estos hallazgos vuelven a destacar la importancia de las características de HRV específicas de la AOS desarrolladas en el marco de la presente Tesis Doctoral.

En virtud de estas consideraciones, podemos afirmar que los diversos enfoques metodológicos aplicados en el transcurso de esta Tesis Doctoral han permitido caracterizar las alteraciones que ocurren en el SNA, obteniendo nuevos parámetros que ofrecen información relevante relacionada con la AOS en niños. En cuanto a la evaluación clínica de la utilidad de la caracterización de la HRV para el diagnóstico de la AOS, la Tabla C.3 muestra los resultados más relevantes de los diferentes estudios para la clasificación de sujetos utilizando tres umbrales de severidad de IAH (1, 5 y 10 e/h) (Martín-Montero et al., 2023, 2021a,b). En general, se observa que los modelos MLP construidos con características biespectrales superaron el rendimiento diagnóstico de los modelos de LDA que utilizaban características espectrales, alcanzándose el mayor rendimiento diagnóstico en los umbrales de IAH de 5 y 10 e/h de toda la Tesis Doctoral. Sin embargo, el mejor rendimiento general en el umbral de 1 e/h se obtuvo con el modelo de LSBoost en términos de sensibilidad, precisión, valor predictivo positivo, área bajo la curva *receiver-operating characteristics*, y *F1-score*, lo que demuestra el alto potencial del modelo de LSBoost para la detección de la AOS pediátrica en su grado de se-

Tabla B.3. Rendimiento diagnóstico alcanzado por los diferentes modelos de *machine-learning* a través de los estudios de la presente Tesis Doctoral en los tres umbrales de severidad de IAH. Todos los resultados se obtuvieron en el grupo de test de cada estudio, es decir, el subgrupo *nonrandomized* de la base de datos CHAT en Martín-Montero et al. (2021b) y Martín-Montero et al. (2021a), y un subgrupo formado por sujetos de los distintos grupos de la base de datos de CHAT en Martín-Montero et al. (2023).

Umbral: IAH = 1 e/h										
Estudio	**Modelo**	**S**	**E**	**P**	**VPP**	**VPN**	**RV$^+$**	**RV$^-$**	**AUC**	***F1-score***
Martín-Montero et al. (2021b)	**LDA**$_{Classic}$	25.70	81.30	44.50	72.90	35.90	1.37	0.91	0.559	0.326
	LDA$_{Specific}$	37.70	80.10	52.00	78.80	39.70	1.89	0.778	0.597	0.437
Martín-Montero et al. (2021a)	**MLP1**$_{Classic}$	52.30	59.40	54.70	71.58	38.87	1.29	0.80	0.600	0.535
	MLP1$_{Specific}$	76.30	38.30	63.40	70.74	45.16	1.24	0.62	0.627	0.693
Martín-Montero et al. (2023)	**LSBoost**	90.76	23.40	80.07	86.26	32.35	1.18	0.39	0.651	0.885
Umbral: IAH = 5 e/h										
Estudio	**Modelo**	**S**	**E**	**P**	**VPP**	**VPN**	**RV$^+$**	**RV$^-$**	**AUC**	***F1-score***
Martín-Montero et al. (2021b)	**LDA**$_{Classic}$	46.40	72.20	68.40	22.50	88.60	1.67	0.74	0.633	0.553
	LDA$_{Specific}$	48.20	80.80	76.00	30.30	90.00	2.51	0.64	0.696	0.590
Martín-Montero et al. (2021a)	**MLP5**$_{Classic}$	50.90	86.20	81.00	39.04	91.00	3.69	0.570	0.774	0.625
	MLP5$_{Specific}$	62.50	84.20	81.00	40.70	92.82	3.96	0.45	0.791	0.706
Martín-Montero et al. (2023)	**LSBoost**	66.67	61.17	63.18	49.66	76.16	1.72	0.54	0.677	0.569
Umbral: IAH = 10 e/h										
Estudio	**Modelo**	**S**	**E**	**P**	**VPP**	**VPN**	**RV$^+$**	**RV$^-$**	**AUC**	***F1-score***
Martín-Montero et al. (2021b)	**LDA**$_{Classic}$	50.70	75.30	73.10	17.10	93.80	2.05	0.65	0.685	0.599
	LDA$_{Specific}$	62.30	84.30	82.30	28.50	95.70	3.97	0.45	0.774	0.709
Martín-Montero et al. (2021a)	**MLP10**$_{Classic}$	43.50	96.50	91.70	55.56	94.45	12.43	0.59	0.847	0.590
	MLP10$_{Specific}$	66.70	91.60	89.30	44.23	96.48	7.94	0.36	0.841	0.764
Martín-Montero et al. (2023)	**LSBoost**	40.00	92.03	84.12	47.37	89.53	5.02	0.65	0.742	0.434

IAH: índice de apnea-hipopnea, S: sensibilidad (%), E: especificidad (%), P: precisión (%), VPP: valor predictivo positivo (%), VPN: valor predictivo negativo (%), RV$^+$: razón de verosimilitud positiva, VR$^-$: razón de verosimilitud negativa, AUC: area bajo la curva *receiver-operating characteristic*, LDA: análisis discriminante lineal, MLP: perceptrón multicapa, LSBoost: *least-squares boosting*.

veridad más leve. Además, al contrastar nuestros resultados con estudios previos de la literatura que emplearon exclusivamente medidas cardiovasculares para el diagnóstico de la AOS en niños, nuestras metodologías propuestas demostraron un rendimiento similar o superior, particularmente para detectar la presencia de la AOS. En cuanto a la utilidad clínica para la clasificación automática de las fases del sueño, el modelo de AdaBoost obtuvo los mejores resultados de clasificación en la fase NREM, con un valor predictivo positivo del 94.55 % y una sensibilidad del 72.08 %. Estos resultados son consistentes con la mayor capacidad discriminatoria entre los segmentos con distinto grado de presencia de eventos apneicos observada en la caracterización de la HRV durante la fase NREM. Por tanto, en

base a estos resultados, podemos concluir que la caracterización de la HRV realizada es útil tanto para la ayuda en el diagnóstico de la AOS infantil, así como para la clasificación de las fases de sueño en niños con AOS.

B.6 Conclusiones

La evaluación minuciosa y la posterior discusión de los resultados obtenidos durante la realización de la presente Tesis Doctoral han permitido la extracción de las principales conclusiones derivadas de la misma, las cuales se presentan a continuación:

1) La nueva banda de HRV específica de la AOS pediátrica BW1 (0.001 – 0.005 Hz) está relacionada con las macro interrupciones del sueño causadas por la AOS. El tratamiento de la enfermedad produce cambios dentro de este rango espectral, los cuales pueden atribuirse a las alteraciones en el IAH obstructivo y en el índice de despertares.

2) La nueva banda de HRV específica de AOS pediátrica BW2 (0.028-0.074 Hz) se encuentra asociada de forma robusta con la duración de los eventos apneicos. Las intervenciones para tratar la AOS generan modificaciones dentro de este rango de frecuencia, principalmente debido a cambios en la severidad de la enfermedad, lo que resulta en una disminución en la actividad de esta banda espectral si la severidad se reduce. Además, esta banda fue la única que mostró efectos de causalidad con la resolución de la AOS, permitiendo diferenciar la actividad de la HRV entre niños que resuelven la enfermedad y niños que no. En consecuencia, BW2 se ha propuesto como un potencial biomarcador de la resolución de la AOS pediátrica. Además, empleando únicamente la señal de HRV, esta banda de frecuencia específica de la AOS ha demostrado ser de gran utilidad para ayudar en el diagnóstico automático de la enfermedad.

3) La nueva banda de HRV específica de AOS pediátrica BWRes (0.04 Hz alrededor del pico respiratorio individual) está relacionada con las desaturaciones de oxígeno y las micro interrupciones del sueño. El tratamiento de la enfermedad provoca cambios dentro de este rango frecuencial, debido principalmente a las variaciones en el índice de despertares. Además, esta banda de frecuencia de HRV específica de AOS ha demostrado ser de gran utilidad a la hora de clasificar fases del sueño en el contexto de la AOS pediátrica mediante el empleo de la señal HRV.

4) Los hallazgos del análisis biespectral revelaron que la AOS pediátrica está asociada con incrementos en la linealidad y la Gaussianidad de la HRV a frecuencias muy bajas. Además, la repetición de eventos apneicos da lugar a un desplazamiento del foco de acoplamiento en el HRV nocturno hacia frecuencias relacionadas con la duración de los eventos. Por otro lado, la presencia de AOS provoca una disminución en la irregularidad de la HRV en componentes frecuenciales asociadas con las bandas BW2 y LF, mientras que hay un descenso de irregularidad en aquellas componentes frecuenciales vinculadas a la respiración.

5) La AOS pediátrica provoca un aumento en la actividad simpática, lo cual se refleja en los patrones de la HRV durante la noche. En principio, se podría esperar que la actividad simpática fuera mayor durante la fase REM debido a que es cuando ocurren con mayor frecuencia los eventos apneicos. Sin embargo, la excitación simpática causada por los eventos apneicos parece enmascararse debido a la actividad simpática basal inherente a esta fase del sueño. Como resultado, se vuelve más viable distinguir las alteraciones asociadas con la AOS en el HRV al analizar información extraída durante la fase NREM. Por lo tanto, se resalta la importancia de considerar tanto la presencia de eventos apneicos como la fase del sueño de forma conjunta al evaluar el HRV en el contexto de la AOS pediátrica.

6) Los métodos de análisis de HRV desarrollados en la presente Tesis Doctoral han demostrado su utilidad para ayudar en el diagnóstico de la AOS pediátrica. Los modelos MLP, construidos usando subconjuntos óptimos de características biespectrales, alcanzaron el mejor rendimiento diagnóstico en los umbrales de severidad de IAH de 5 ($P = 81\,\%$, $AUC = 0{,}791$) y 10 e/h ($P = 91{,}70\,\%$, $AUC = 0{,}847$). Por otro lado, el modelo de LSBoost mostró el mayor rendimiento diagnóstico para el punto de corte de severidad de IAH de 1 e/h ($P = 80{,}07\,\%$, $AUC = 0{,}651$), resaltando la efectividad de la caracterización de HRV a nivel de segmento para detectar la presencia de AOS, incluso en los casos más leves.

Basándonos en las consideraciones expuestas, se puede concluir que la caracterización de la HRV mediante los nuevos índices desarrollados durante la realización de la presente Tesis Doctoral, y calculados en las nuevas bandas de frecuencia específicas de la AOS pediátrica, permiten una evaluación más precisa de las alteraciones en el SNA a causa de la AOS. Además, estos descubrimientos

tienen el potencial de contribuir a la simplificación del diagnóstico de la enfermedad en niños.

Bibliography

Acharya, U. R., Joseph, K. P., Kannathal, N., Lim, C. M., Suri, J. S., 2006. Heart rate variability: A review. Medical & Biological Engineering & Computing 44 (12), 1031–1051.

Adadi, A., Berrada, M., 2018. Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). IEEE Access 6, 52138–52160.

Aljadeff, G., Gozal, D., Schechtman, V. L., Burrell, B., Harper, R. M., Davidson Ward, S. L., 1997. Heart Rate Variability in Children With Obstructive Sleep Apnea. Sleep 20 (2), 151–157.

Allen, J., mar 2007. Photoplethysmography and its application in clinical physiological measurement. Physiological Measurement 28 (3), R1–R39.

Alonso-Álvarez, M. L., Terán-Santos, J., Ordax Carbajo, E., Cordero-Guevara, J. A., Navazo-Egüia, A. I., Kheirandish-Gozal, L., Gozal, D., apr 2015. Reliability of Home Respiratory Polygraphy for the Diagnosis of Sleep Apnea in Children. Chest 147 (4), 1020–1028.

Alpaydin, E., 2014. Introduction to Machine Learning, 3rd Edition. MIT Press.

Álvarez, D., Alonso-Álvarez, M. L., Gutiérrez-Tobal, G. C., Crespo, A., Kheirandish-Gozal, L., Hornero, R., Gozal, D., TerÁn-Santos, J., Del Campo, F., 2017. Automated screening of children with obstructive sleep apnea using nocturnal oximetry: An alternative to respiratory polygraphy in unattended settings. Journal of Clinical Sleep Medicine 13 (5), 693–702.

Álvarez, D., Hornero, R., Marcos, J. V., Wessel, N., Penzel, T., Glos, M., Del Campo, F., oct 2013. Assessment of feature selection and classification approaches to enhance information from overnight oximetry in the context of apnea diagnosis. International Journal of Neural Systems 23 (05), 1350020.

American Thoracic Society, feb 1996. Standards and indications for cardiopulmonary sleep studies in children. American Journal of Respiratory and Critical Care Medicine 153 (2), 866–878.

Atri, R., Mohebbi, M., sep 2015. Obstructive sleep apnea detection using spectrum and bispectrum analysis of single-lead ECG signal. Physiological Measurement 36 (9), 1963–1980.

Baharav, A., Kotagal, S., Rubin, B. K., Pratt, J., Akselrod, S., 1999. Autonomic cardiovascular control in children with obstructive sleep apnea. Clinical Autonomic Research 9 (6), 345–351.

Baron, R. M., Kenny, D. A., oct 1986. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. Journal of Personality and Social Psychology 51 (6), 1173–1182.

Barroso-García, V., Gutiérrez-Tobal, G. C., Kheirandish-Gozal, L., Álvarez, D., Vaquerizo-Villar, F., Núñez, P., del Campo, F., Gozal, D., Hornero, R., 2020. Usefulness of recurrence plots from airflow recordings to aid in paediatric sleep apnoea diagnosis. Computer Methods and Programs in Biomedicine 183, 105083.

Barroso-García, V., Gutiérrez-Tobal, G. C., Kheirandish-Gozal, L., Vaquerizo-Villar, F., Álvarez, D., del Campo, F., Gozal, D., Hornero, R., feb 2021. Bispectral analysis of overnight airflow to improve the pediatric sleep apnea diagnosis. Computers in Biology and Medicine 129, 104167.

Baumert, M., Pamula, Y., Martin, J., Kennedy, D., Ganesan, A., Kabir, M., Kohler, M., Immanuel, S. A., 2016. The effect of adenotonsillectomy for childhood sleep apnoea on cardiorespiratory control. ERJ Open Research 2 (2), 1–9.

Benitez, D., Gaydecki, P. A., Zaidi, A., Fitzpatrick, A. P., 2001. The use of the Hilbert transform in ECG signal analysis. Computers in Biology and Medicine 31 (5), 399–406.

Berry, R. B., Budhiraja, R., Gottlieb, D. J., Gozal, D., Iber, C., Kapur, V. K., Marcus, C. L., Mehra, R., Parthasarathy, S., Quan, S. F., Redline, S., Strohl, K. P., Ward, S. L. D., Tangredi, M. M., oct 2012. Rules for Scoring Respiratory Events in Sleep: Update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events. Journal of Clinical Sleep Medicine 08 (05), 597–619.

Berry, R. B., Quan, S. F., Abreu, A., 2020. The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications, Version 2.6. American Academy of Sleep Medicine.

Bhattacharjee, R., Kheirandish-Gozal, L., Pillar, G., Gozal, D., mar 2009. Cardiovascular Complications of Obstructive Sleep Apnea Syndrome: Evidence from Children. Progress in Cardiovascular Diseases 51 (5), 416–433.

Bishop, C., 2006. Pattern Recognition and Machine Learning. Springer, New York, NY.

Bonnet, M., Arand, D., may 1997. Heart rate variability: sleep stage, time of night, and arousal influences. Electroencephalography and Clinical Neurophysiology 102 (5), 390–396.

Bronzino, J. D., Peterson, D. R., 2014. Biomedical engineering fundamentals, 4th Edition. CRC press.

Bühlmann, P., Hothorn, T., nov 2007. Boosting Algorithms: Regularization, Prediction and Model Fitting. Statistical Science 22 (4), 477–505.

Bühlmann, P., Yu, B., jun 2003. Boosting With the L 2 Loss. Journal of the American Statistical Association 98 (462), 324–339.

Chiner, E., Cánovas, C., Molina, V., Sancho-Chust, J. N., Vañes, S., Pastor, E., Martinez-Garcia, M. A., jul 2020. Home Respiratory Polygraphy is Useful in the Diagnosis of Childhood Obstructive Sleep Apnea Syndrome. Journal of Clinical Medicine 9 (7), 2067.

Chua, K. C., feb 2009. Cardiac Health Diagnosis Using Higher Order Spectra and Support Vector Machine. The Open Medical Informatics Journal 3 (1), 1–8.

Chua, K. C., Chandran, V., Acharya, U. R., Lim, C. M., jan 2008. Cardiac state diagnosis using higher order spectra of heart rate variability. Journal of Medical Engineering & Technology 32 (2), 145–155.

Chua, K. C., Chandran, V., Acharya, U. R., Lim, C. M., 2010. Application of higher order statistics/spectra in biomedical signals-A review. Medical Engineering and Physics 32 (7), 679–689.

Church, G. D., jan 2012. The Role of Polysomnography in Diagnosing and Treating Obstructive Sleep Apnea in Pediatric Patients. Current Problems in Pediatric and Adolescent Health Care 42 (1), 2–25.

Cohen, G., de Chazal, P., 2015. Automated detection of sleep apnea in infants: A multi-modal approach. Computers in Biology and Medicine 63, 118–123.

Cohen, J., apr 1960. A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement 20 (1), 37–46.

Cohen, J., 1998. Statistical Power Analysis for the Behavioral Sciences. Routledge.

Constantin, E., McGregor, C. D., Cote, V., Brouillette, R. T., 2008. Pulse rate and pulse rate variability decrease after adenotonsillectomy for obstructive sleep apnea. Pediatric Pulmonology 43 (5), 498–504.

Crespo, A., Álvarez, D., Kheirandish-Gozal, L., Gutiérrez-Tobal, G. C., Cerezo-Hernández, A., Gozal, D., Hornero, R., del Campo, F., dec 2018. Assessment of oximetry-based statistical classifiers as simplified screening tools in the management of childhood obstructive sleep apnea. Sleep and Breathing 22 (4), 1063–1073.

Dehkordi, P., Garde, A., Karlen, W., Petersen, C. L., Wensley, D., Dumont, G. A., Mark Ansermino, J., 2016. Evaluation of cardiac modulation in children in response to apnea/hypopnea using the Phone Oximeter™. Physiological Measurement 37 (2), 187–202.

Deviaene, M., Testelmans, D., Buyse, B., Borzee, P., Van Huffel, S., Varon, C., mar 2019. Automatic Screening of Sleep Apnea Patients Based on the SpO 2 Signal. IEEE Journal of Biomedical and Health Informatics 23 (2), 607–617.

Efron, B., Tibshirani, R., may 1994. An Introduction to the Bootstrap. Chapman and Hall/CRC.

El-Hamad, F., Immanuel, S., Liu, X., Pamula, Y., Kontos, A., Martin, J., Kennedy, D., Kohler, M., Porta, A., Baumert, M., oct 2017. Altered Nocturnal Cardiovascular Control in Children With Sleep-Disordered Breathing. Sleep 40 (10).

Elith, J., Leathwick, J. R., Hastie, T., jul 2008. A working guide to boosted regression trees. Journal of Animal Ecology 77 (4), 802–813.

Fawcett, T., jun 2006. An introduction to ROC analysis. Pattern Recognition Letters 27 (8), 861–874.

Fleming, S., Thompson, M., Stevens, R., Heneghan, C., Plüddemann, A., MacOnochie, I., Tarassenko, L., Mant, D., 2011. Normal ranges of heart rate and respiratory rate in children from birth to 18 years of age: A systematic review of observational studies. The Lancet 377 (9770), 1011–1018.

Flemons, W. W., Littner, M. R., oct 2003. Measuring Agreement Between Diagnostic Devices. Chest 124 (4), 1535–1542.

Freund, Y., Schapire, R. E., aug 1997. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. Journal of Computer and System Sciences 55 (1), 119–139.

Friedman, J. H., mar 1989. Regularized Discriminant Analysis. Journal of the American Statistical Association 84 (405), 165–175.

Friedman, J. H., Meulman, J. J., may 2003. Multiple additive regression trees with application in epidemiology. Statistics in Medicine 22 (9), 1365–1381.

Garde, A., Dehkordi, P., Ansermino, J., Dumont, G., jun 2017. Correntropy-Based Pulse Rate Variability Analysis in Children with Sleep Disordered Breathing. Entropy 19 (6), 282.

Garde, A., Dehkordi, P., Karlen, W., Wensley, D., Ansermino, J. M., Dumont, G. A., 2014. Development of a screening tool for sleep disordered breathing in children using the phone oximeterTM. PLoS ONE 9 (11).

Garde, A., Hoppenbrouwer, X., Dehkordi, P., Zhou, G., Rollinson, A. U., Wensley, D., Dumont, G. A., Ansermino, J. M., aug 2019. Pediatric pulse oximetry-based OSA screening at different thresholds of the apnea-hypopnea index with an expression of uncertainty for inconclusive classifications. Sleep Medicine 60, 45–52.

Georgiou, K., Larentzakis, A. V., Khamis, N. N., Alsuhaibani, G. I., Alaska, Y. A., Giallafos, E. J., jan 2018. Can Wearable Devices Accurately Measure Heart Rate Variability? A Systematic Review. Folia Medica 60 (1).

Gil, E., Bailón, R., Vergara, J. M., Laguna, P., 2010. PTT variability for discrimination of sleep apnea related decreases in the amplitude fluctuations of PPG signal in children. IEEE Transactions on Biomedical Engineering 57 (5), 1079–1088.

Gil, E., Mendez, M., Vergara, J. M., Cerutti, S., Bianchi, A. M., Laguna, P., 2009. Discrimination of sleep-apnea-related decreases in the amplitude fluctuations of ppg signal in children by HRV analysis. IEEE Transactions on Biomedical Engineering 56 (4), 1005–1014.

Gnauck, A., oct 2004. Interpolation and approximation of water quality time series and process identification. Analytical and Bioanalytical Chemistry 380 (3), 484–492.

Goren, Y., Davrath, L. R., Pinhas, I., Toledo, E., Akselrod, S., 2006. Individual time-dependent spectral boundaries for improved accuracy in time-frequency analysis of heart rate variability. IEEE Transactions on Biomedical Engineering 53 (1), 35–42.

Gozal, D., Hakim, F., Kheirandish-Gozal, L., 2013. Chemoreceptors, baroreceptors and autonomic deregulation in children with obstructive sleep apnea. Respiratory Physiology and Neurobiology 185 (1), 177–185.

Guilleminault, C., Winkle, R., Connolly, S., Melvin, K., Tilkian, A., 1984. Cyclical Variation of the Heart Rate in Sleep Apnoea Syndrome. The Lancet 323 (8369), 126–131.

Gutiérrez-Tobal, G., Álvarez, D., Gomez-Pilar, J., del Campo, F., Hornero, R., jan 2015. Assessment of Time and Frequency Domain Entropies to Detect Sleep Apnoea in Heart Rate Variability Recordings from Men and Women. Entropy 17 (1), 123–141.

Gutierrez-Tobal, G. C., Alvarez, D., Crespo, A., del Campo, F., Hornero, R., mar 2019. Evaluation of Machine-Learning Approaches to Estimate Sleep Apnea Severity From At-Home Oximetry Recordings. IEEE Journal of Biomedical and Health Informatics 23 (2), 882–892.

Gutiérrez-Tobal, G. C., Álvarez, D., Marcos, J. V., Del Campo, F., Hornero, R., 2013. Pattern recognition in airflow recordings to assist in the sleep apnoea-hypopnoea syndrome diagnosis. Medical and Biological Engineering and Computing 51 (12), 1367–1380.

Gutiérrez-Tobal, G. C., Álvarez, D., Vaquerizo-Villar, F., Crespo, A., Kheirandish-Gozal, L., Gozal, D., del Campo, F., Hornero, R., nov 2021. Ensemble-learning regression to estimate sleep apnea severity using at-home oximetry in adults. Applied Soft Computing 111, 107827.

Gutiérrez-Tobal, G. C., Hornero, R., Álvarez, D., Marcos, J. V., Del Campo, F., 2012. Linear and nonlinear analysis of airflow recordings to help in sleep apnoea-hypopnoea syndrome diagnosis. Physiological Measurement 33 (7), 1261–1275.

Gutiérrez-Tobal, G. C., Álvarez, D., Kheirandish-Gozal, L., del Campo, F., Gozal, D., Hornero, R., aug 2022. Reliability of machine learning to diagnose pediatric obstructive sleep apnea: Systematic review and meta-analysis. Pediatric Pulmonology 57 (8), 1931–1943.

Guyon, I., Elisseeff, A., 2003. An introduction to variable and feature selection. Journal of Machine Learning Research 3 (Mar), 1157–1182.

Haddad, G. G., Jeng, H. J., Lai, T. L., Mellins, R. B., jun 1987. Determination of Sleep State in Infants Using Respiratory Variability. Pediatric Research 21 (6), 556–562.

Harper, R. M., Schechtman, V. L., Kluge, K. A., 1987. Machine classification of infant sleep state using cardiorespiratory measures. Electroencephalography and Clinical Neurophysiology 67 (4), 379–387.

Horne, R. S., Shandler, G., Tamanyan, K., Weichard, A., Odoi, A., Biggs, S. N., Davey, M. J., Nixon, G. M., Walter, L. M., jan 2018. The impact of sleep disordered breathing on cardiovascular health in overweight children. Sleep Medicine 41, 58–68.

Hornero, R., Kheirandish-Gozal, L., Gutiérrez-Tobal, G. C., Philby, M. F., Alonso-Álvarez, M. L., Alvarez, D., Dayyat, E. A., Xu, Z., Huang, Y. S., Kakazu, M. T., Li, A. M., Van Eyck, A., Brockmann, P. E., Ehsan, Z., Simakajornboon, N., Kaditis, A. G., Vaquerizo-Villar, F., Sedano, A. C., Capdevila, O. S., Von Lukowicz, M., Terán-Santos, J., Campo, F. D., Poets, C. F., Ferreira, R., Bertran, K., Zhang, Y., Schuen, J., Verhulst, S., Gozal, D., 2017. Nocturnal oximetry-based evaluation of habitually snoring children. American Journal of Respiratory and Critical Care Medicine 196 (12), 1591–1598.

Hunter, S. J., Gozal, D., Smith, D. L., Philby, M. F., Kaylegian, J., Kheirandish-Gozal, L., 2016. Effect of sleep-disordered breathing severity on cognitive performance measures in a large community cohort of young school-aged children. American Journal of Respiratory and Critical Care Medicine 194 (6), 739–747.

Iber, C., Ancoli-Israel, S., Chesson, A. L., Quan, S. F., 2007. The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications.

Imai, K., Keele, L., Tingley, D., 2010a. A General Approach to Causal Mediation Analysis. Psychological Methods 15 (4), 309–334.

Imai, K., Keele, L., Yamamoto, T., 2010b. Identification, inference and sensitivity analysis for causal mediation effects. Statistical Science 25 (1), 51–71.

Isaiah, A., Bertoni, D., Pereira, K. D., Diaz-Abad, M., Mitchell, R. B., Das, G., 2020. Treatment-Related Changes in Heart Rate Variability in Children with Sleep Apnea. Otolaryngology - Head and Neck Surgery 162 (5), 737–745.

Isler, J. R., Thai, T., Myers, M. M., Fifer, W. P., dec 2016. An automated method for coding sleep states in human infants based on respiratory rate variability. Developmental Psychobiology 58 (8), 1108–1115.

Ji-Wu Zhang, Chong-Xun Zheng, An Xie, mar 2000. Bispectrum analysis of focal ischemic cerebral EEG signal using third-order recursion method. IEEE Transactions on Biomedical Engineering 47 (3), 352–359.

Jiménez-García, J., Gutiérrez-Tobal, G. C., García, M., Kheirandish-Gozal, L., Martín-Montero, A., Álvarez, D., del Campo, F., Gozal, D., Hornero, R., jun 2020. Assessment of Airflow and Oximetry Signals to Detect Pediatric Sleep Apnea-Hypopnea Syndrome Using AdaBoost. Entropy 22 (6), 670.

Jobson, J. D., 2012. Applied multivariate data analysis: volume II: Categorical and Multivariate Methods. Springer Texts in Statistics. Springer New York, New York, NY.

Jon, C., 2009. Polysomnography in Children. In: Pediatric Otolaryngology for the Clinician. Humana Press, Totowa, NJ, pp. 35–47.

Kaditis, A., Kheirandish-Gozal, L., Gozal, D., jun 2016. Pediatric OSAS: Oximetry can provide answers when polysomnography is not available. Sleep Medicine Reviews 27, 96–105.

Kaditis, A. G., Chaidas, K., Alexopoulos, E. I., Varlami, V., Malakasioti, G., Gourgoulianis, K., 2011. Effects of adenotonsillectomy on R-R interval and brain natriuretic peptide levels in children with sleep apnea: A preliminary report. Sleep Medicine 12 (7), 646–651.

Kapur, V. K., Auckley, D. H., Chowdhuri, S., Kuhlmann, D. C., Mehra, R., Ramar, K., Harrod, C. G., mar 2017. Clinical Practice Guideline for Diagnostic Testing for Adult Obstructive Sleep Apnea: An American Academy of Sleep Medicine Clinical Practice Guideline. Journal of Clinical Sleep Medicine 13 (03), 479–504.

Katz, E. S., Lutz, J., Black, C., Marcus, C. L., apr 2003. Pulse Transit Time as a Measure of Arousal and Respiratory Effort in Children with Sleep-Disordered Breathing. Pediatric Research 53 (4), 580–588.

Katz, E. S., Mitchell, R. B., D'Ambrosio, C. M., apr 2012. Obstructive Sleep Apnea in Infants. American Journal of Respiratory and Critical Care Medicine 185 (8), 805–816.

Kheirandish-Gozal, L., 2010. What is "Abnormal" in pediatric sleep? Respiratory Care 55 (10), 1366–1374.

Kirk, V. G., Edgell, H., Joshi, H., Constantin, E., Katz, S. L., MacLean, J. E., dec 2020. Cardiovascular changes in children with obstructive sleep apnea and obesity after treatment with noninvasive ventilation. Journal of Clinical Sleep Medicine 16 (12), 2063–2071.

Kontos, A., Baumert, M., Lushington, K., Kennedy, D., Kohler, M., Cicua-Navarro, D., Pamula, Y., Martin, J., feb 2020. The Inconsistent Nature of Heart Rate Variability During Sleep in Normal Children and Adolescents. Frontiers in Cardiovascular Medicine 7 (February), 1–11.

Kwok, K. L., Yung, T. C., Ng, D. K., Chan, C. H., Lau, W. F., Fu, Y. M., 2011. Heart Rate Variability in Childhood Obstructive Sleep Apnea. Pediatric Pulmonology 46 (3), 205–210.

Lázaro, J., Gil, E., Vergara, J. M., Laguna, P., 2014. Pulse rate variability analysis for discrimination of sleep-apnea-related decreases in the amplitude fluctuations of pulse photoplethysmographic signal in children. IEEE Journal of Biomedical and Health Informatics 18 (1), 240–246.

Levene, H., 1960. Robust tests for equality of variances. Contributions to probability and statistics, 278–292.

Lewicke, A., Sazonov, E., Corwin, M., Neuman, M., Schuckers, S., jan 2008. Sleep Versus Wake Classification From Heart Rate Variability Using Computational Intelligence: Consideration of Rejection in Classification Models. IEEE Transactions on Biomedical Engineering 55 (1), 108–118.

Liao, D., Li, X., Rodriguez-Colon, S. M., Liu, J., Vgontzas, A. N., Calhoun, S., Bixler, E. O., 2010a. Sleep-disordered breathing and cardiac autonomic modulation in children. Sleep Medicine 11 (5), 484–488.

Liao, D., Li, X., Vgontzas, A. N., Liu, J., Rodriguez-Colon, S., Calhoun, S., Bixler, E. O., 2010b. Sleep-disordered breathing in children is associated with impairment of sleep stage-specific shift of cardiac autonomic modulation. Journal of Sleep Research 19 (2), 358–365.

Lilliefors, H., 1967. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. Journal of the American Statistical Association 62 (318), 399–402.

Liu, X., Immanuel, S., Kennedy, D., Martin, J., Pamula, Y., Baumert, M., 2018. Effect of adenotonsillectomy for childhood obstructive sleep apnea on nocturnal heart rate patterns. Sleep 41 (11), 1–8.

Lumeng, J. C., Chervin, R. D., 2008. Epidemiology of pediatric obstructive sleep apnea. Proceedings of the American Thoracic Society 5 (2), 242–252.

Luz Alonso-Álvarez, M., Canet, T., Cubell-Alarco, M., Estivill, E., Fernández-Julián, E., Gozal, D., Jurado-Luque, M. J., Lluch-Roselló, M. A., Martínez-Pérez, F., Merino-Andreu, M., Pin-Arboledas, G., Roure, N., Sanmartí, F. X., Sans-Capdevila, Ó., Segarra-Isern, F., Tomás-Vila, M., Terán-Santos, J., jan 2011. Consensus document on sleep apnea-hypopnea syndrome in children. Archivos de Bronconeumología 47 (5), 2–18.

Malik, M., Bigger, J. T., Camm, A. J., Kleiger, R. E., Malliani, A., Moss, A. J., Schwartz, P. J., mar 1996. Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. European Heart Journal 17 (3), 354–381.

Marcos, J. V., Hornero, R., Álvarez, D., del Campo, F., Zamarrón, C., 2009. Assessment of four statistical pattern recognition techniques to assist in obstructive sleep apnoea diagnosis from nocturnal oximetry. Medical Engineering and Physics 31 (8), 971–978.

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., Shiffman, R. N., Spruyt, K., 2012. Diagnosis and management of childhood obstructive sleep apnea syndrome. Pediatrics 130 (3), e714–e755.

Marcus, C. L., Carroll, J. L., Donnelly, D., Loughlin, G. M., 2008. Sleep in Children and Sleep and Breathing in Children: Developmental Changes in Sleep Patterns, 2nd Edition. CRC Press, New York.

Marcus, C. L., Moore, R. H., Rosen, C. L., Giordani, B., Garetz, S. L., Taylor, H. G., Mitchell, R. B., Amin, R., Katz, E. S., Arens, R., Paruthi, S., Muzumdar, H., Gozal, D., Thomas, N. H., Janice Ware, D. B., Snyder, K., Elden, L., Sprecher, R. C., Willging, P., Jones, D., Bent, J. P., Hoban, T., Chervin, R. D., Ellenberg, S. S., Redline, S., 2013. A randomized trial of adenotonsillectomy for childhood sleep apnea. New England Journal of Medicine 368 (25), 2366–2376.

Martín-González, S., Navarro-Mesa, J. L., Juliá-Serdá, G., Ramírez-Ávila, G. M., Ravelo-García, A. G., apr 2018. Improving the understanding of sleep apnea characterization using Recurrence Quantification Analysis by defining overall acceptable values for the dimensionality of the system, the delay, and the distance threshold. PLOS ONE 13 (4), e0194462.

Martín-Montero, A., Armañac-Julián, P., Gil, E., Kheirandish-Gozal, L., Álvarez, D., Lázaro, J., Bailón, R., Gozal, D., Laguna, P., Hornero, R., Gutiérrez-Tobal, G. C., mar 2023. Pediatric sleep apnea: Characterization of apneic events and sleep stages using heart rate variability. Computers in Biology and Medicine 154 (March), 106549.

Martín-Montero, A., Gutiérrez-Tobal, G. C., Gozal, D., Barroso-García, V., Álvarez, D., del Campo, F., Kheirandish-Gozal, L., Hornero, R., aug 2021a. Bispectral Analysis of Heart Rate Variability to Characterize and Help Diagnose Pediatric Sleep Apnea. Entropy 23 (8), 1016.

Martín-Montero, A., Gutiérrez-Tobal, G. C., Kheirandish-Gozal, L., Jiménez-García, J., Álvarez, D., del Campo, F., Gozal, D., Hornero, R., may 2021b. Heart rate variability spectrum characteristics in children with sleep apnea. Pediatric Research 89 (7), 1771–1779.

Martín-Montero, A., Gutiérrez-Tobal, G. C., Kheirandish-Gozal, L., Vaquerizo-Villar, F., Álvarez, D., del Campo, F., Gozal, D., Hornero, R., feb 2022. Heart rate variability as a potential biomarker of pediatric obstructive sleep apnea resolution. Sleep 45 (2), 1–9.

Mason, R. E., Likar, I., feb 1966. A new system of multiple-lead exercise electrocardiography. American Heart Journal 71 (2), 196–205.

Mejía-Mejía, E., Budidha, K., Abay, T. Y., May, J. M., Kyriacou, P. A., jul 2020. Heart Rate Variability (HRV) and Pulse Rate Variability (PRV) for the Assessment of Autonomic Responses. Frontiers in Physiology 11.

Michels, N., Clays, E., De Buyzere, M., Huybrechts, I., Marild, S., Vanaelst, B., De Henauw, S., Sioen, I., 2013. Determinants and reference values of short-term heart rate variability in children. European Journal of Applied Physiology 113 (6), 1477–1488.

Milagro, J., Gil, E., Lazaro, J., Seppa, V. P., Pekka Malmberg, L., Pelkonen, A. S., Kotaniemi-Syrjanen, A., Makela, M. J., Viik, J., Bailon, R., 2018. Nocturnal Heart Rate Variability Spectrum Characterization in Preschool Children with Asthmatic Symptoms. IEEE Journal of Biomedical and Health Informatics 22 (5), 1332–1340.

Milagro, J., Gracia, J., Seppa, V.-P., Karjalainen, J., Paassilta, M., Orini, M., Bailon, R., Gil, E., Viik, J., 2019. Noninvasive Cardiorespiratory Signals Analysis for Asthma Evolution Monitoring in Preschool Children. IEEE Transactions on Biomedical Engineering 9294 (c), 1–1.

Morelli, D., Rossi, A., Cairo, M., Clifton, D. A., jul 2019. Analysis of the Impact of Interpolation Methods of Missing RR-intervals Caused by Motion Artifacts on HRV Features Estimations. Sensors 19 (14), 3163.

Mukkamala, R., Hahn, J.-O., Inan, O. T., Mestha, L. K., Kim, C.-S., Toreyin, H., Kyal, S., aug 2015. Toward Ubiquitous Blood Pressure Monitoring via Pulse Transit Time: Theory and Practice. IEEE Transactions on Biomedical Engineering 62 (8), 1879–1901.

Muzumdar, H. V., Sin, S., Nikova, M., Gates, G., Kim, D., Arens, R., 2011. Changes in heart rate variability after adenotonsillectomy in children with obstructive sleep apnea. Chest 139 (5), 1050–1059.

Najarian, K., Splinter, R., apr 2016. Biomedical Signal and Image Processing. CRC Press, Boca Raton.

Narkiewicz, K., Somers, V. K., 1997. The sympathetic nervous system and obstructive sleep apnea: Implications for hypertension. Journal of Hypertension 15 (12 II), 1613–1619.

Nisbet, L. C., Yiallourou, S. R., Nixon, G. M., Biggs, S. N., Davey, M. J., Trinder, J., Walter, L. M., Horne, R. S., 2013. Nocturnal autonomic function in preschool children with sleep-disordered breathing. Sleep Medicine 14 (12), 1310–1316.

Nisbet, L. C., Yiallourou, S. R., Walter, L. M., Horne, R. S., 2014. Blood pressure regulation, autonomic control and sleep disordered breathing in children. Sleep Medicine Reviews 18 (2), 179–189.

O'Brien, L. M., Gozal, D., 2005. Autonomic dysfunction in children with sleep-disordered breathing. Sleep 28 (6), 747–752.

O'Driscoll, D. M., Foster, A. M., Ng, M. L., Yang, J. S. C., Bashir, F., Wong, S., Nixon, G. M., Davey, M. J., Anderson, V., Walker, A. M., Trinder, J., Horne, R. S. C., dec 2009. Central apnoeas have significant effects on blood pressure and heart rate in children. Journal of Sleep Research 18 (4), 415–421.

Pavone, M., Ullmann, N., Verrillo, E., De Vincentiis, G., Sitzia, E., Cutrera, R., 2017. At-home pulse oximetry in children undergoing adenotonsillectomy for obstructive sleep apnea. European Journal of Pediatrics 176 (4), 493–499.

Penzel, T., Kantelhardt, J., Grote, L., Peter, J., Bunde, A., 2003. Comparison of Detrended fluctuation Analysis and Spectral Analysis of Heart Rate Variability in Sleep and Sleep Apnea. IEEE Transactions on Biomedical Engineering 50 (10), 1143–1151.

Qin, H., Steenbergen, N., Glos, M., Wessel, N., Kraemer, J. F., Vaquerizo-Villar, F., Penzel, T., jul 2021. The Different Facets of Heart Rate Variability in Obstructive Sleep Apnea. Frontiers in Psychiatry 12 (July), 1–20.

Quante, M., Wang, R., Weng, J., Rosen, C. L., Amin, R., Garetz, S. L., Katz, E., Paruthi, S., Arens, R., Muzumdar, H., Marcus, C. L., Ellenberg, S., Redline, S., sep 2015. The Effect of Adenotonsillectomy for Childhood Sleep Apnea on Cardiometabolic Measures. Sleep 38 (9), 1395–1403.

Rangayyan, R., may 2015. Biomedical Signal Analysis. John Wiley & Sons, Inc., Hoboken, NJ, USA.

Redline, S., Amin, R., Beebe, D., Chervin, R. D., Garetz, S. L., Giordani, B., Marcus, C. L., Moore, R. H., Rosen, C. L., Arens, R., Gozal, D., Katz, E. S., Mitchell, R. B., Muzumdar, H., Taylor, H., Thomas, N., Ellenberg, S., 2011. The Childhood Adenotonsillectomy Trial (CHAT): Rationale, Design, and Challenges of a Randomized Controlled Trial Evaluating a Standard Surgical Procedure in a Pediatric Population. Sleep 34 (11), 1509–1517.

Ribeiro, M., Singh, S., Guestrin, C., 2016. "Why Should I Trust You?": Explaining the Predictions of Any Classifier. In: Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations. Association for Computational Linguistics, Stroudsburg, PA, USA, pp. 97–101.

Riedl, M., Müller, A., Kraemer, J. F., Penzel, T., Kurths, J., Wessel, N., 2014. Cardio-respiratory coordination increases during sleep apnea. PLoS ONE 9 (4), e93866.

Sazonova, N. A., Sazonov, E. E., Tan, B., Schuckers, S. A. C., study group, C., aug 2006. Sleep State Scoring in Infants from Respiratory and Activity Measurements. In: 2006 International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, pp. 2462–2465.

Schechtman, V. L., Harper, R. M., jan 1992. The Maturation of Correlations between Cardiac and Respiratory Measures Across Sleep States in Normal Infants. Sleep 15 (1), 41–47.

Shaffer, F., Ginsberg, J. P., sep 2017. An Overview of Heart Rate Variability Metrics and Norms. Frontiers in Public Health 5 (September), 1–17.

Shouldice, R. B., O'Brien, L. M., O'Brien, C., De Chazal, P., Gozal, D., Heneghan, C., 2004. Detection of Obstructive Sleep Apnea in Pediatric Subjects using Surface Lead Electrocardiogram Features. Sleep 27 (4), 784–792.

Smith, D. F., Amin, R. S., aug 2019. OSA and Cardiovascular Risk in Pediatrics. Chest 156 (2), 402–413.

Somers, V. K., Mark, A., Abboud, F., 1998. Sympathetic activation by hypoxia and hypercapnia–implications for sleep apnea. Clin Exp Hypertens 10 (1), 413–422.

Sörnmo, L., Laguna, P., 2005. Bioelectrical Signal Processing, 1st Edition. Elsevier Academic Press.

Standards of Practice Committee of the American Sleep Disorders Association, 1994. Practice parameters for the use of portable recording in the assessment of obstructive sleep apnea. Sleep 17 (4), 372–377.

Stein, P. K., Pu, Y., 2012. Heart rate variability, sleep and sleep disorders. Sleep Medicine Reviews 16 (1), 47–66.

Sullivan, G. M., Feinn, R., sep 2012. Using Effect Size—or Why the P Value Is Not Enough. Journal of Graduate Medical Education 4 (3), 279–282.

Tan, H.-L., Gozal, D., Ramirez, H. M., Bandla, H. P. R., Kheirandish-Gozal, L., 2014. Overnight polysomnography versus respiratory polygraphy in the diagnosis of pediatric obstructive sleep apnea. Sleep 37 (2), 255–260.

Tan, H.-L., Kheirandish-Gozal, L., Gozal, D., dec 2015. Pediatric Home Sleep Apnea Testing: Slowly Getting There! Chest 148 (6), 1382–1395.

Tauman, R., Gozal, D., jun 2011. Obstructive sleep apnea syndrome in children. Expert Review of Respiratory Medicine 5 (3), 425–440.

Terrill, P. I., Wilson, S. J., Suresh, S., Cooper, D. M., Dakin, C., aug 2012. Application of recurrence quantification analysis to automatically estimate infant sleep states using a single channel of respiratory data. Medical & Biological Engineering & Computing 50 (8), 851–865.

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., Imai, K., 2014. Mediation: R package for causal mediation analysis. Journal of Statistical Software 59 (5), 1–38.

Van Eyck, A., Van Hoorenbeeck, K., De Winter, B. Y., Van Gaal, L., De Backer, W., Verhulst, S. L., 2016. Sleep disordered breathing and autonomic function in overweight and obese children and adolescents. ERJ Open Research 2 (4), 1–8.

Vaquerizo-Villar, F., Álvarez, D., Kheirandish-Gozal, L., Gutiérrez-Tobal, G. C., Barroso-García, V., Crespo, A., del Campo, F., Gozal, D., Hornero, R., nov 2018a. Detrended fluctuation analysis of the oximetry signal to assist in paediatric sleep apnoea–hypopnoea syndrome diagnosis. Physiological Measurement 39 (11), 114006.

Vaquerizo-Villar, F., Álvarez, D., Kheirandish-Gozal, L., Gutiérrez-Tobal, G. C., Barroso-García, V., Crespo, A., del Campo, F., Gozal, D., Hornero, R., mar 2018b. Utility of bispectrum in the screening of pediatric sleep apnea-hypopnea syndrome using oximetry recordings. Computer Methods and Programs in Biomedicine 156, 141–149.

Vaquerizo-Villar, F., Alvarez, D., Kheirandish-Gozal, L., Gutierrez-Tobal, G. C., Barroso-Garcia, V., Santamaria-Vazquez, E., Campo, F. D., Gozal, D., Hornero, R., aug 2021. A Convolutional Neural Network Architecture to Enhance Oximetry Ability to Diagnose Pediatric Obstructive Sleep Apnea. IEEE Journal of Biomedical and Health Informatics 25 (8), 2906–2916.

Vaquerizo-Villar, F., Gozal, D., Hornero, R., Álvarez, D., Kheirandish-Gozal, L., Gutiérrez-Tobal, G. C., Barroso-García, V., Crespo, A., del Campo, F., 2018c. Wavelet analysis of oximetry recordings to assist in the automated detection of moderate-to-severe pediatric sleep apnea-hypopnea syndrome. PLoS ONE 13 (12), 1–18.

Vitelli, O., Del Pozzo, M., Baccari, G., Rabasco, J., Pietropaoli, N., Barreto, M., Villa, M. P., 2016. Autonomic imbalance during apneic episodes in pediatric obstructive sleep apnea. Clinical Neurophysiology 127 (1), 551–555.

Vlahandonis, A., Walter, L. M., Horne, R. S., feb 2013. Does treatment of SDB in children improve cardiovascular outcome? Sleep Medicine Reviews 17 (1), 75–85.

Vlahandonis, A., Yiallourou, S. R., Sands, S. A., Nixon, G. M., Davey, M. J., Walter, L. M., Horne, R. S., 2014. Long-term changes in heart rate variability in elementary school-aged children with sleep-disordered breathing. Sleep Medicine 15 (1), 76–82.

Walter, L. M., Biggs, S. N., Nisbet, L. C., Weichard, A. J., Hollis, S. L., Davey, M. J., Anderson, V., Nixon, G. M., Horne, R. S. C., mar 2016. Improved long-term autonomic function following resolution of sleep-disordered breathing in preschool-aged children. Sleep and Breathing 20 (1), 309–319.

Walter, L. M., Nixon, G. M., Davey, M. J., Anderson, V., Walker, A. M., Horne, R. S., 2013. Autonomic dysfunction in children with sleep disordered breathing. Sleep and Breathing 17 (2), 605–613.

Wang, R., Wang, J., Li, S., Yu, H., Deng, B., Wei, X., jan 2015. Multiple feature extraction and classification of electroencephalograph signal for Alzheimers' with spectrum and bispectrum. Chaos 25 (1), 013110.

Welch, P. D., 1967. The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms. IEEE Transactions on Audio and Electroacoustics AU-15 (2), 70–73.

Witten, I. H., Frank, E., Hall, M. A., 2011. Data Mining: Practical Machine Learning Tools and Techniques, third edit Edition. Elsevier.

Wu, Y., Tian, L., Ma, D., Wu, P., Tang, Y., Cui, X., Xu, Z., jun 2022. Autonomic nervous function and low-grade inflammation in children with sleep-disordered breathing. Pediatric Research 91 (7), 1834–1840.

Xu, Z., Gutiérrez-Tobal, G. C., Wu, Y., Kheirandish-Gozal, L., Ni, X., Hornero, R., Gozal, D., 2019. Cloud algorithm-driven oximetry-based diagnosis of obstructive sleep apnoea in symptomatic habitually snoring children. The European respiratory journal 53 (2).

Yu, L., Liu, H., 2004. Efficient Feature Selection via Analysis of Relevance and Redundancy. The Journal of Machine Learning Research 5, 1205–1224.

Zweig, M. H., Campbell, G., 1993. Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. Clinical Chemistry 39 (4), 561–577.

Şaylan, B., Çevik, A., Tavli, V., Vitrinel, A., 2011. Spectral and time-domain analyses of heart-rate variability in children with severe upper airway obstruction. Balkan Medical Journal 28 (2), 148–150.

Index

Pediatric obstructive sleep apnea (OSA) is a respiratory disease characterized by episodes of complete airflow cessation, known as apneas, or significant airflow reduction, termed hypopneas. This disorder has a high prevalence among children, affecting 5.7% of the pediatric population. When left untreated, it can lead to various adverse cardiovascular consequences, as well as cognitive and developmental impairments.

Currently, overnight polysomnography (PSG) serves as the gold standard for diagnosing pediatric OSA. Despite its usefulness, PSG has certain drawbacks, including being time-consuming, complex, costly, and uncomfortable for pediatric patients. Consequently, there has been a growing interest in exploring alternative techniques to simplify OSA diagnosis.

In this regard, the present Doctoral Thesis aims to comprehensively examine the behavior of heart rate variability (HRV) in children, seeking valuable insights into cardiac alterations associated with pediatric OSA that could aid in its diagnosis. Specifically, OSA-related HRV features in the spectral and bispectral domain are proposed to investigate cardiac behavior in the presence of OSA and assess the impact of OSA treatment. Additionally, the clinical utility of these new HRV features is evaluated by using machine-learning models for automated pediatric OSA diagnosis and sleep stage classification. In the light of the results, we can conclude that the newly proposed HRV features effectively capture alterations in the autonomic nervous system related to pediatric OSA. These features not only assist in evaluating the effects of OSA treatment but also contribute to the improved diagnosis of this condition.

UVa

Doctoral Dissertation
Compendium of publications
International Mention