USING OXIMETRY DYNAMICS TO SCREEN FOR SLEEP DISORDERED BREATHING AT VARYING THRESHOLDS OF SEVERITY

Ainara Garde[*], *Parastoo Dehkordi*[*], *David Wensley*[†], *J. Mark Ansermino*[*] *and Guy A. Dumont*[*]

[*]Electrical and Computer Engineering in Medicine Group, Department of Electrical and Computer Engineering, University of British Columbia
and the Pediatric Anesthesia Research Team, BC Children's Hospital
[†] Division of Critical Care, BC Children's Hospital

ABSTRACT

Sleep disordered breathing (SDB) is highly prevalent in children and causes daytime sleepiness, growth failure and developmental delay. Polysomnography (PSG), the gold standard to diagnose SDB, provides an estimate of severity, called the Apnea Hypoapnea Index (AHI). PSG is costly and resource intensive; therefore we propose using the *Phone Oximeter*, a pulse oximeter integrated into a phone that measures blood oxygen saturation (SpO_2), as an at-home screening tool. In clinical practice, an AHI of 2-5 indicates mild SDB, but an AHI $\geq$ 5 usually prompts SDB treatment. Thus, we studied the performance of the *Phone Oximeter* as an SDB screening tool at varying thresholds of AHI. We analyzed the SpO_2 of 146 children, recorded by the *Phone Oximeter*, alongside conventional PSG. Time-frequency characterization of SpO_2 dynamics resulted in identification of 77% and 86% of children with an AHI $\geq$ 2 and AHI $\geq$ 5, respectively, using a multiple logistic regression model.

Index Terms— Sleep disordered breathing, Apnea hypoapnea index, *Phone Oximeter*, Polysomnography, Pulse oximetry

1. INTRODUCTION

Sleep disordered breathing (SDB) describes a family of disorders characterized by frequent episodes of partial or complete cessation of breathing during sleep. Obstructive sleep apnea syndrome, the most common type of SDB in children, is characterized by upper airway obstruction that disturbs normal respiratory gas exchange [1]. Adenotonsillar hypertrophy is the most common cause of SDB in children [2], [3].

SDB is highly prevalent in pediatric patients (2% in children [4], [5]; 2.5%-6% in adolescents [6]) and poses a serious threat to healthy growth and development of these children. Pediatric SDB has important short- and long-term adverse effects, including daytime sleepiness, behavioural problems, neurocognitive impairment, heart failure, and developmental delay [2], [3], [7].

The gold standard for diagnosis of SDB is an overnight sleep study referred to as Polysomnography (PSG). PSGs are resource intensive (costing the healthcare system $500-1000/night [8]), and require that a child travel to and spend the night in a specialized sleep laboratory. This greatly limits accessibility for families who live outside of major cities, and restricts the number of patients who can be tested each year.

Pulse oximetry is a simple, non-invasive method of measuring blood oxygen saturation (SpO_2) and blood volume changes in tissue, using a photoplethysmographic (PPG) signal. The *Phone Oximeter* integrates a commercially available and Federal Drug Administration (FDA) approved microcontroller-based pulse oximeter (Masimo Set uSpO_2 Pulse Oximetry Cable) with a mobile smartphone [9]. In previous research, we demonstrated that the characterization of nocturnal SpO_2 recorded by the *Phone Oximeter* could identify children with SDB, with an AHI $\geq$ 5 considered to be a positive case for SDB [10].

There is no clear consensus on the appropriate AHI threshold to decipher between abnormal or mild SDB (AHI $\geq$ 1, AHI $\geq$ 2 or AHI $\geq$ 5). However, the majority of studies suggest that the threshold which indicates a need for SDB treatment (e.g., adenotonsillectomy), is a PSG-determined AHI $\geq$ 5. Therefore, in this study we aim to further investigate the potential of using the *Phone Oximeter* to identify children with SDB at varying AHI thresholds, through the characterization of SpO_2 dynamics.

2. MATERIAL AND METHODS

2.1. Dataset

Children (n=160) with signs of SDB, who had been referred to the British Columbia Children's Hospital for a PSG recording, were recruited following approval of the University of British Columbia and Children's and Women's Health Centre of British Columbia Research Ethics Board (H11-01769).

This work was supported in part by The Natural Sciences and Engineering Research Council of Canada (NSERC), the Canadian Institutes of Health Research (CIHR) through the Collaborative Health Research Projects Program and by a Mitacs Elevate Fellowship.

978-0-9928626-3-3/15/$31.00 ©2015 IEEE

Data acquisition was carried out in the Sleep Unit (a dedicated facility attached to the Medical Day Unit) where formal PSGs are performed. A standard PSG study includes the overnight measurement of electrocardiography (ECG), electroencephalography (EEG), pulse oximetry, chest movement, and nasal airflow, as well as video recording using the Embla Sandman S4500. Based on the American Academy of Sleep Medicine (AASM) criteria, a sleep technician annotated and scored the PSG recordings, providing the average of apnea/hypopnea events observed per hour, known as the apnea/hypoapnea index (AHI). As is standard clinically in children, OSA events were scored when the respiratory signal amplitude dropped by $\geq 90\%$ of pre-event baseline and lasted ≥ 2 breaths (based on baseline breathing during current sleep stage). Hypopnea events were scored when respiratory signal amplitude reduced $\geq 50\%$ relative to the 2 preceding breaths. Respiratory oxygen desaturation events were defined as oxygen desaturation decreases $\geq 3\%$. Both AHI ≥ 2 and AHI ≥ 5 are examined separately in this study as positive SDB thresholds, with a greater AHI indicating higher SDB severity.

A second pulse oximeter sensor was applied to the finger adjacent to the one used during a standard PSG. This sensor was attached to the *Phone Oximeter* and recorded SpO_2 simultaneously, at a sample frequency of 1 Hz, with a resolution of 0.1%. Fourteen children were excluded from analyses based on having a total sleep duration, or signal data duration (from PSG or the *Phone Oximeter*), shorter than 3 hours. After this exclusion, a total of 146 children were analyzed in this study. Table 1 summarizes the demographic and clinical data of the children who were included, as well the AHI index derived from the PSG diagnosis.

Table 1. Demographic and clinical information of the dataset represented as mean $\pm$ standard deviation:

	AHI < 2	$2 \leq$ AHI < 5	AHI ≥ 5
Children (n)	64	26	56
Age	9.1 ± 4.0	9.8 ± 4.1	8.8 ± 4.6
Male/Female	35/29	14/12	38/18
BMI	19.7 ± 5.1	20.8 ± 9.4	23.2 ± 8.3
AHI	0.8 ± 0.5	2.9 ± 0.7	19.7 ± 19.5

2.2. SpO_2 dynamics characterization

The overnight SpO_2 recorded by the *Phone Oximeter* was characterized in the time-frequency domain using a 2-minute sliding window with 1-minute overlap. The SpO_2 samples below 70% or above 100%, and sudden SpO_2 variations (more than 4% difference between consecutive samples) were considered artifacts and consequently removed from the signal. Segments with more than 50% of the signal labeled as artifacts were excluded from the analysis. The SpO_2 signal was characterized in the spectral domain through the power spectral density (PSD), and in the time domain using variability measures and indices related to desaturation episodes proposed in previous works to predict the presence of the SDB [11], [10]. Figure 1 illustrates an example of the overnight SpO_2 signal of three children with differing SDB severity.

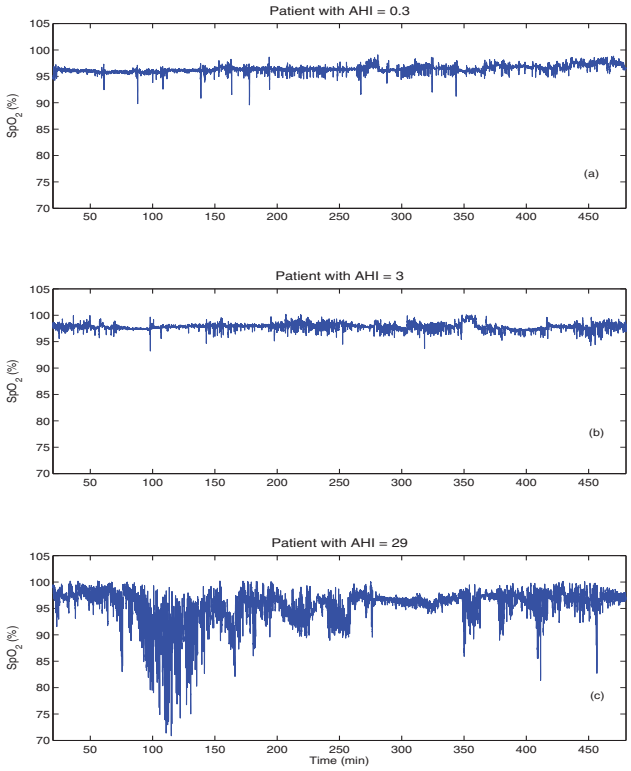

Fig. 1. Overnight SpO_2 signal of a children with (a) AHI < 2, (b) $2 \leq$ AHI < 5 and (c) AHI ≥ 5.

2.2.1. *SpO_2 characterization in the time domain*

Previously defined oximetry indices, such as the number of SpO_2 desaturations 3% and 4% below baseline, cumulative time spent below an SpO_2 of 96% and 94%, and Δ index were calculated for each time window [11], [10]. The Δ index previously proposed by [12] to quantify SpO_2 variability, was computed as the average of absolute differences of the mean SpO_2 between successive 12-sec intervals. The mean, median, standard deviation and interquartile range of the SpO_2 within each time window was also calculated. In addition, the Central Tendency Measure (CTM)—a non-linear method that provides quantitative variability information was also applied to the SpO_2 [13] signal.

2.2.2. *SpO_2 characterization in the frequency domain*

To account for respiratory pattern changes in the spectral domain, a time-varying spectral analysis was implemented. Using the 2-minute sliding time window, the SpO_2 signal was

divided into small segments that could be assumed to be stationary and therefore permitted computation of PSD. To provide a better frequency resolution, a parametric power spectral estimation was performed through autoregressive modeling. Figure 2 shows the spectral analysis of the SpO_2 signal applied to the same three children as above, again with different severities of SDB (see Figure 1).

Power spectral density (PSD) The signal $x(n)$ is modeled through an autoregressive model by

$$x(n) = -\sum_{k=1}^{p} a[k]x(n-k) + e(n) \qquad (1)$$

where $e(n)$ denotes zero-mean white noise with variance σ_e^2, $a[k]$ the AR coefficients and p the model order. Once the autoregressive coefficients and the variance σ_e^2 have been estimated, the PSD of an autoregressive process is computed by means of

$$\hat{P}_{x\,AR}(f) = \frac{\sigma_e^2}{|1 + \sum_{k=1}^{p} a[k] \cdot e^{-j \cdot 2\pi f k T}|^2} \cdot \qquad (2)$$

being T the sampling period. The selection of model order is a trade-off between the estimated prediction error and the spurious peaks (or overfitting). The optimum model order was evaluated according to Rissanen's minimum description length criterion ($p = 6$). A significant power increase in a frequency band ranging from 0.014 to 0.033 Hz was previously documented in subjects suffering from SDB, due to the modulation provoked by continuous oxygen desaturations [11], [10]. Therefore, several spectral parameters were extracted from the PSD: 1) the power in the discriminant frequency band, which consisted of a frequency interval (0.02 Hz) centered around the modulation frequency peak, tracked in the band from 0.005 to 0.1 Hz; 2) the total power of the PSD; 3) the ratio between the power of the discriminant frequency band and total power; and 4) the Shannon entropy of the PSD.

2.3. Data analysis

The described parameter set is computed for every 2-minute time window. In order to evaluate overnight SpO_2 dynamics, the mean (m) and standard deviation (std) of each parameter was studied. These features were included as predictors for a multiple logistic regression model. The Glmnet package for Matlab, implemented by Qian et al. [14], was applied to identify the individual features that distinguish between children with and without SDB, defined at varying AHI thresholds [15]. The most relevant features selected by LASSO (Least Absolute Shrinkage and Selection Operator), were applied to generate a multiple logistic regression model. Two models were created to identify children with an AHI ≥ 2 and children with an AHI ≥ 5, respectively. The Glmnet package

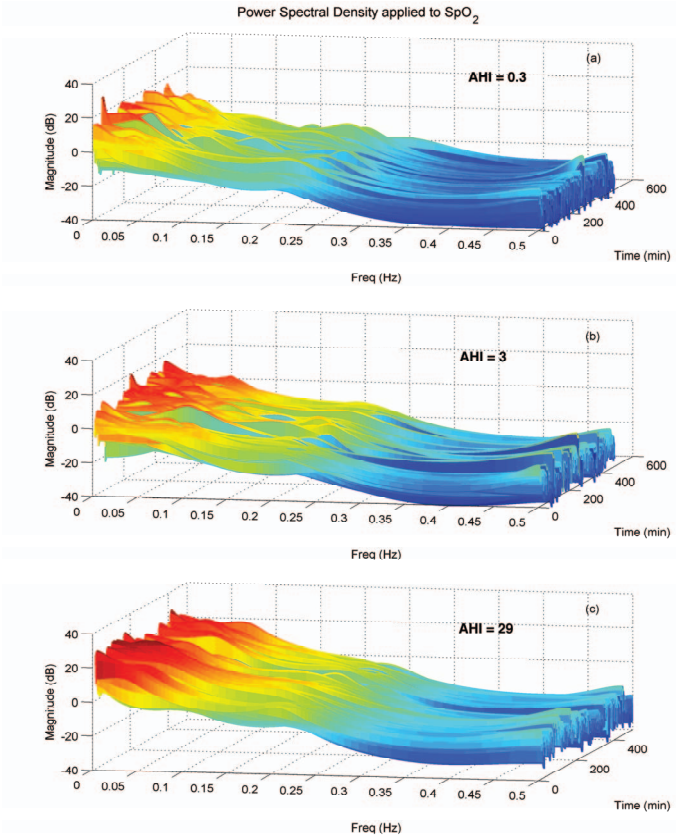

Fig. 2. (a) Power spectral density applied to the overnight SpO_2 signal of a children with (a) AHI <2, (b) $2\leq$ AHI <5 and (c) AHI ≥ 5.

was used to perform LASSO and evaluate the models' performance using a 10-fold cross-validation.

3. RESULTS

The SpO_2 dynamics of 146 children were analyzed and used to identify those with SDB, defined at varying AHI thresholds. The optimal multiple logistic regression model for identifying children with an AHI ≥ 2, while using a minimum number of features, was achieved with eight features. These features included the ratio between the power of the modulation frequency band and total power (m and std), the cumulative time spent below an SpO_2 of 94% (m and std), the Δ index (m), the inter quartile range of the SpO_2 (std), number of desaturations 3% below baseline (std) and the Shannon entropy of the PSD (std). This model correctly classified 76% of the children with and without an AHI ≥ 2. The Hosmer-Lemeshow goodness-of-fit test had a p-value of p=0.23, showing a suitable fit.

Fourteen features were selected by the multiple logistic regression model to identify children with an AHI ≥ 5. These features included the ratio between the power of the modu-

SDB	Accuracy (%)	Specificity (%)	Sensitivity (%)
AHI ≥ 2	75.9	75.0	76.8
AHI ≥ 5	83.4	81.1	85.7

Table 2. Results of the multiple logistic regression model using the most relevant parameter set to identify children with SDB at varying thresholds of AHI. The mean value of accuracy, specificity and sensitivity obtained with the 10-fold cross-validation is provided. Threshold values of 0.5 and 0.3 were used to identify children with an AHI ≥ 2 and ≥ 5, respectively.

lation frequency band and total power (m and std), the inter quartile range of the SpO_2 (m and std), the time spent below an SpO_2 of 94% and 96% (m and std), the number of desaturations 2% below baseline (m and std), the CTM of the SpO_2 (m), the total power of the PSD (std), the Shannon entropy of the PSD (std), the standard deviation of the SpO_2 (std), the Δ index (std) and the number of desaturations 3% below baseline (std). This model correctly classified 83% of the children with and without an AHI ≥ 5. The Hosmer-Lemeshow goodness-of-fit test had a p-value of p=0.68, showing an appropriate calibration.

Children with a higher AHI, showed significantly higher power in the modulation frequency band than children with a lower AHI (see figure 2). They also showed higher SpO_2 variability reflected by the Δ index, CTM, number of desaturations 2% and 3% below baseline and cumulative time spent below an SpO_2 of 94% and 96%. This is reflected in Figure 3 showing the distribution of some selected parameters. However, from the ROC curves obtained with each model (Figure 4), it can be observed that identifying children with SDB defined at a lower AHI threshold is more challenging.

4. DISCUSSION

This study shows that characterizing the dynamics of the SpO_2 pattern permits identification of children with SDB defined at varying AHI thresholds. It is feasible to detect children who should be referred to the hospital for PSG or treatment (AHI $\geq$ 5), but it becomes more challenging to identify cases with lower AHI. The most relevant features automatically selected were related mainly to the SpO_2 variability and modulation represented in the spectral domain. Children with higher AHI compared to children with lower AHI, showed greater SpO_2 variability and dispersion in the time domain, accompanied by higher SpO_2 spectral power at low frequencies (modulation band) in the frequency domain. These differences are less notable in children with mild forms of SDB ($2 \leq$ AHI < 5). From Figure 2, it can be observed that the power within the modulation frequency band increases in correlation with AHI and consequently with SDB severity. Therefore, children with mild forms of SDB ($2 \leq$ AHI < 5), showed a smaller power increase in the modulation frequency

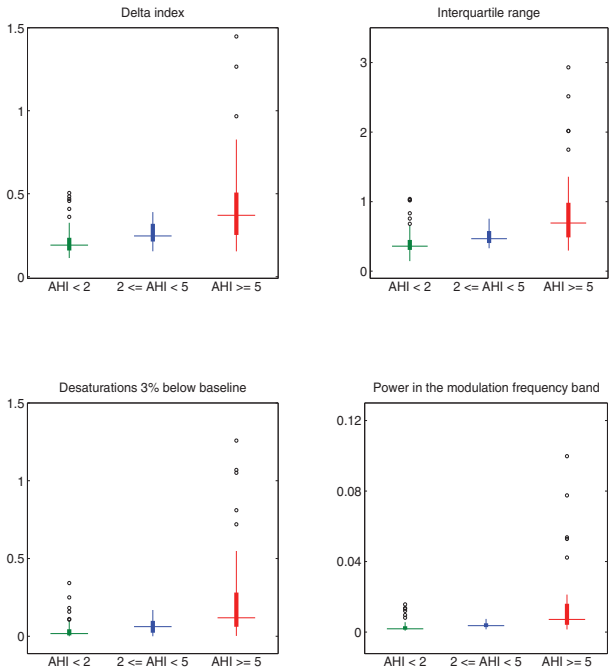

Fig. 3. Boxplot of the mean value of the (a) Δ index, (b) interquartile range of the SpO_2, (c) number of desaturations 3% below the $Sp0_2$ baseline value, and (d) power in the modulation frequency band defined as the frequency interval (0.02 Hz) centered around the modulation frequency peak, for children with different SDB severities (children with AHI $<$ 2 represented in green, $2 \leq$ AHI < 5 in blue, and AHI ≥ 5 in red. Quartile values are displayed as bottom, middle and top horizontal line of the boxes. Whiskers are used to represent the most extreme values within 1.5 times the interquartile range from the median. Outliers (data with values beyond the ends of the whiskers) are displayed as black circles.

band. The same effect can be observed in Figure 3 comparing the distribution of the parameters reflecting SpO_2 variability of children with SDB but low AHI and children without SDB. Thus, identifying all children in a sample with an AHI higher than 2 is clearly more challenging than identifying all children in a sample with an AHI higher than 5.

In conclusion, this SpO_2 pattern characterization will permit use of the *Phone Oximeter* as an at-home SDB screening tool, defining SDB at different AHI thresholds. This tool could be helpful in identifying children with suspected SDB, and inform monitoring and/or treatment plans.

5. ACKNOWLEDGEMENTS

The authors would like to thank the clinical staff of the sleep laboratory at British Columbia Children's Hospital for their collaboration and assistance with data acquisition, Aryannah Umedaly and Mika Johnson for helping to revise this manuscript and Guohai Zhou for his advice regarding sta-

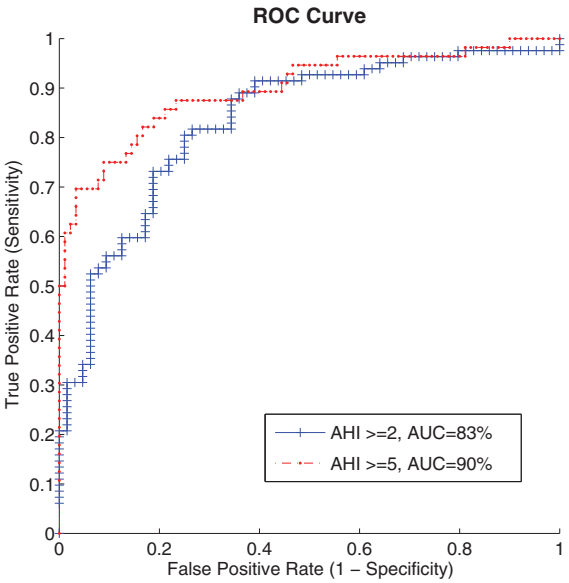

Fig. 4. The ROC obtained with the most discriminating feature set applied to the multiple logistic regression model to identify children with SDB, using 10-fold cross-validation. Represented in black the performance when SDB is defined as AHI ≥ 2, and in red when SDB is defined as AHI ≥ 5.

tistical analysis. The authors would also like to especially thank the Pediatric Anesthesia Research Team at the British Columbia Children's Hospital for their efforts in data collection and continuous support.

REFERENCES

[1] R T Brouillette, A Morielli, A Leimanis, K A Waters, R Luciano, and F M Ducharme, "Nocturnal pulse oximetry as an abbreviated testing modality for pediatric obstructive sleep apnea," *Pediatrics*, vol. 105, no. 2, pp. 405–412, Feb. 2000.

[2] C L Marcus, L J Brooks, A Kari A Draper, D Gozal, A C Halbower, J Jones, M S Schechter, S H Sheldon, K Spruyt, S D Ward, C Lehmann, and R N Shiffman, "Diagnosis and management of childhood obstructive sleep apnea syndrome.," *Pediatrics*, vol. 130, no. 3, pp. 576–584, 2012.

[3] C L Marcus, R H Moore, C L Rosen, B Giordani, S L Garetz, H G Taylor, R B Mitchell, R Amin, E S Katz, R Arens, S Paruthi, H Muzumdar, D Gozal, N H Thomas, J Ware, D Beebe, K Snyder, L Elden, R C Sprecher, P Willging, D Jones, J P Bent, T Hoban, R D Chervin, S S Ellenberg, and S Redline, "A randomized trial of adenotonsillectomy for childhood sleep apnea," *The New England Journal of Medicine*, vol. 368, no. 25, pp. 2366–2376, June 2013.

[4] C L Rosen, E K Larkin, H L Kirchner, J L Emancipator, S F Bivins, S A Surovec, R J Martin, and S Redline, "Prevalence and risk factors for sleep-disordered breathing in 8- to 11-year-old children: association with race and prematurity," *The Journal of Pediatrics*, vol. 142, no. 4, pp. 383–389, Apr. 2003.

[5] J H Wildhaber and A Moeller, "Sleep and respiration in children: time to wake up!," *Swiss Medical Weekly*, vol. 137, no. 49-50, pp. 689–694, Dec. 2007.

[6] E O Johnson and R Thomas, "An epidemiologic study of sleep-disordered breathing symptoms among adolescents," *Sleep*, vol. 29, no. 9, pp. 1135–1142, Sept. 2006.

[7] Section on Pediatric Pulmonology and Subcommittee on Obstructive Sleep Apnea Syndrome, "Clinical practice guideline: Diagnosis and management of childhood obstructive sleep apnea syndrome," *Pediatrics*, vol. 109, no. 4, pp. 704–712, 2002.

[8] Medical Services Commission of British Columbia, "Respirology," 2013.

[9] W Karlen, G A Dumont, C Petersen, J Gow, J Lim, J Sleiman, and M J Ansermino, "Human-centered phone oximeter interface design for the operating room," in *International Conference on Health Informatics*, 2011, pp. 433–438.

[10] A Garde, W Karlen, P Dehkordi, D Wensley, J M Ansermino, and G A Dumont, "Oxygen saturation in children with and without obstructive sleep apnea using the phone-oximeter," in *International Conference of the IEEE Engineering in Medicine and Biology Society*, 2013, pp. 2531–2534.

[11] D Alvarez, R Hornero, J V Marcos, N Wessel, T Penzel, M Glos, and F Del Campo, "Assessment of feature selection and classification approaches to enhance information from overnight oximetry in the context of apnea diagnosis," *International Journal of Neural Systems*, vol. 23, pp. 1350020, 2013.

[12] P Levy, J L Pepin, C Deschaux-Blanc, B Paramelle, and C Brambilla, "Accuracy of oximetry for detection of respiratory disturbances in sleep apnea syndrome," *Chest*, vol. 109, no. 2, pp. 395–399, 1996.

[13] J Jeong, J C Gore, and B S Peterson, "A method for determinism in short time series, and its application to stationary eeg," *IEEE Trans. Biomed. Engineering*, vol. 49, no. 11, pp. 1374–1379, 2002.

[14] J Qian, T Hastie, J Friedman, R Tibshirani, and N Simon, "Glmnet for matlab," .

[15] J H Friedman, T Hastie, and R Tibshirani, "Regularization paths for generalized linear models via coordinate descent," *Journal of Statistical Software*, vol. 33, no. 1, pp. 1–22, 2 2010.