ABSTRACT

Title of Thesis:	THE USE OF ACTIGRAPHY FOR RISK STRATIFICATION IN PEDIATRIC OBSTRUCTIVE SLEEP APNEA
	Dylan Bertoni, Master of Science, 2018
Thesis Directed By:	Dr. Amal Isaiah, Department of Otorhinolaryngology—Head and Neck Surgery
	Professor John P. Fisher, Department of Bioengineering

Objectives. (i) To determine the feasibility of using actigraphy to identify sleep interruption in children with suspected obstructive sleep apnea (OSA); (ii) to assess the utility of actigraphy for prediction of OSA severity.

Subjects and Methods. Ten healthy children aged 2 to 15 years with suspicion for OSA underwent polysomnography (PSG) with actigraphy. Statistical learning algorithms were used to (i) identify sleep-related respiratory events (awake, asleep, hypopneas, and apneas), and (ii) predict OSA severity (mild, moderate, and severe) utilizing actigraphy counts.

Results. No adverse events were identified. Actigraphy counts were obtained in all 10 children. Linear discriminant analysis identified 100% of patients with severe OSA. Actigraphy counts reliably identified hypopneas and awakenings but not apneas.

Conclusions. Actigraphy counts may provide effective risk stratification for pediatric OSA. Further investigations are necessary to investigate the utility of using actigraphy and pulse oximetry together to identify all respiratory events during sleep.

THE USE OF ACTIGRAPHY FOR RISK STRATIFICATION IN PEDIATRIC OBSTRUCTIVE SLEEP APNEA

by

Dylan G. Bertoni

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Science
2018

Advisory Committee:
Dr. Amal Isaiah, Co-Chair
Professor John P. Fisher, Co-Chair
Dr. Gregg Duncan

Acknowledgements

I would like to sincerely thank Dr. Amal Isaiah for the immense effort he has put into being an advisor and mentor to me, both as a bioengineering student and as a student doctor. I would also like to thank Dr. Fisher for all of his help and advice throughout this year. I would like to thank Valerie Rogers, Ph.D. for providing Actiwatch-2® devices, Libby Nichols for her help in collecting data, Dr. Kevin Pereira and Judith Oakley, N.P. for their assistance in recruiting patients, and the University of Maryland Department of Otorhinolaryngology—Head and Neck Surgery for their support throughout this project.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

SDB—sleep disordered breathing

PS—primary snoring

OSA—obstructive sleep apnea

PSG—polysomnography

PAP—positive airway pressure

T&A—tonsillectomy and adenoidectomy

AHI—apnea-hypopnea index

PRAE—perioperative respiratory adverse event

AAO—HNS—American Academy of Otolaryngology—Head and Neck Surgery

BMI—body mass index

PSQ—pediatric sleep questionnaire

HSAT—home sleep apnea test

TST—total sleep time

TIB—time in bed

SE—sleep efficiency

SL—sleep latency

WASO—wake after sleep onset

ODI—oxygen desaturation index

MEMS—miniaturized poly-silicone-surface micromachined sensor

UMMC—University of Maryland Medical Center

ML—machine learning

LDA—linear discriminant analysis

ROC—receiver operating characteristic

Chapter 1: Introduction

Background

Sleep Disordered Breathing

Sleep disordered breathing (SDB) in children is a spectrum of abnormal breathing patterns, characterized by a limitation of airflow during sleep. Upper airway resistance and collapse of the pharynx can cause varying degrees of obstruction, which can be exacerbated by the presence of adenotonsillar hypertrophy, obesity, and craniofacial or neuromuscular disorders. The spectrum of SDB ranges from primary snoring (PS), to upper airway resistance syndrome (UARS), to the most severe diagnosis of obstructive sleep apnea (OSA).[1] Obstructive sleep apnea can be further stratified into mild, moderate, and severe categories. This spectrum is based on the frequency and severity of airflow obstructions, arousals, and gas exchange abnormalities during sleep.[1]

The prevalence of sleep disordered breathing has increased over the past 2 decades, associated with increases in obesity rates among children. A meta-analysis estimated the prevalence of SDB to be about 7.45%, with an underlying assumption that a large proportion of children with SDB remain undiagnosed.[2] Studies indicate that SDB has been shown to have increased prevalence among obese children,[3] African-American children,[3] children born prematurely,[4] and those exposed to cigarette smoke.[5]

Sleep disordered breathing is diagnosed clinically by history and physical examination as the initial method to establish the site of obstruction. Habitual snoring, weight and body habitus, tonsil size, tongue position, mouth breathing, drooling, neuromuscular deficits, and other findings can all indicate the presence of SDB.[6] Definitive diagnosis of the type of SDB is determined by polysomnography (PSG) or sleep study.

Treatment of SDB depends on the etiology of obstruction. Pharmacotherapy, such as intranasal steroids to reduce upper airway inflammation, can be used to treat mild disease.[7] Positive airway pressure (PAP) can be an effective therapy, although the compliance among children is poor.[8] Orthodontic treatments such as rapid maxillary distraction (RMD), have also been tested with varying degrees of effectiveness, depending on compliance and the type of SDB present.[9] Surgical therapy remains the first-line treatment for more severe disease.

Pediatric Obstructive Sleep Apnea Syndrome

Pediatric OSA is the most severe form of SDB and is characterized by increased upper airway resistance, followed by intermittent partial or complete upper airway obstruction, which disrupts normal ventilation and sleep patterns. The prevalence of OSA is estimated to be up to 5% in children.[2] The principal risk factor for pediatric OSA is the pathologic enlargement of the lymphoid tissue within the tonsils and adenoids,[10] as seen in Fig. 1. Other less common risk factors for OSA include craniofacial dysmorphisms and neuromuscular abnormalities. In children with OSA,

enlarged lymphoid tissue leads to increased airway resistance and collapse, and therefore, the standard of treatment is a tonsillectomy and adenoidectomy (T&A).[11]

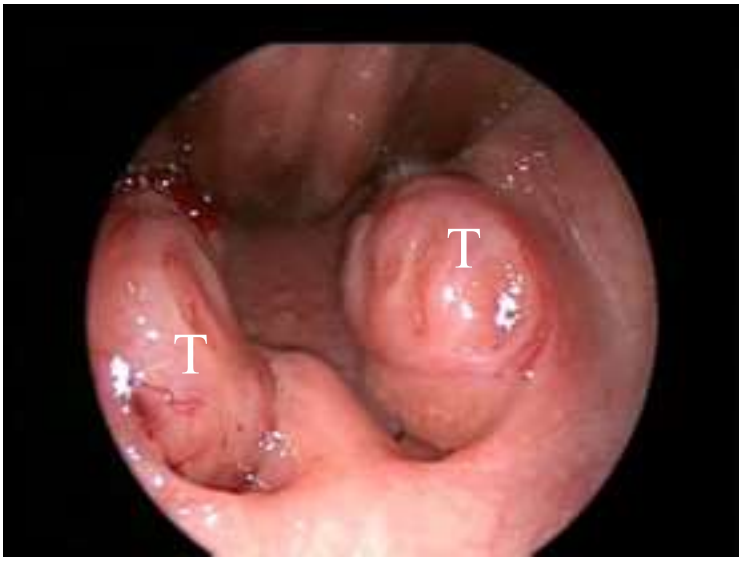

Figure 1: Image of the palatine tonsils (*T*) showing restricted dimensions of the oropharyngeal airway (picture courtesy of Dr. Amal Isaiah).

Pediatric OSA is characterized by repeated episodes of partial or complete airway obstruction during sleep, resulting in a cycle of hypoxemia, respiratory compensation, heart rate variability, and awakening.[6] Left undiagnosed or untreated, the earliest adverse effects of OSA include neurobehavioral problems[12] and cognitive dysfunction.[13] Currently, the gold standard for diagnosis is an overnight laboratory-based PSG.[14] The principal metric derived from a PSG is the apnea-hypopnea index (AHI), which reflects the severity and frequency of airflow limitation. Although there is no universally accepted grading system for OSA, an AHI greater than 1 but less than 5 is considered mild, greater than or equal to 5 but less than 10 is considered moderate, and a value greater than or equal to 10 is considered severe.[15]

Tonsillectomy and adenoidectomy results in relief of upper airway obstruction and

significant improvement or resolution of symptoms in the majority of children with OSA.[16] With over 500,000 performed each year in the United States, T&A is one of the most common surgical procedures performed in children.[17]

Due to the effects of chronic airway obstruction and hypoxemia on brainstem respiratory control mechanisms, children with OSA are more prone to perioperative respiratory adverse events (PRAEs), and may require overnight admission following surgery.[18] The most recent guidelines from the American Academy of Otolaryngology—Head and Neck Surgery (AAO-HNS) recommend overnight monitoring for children with AHI $\geq$ 10.[19] Although PSG is helpful in the identification and stratification of children with OSA prior to T&A, obtaining a PSG for every child undergoing the procedure is not feasible. This is due to the enormous need for financial, personnel, and technical resources. In fact, only about 10% of children undergoing T&A receive a preoperative PSG.[20] In addition, children without risk factors for OSA may have unexpectedly severe OSA (AHI $\geq$ 30), emphasizing the need for preoperative risk stratification.[21] Therefore, an ideal screening process balances cost-effectiveness and improved risk stratification of OSA so that eligible children may undergo further objective assessment utilizing PSG.

Sequelae of Untreated OSA

Many short- and long-term sequelae associated with untreated OSA in children support prompt diagnosis and initiation of early treatment. Cha et al.[22] demonstrated significant correlation between AHI and mean diffusivity in select regions within the brain. For example, the dentate gyrus is a component of the hippocampal formation,

and contributes to several higher brain functions including learning, memory, and spatial coding. A lower mean diffusion of water within the dentate gyrus, as measured by diffusion tensor imaging, correlated with AHI and was associated with a lower verbal learning and memory score on neuropsychological assessments. Gatica et al.[23] established a negative correlation between sleep-related breathing disorders and academic performance in young children. Significant associations were observed between sleep-related breathing disorders and low performances in math, language arts, and science. Furthermore, a study by Landau et al.[24] found that preschool-aged children with OSA demonstrated significantly impaired executive function, attention, and receptive vocabulary; had more behavioral problems; and scored lower on a health-related quality of life questionnaire than matched controls.

Long-term OSA is associated with sequelae beyond neuropsychiatric and behavioral issues. Nieminen et al.[25] found that children with OSA had significant impairments in secretions of different growth hormones. Adenotonsillectomy resulted in weight gain and restored hormone secretion. Peña-Zarza et al.[26] found that increasing glycated hemoglobin (Hb_{A1C}), which is a useful homeostatic biomarker of blood-glucose levels, was associated with increasing AHI. High blood glucose levels are pathologic for diabetes. Although both diabetes and OSA have associations with obesity, the correlations between Hb_{A1C} and AHI remained significant when adjusting by age and body mass index (BMI).

Diagnosis of OSA

History and Physical Examination

A summary of the presenting signs and symptoms of OSA is provided in Table 1 below. Adolescents and older children may be able to recognize these problems and articulate complaints, whereas these symptoms may be reported by parents of younger children.

Nighttime symptoms	Frequent snoring (≥3 nights per week) Labored breathing during sleep Gasps/snorting noises/observed apneas Sleep enuresis (involuntary urination) Nocturnal sweating Sleep walking Sleep talking Sleep terrors Crying spells Drooling Mouth-breathing Insomnia Sleeping in seated position/with neck hyperextended Cyanosis Headache on awakening
Daytime symptoms	Daytime sleepiness Mouth-breathing Difficulty waking in the morning Poor eating/failure to thrive Symptoms attributable to attention deficit hyperactivity disorder (ADHD) Delayed puberty Learning problems Behavioral problems
Physical exam findings	Obesity/above average BMI Elevated blood pressure Enlarged tonsils

Table 1: Summary of presenting signs and symptoms of OSA.[6,27]

A number of studies have assessed the utility of history and physical examination findings, together referred to as clinical parameters, in predicting both the presence and the severity of OSA. For example, Wang et al.[28] attempted to quantify the relationship between clinical findings and OSA, with predictive accuracy never exceeding 50%. These findings were additionally confirmed by a subsequent meta-analysis. In the childhood adenotonsillectomy trial (CHAT),[15] a randomized-controlled trial comparing the benefit of early T&A and watchful waiting with supportive care, risk factors for OSA could be identified, but the association between these parameters and the severity of OSA could not be established. Furthermore, the use of validated instruments such as the pediatric sleep questionnaire[29] and the OSA score[30] have had some utility as research tools to determine the risk of OSA, although they have not universally supplanted objective testing.

Screening Instruments

A number of screening questionnaires have been developed to screen children for OSA. Brouilette et al.[30] used a standardized questionnaire that demonstrated significantly increased frequencies of the following symptoms when compared to age- and sex-matched controls: difficulty breathing during sleep, apnea observed by the parents, snoring, restless sleep, chronic rhinorrhea (runny nose), and mouth breathing while awake. They developed the OSA-score using discriminant analysis. Scores over 3.5 were highly predictive of OSA requiring T&A, no child with a score less than -1 had OSA, and children with scores between -1 and 3.5 required PSG. Subsequent studies, however, have demonstrated that the OSA-score is limited in its ability to distinguish PS from OSA.[31]

Chervin et al.[32] developed the Pediatric Sleep Questionnaire (PSQ), a 22-item questionnaire in which scores range from 0 to 1, with higher scores indicating greater severity. The PSQ items address snoring frequency and quality, breathing problems, mouth breathing, daytime sleepiness, inattention/hyperactivity, and other symptoms to identify children with SDB. They used logistical regression models to identify question-items most suitable for use in identifying SDB. Both sensitivity and specificity were high (0.85 and 0.87 respectively) when 8 or more positive answers to the 22 question items were considered abnormal. Although the PSQ can be used to identify patients at risk for OSA, it alone cannot diagnose or distinguish between mild, moderate and severe OSA.

The OSA-18 quality-of-life questionnaire (OSA-18) is an 18-item questionnaire that uses a Likert-type scoring system to collect information considered to be elements in quality of life: sleep disturbance, physical symptoms, emotional symptoms, daytime function, and caregiver concerns. The output is a summary score that ranges from 18 (no impact on quality of life) to 126 (major negative impact).[33] While it can be helpful in aiding clinical diagnosis, the OSA-18 does not accurately detect which children have abnormal oximetry scores, and cannot be used in place of objective testing to identify moderate-to-severe OSA in children.[34]

Home Sleep Apnea Testing (HSAT)

Pulse oximetry has been explored as a diagnostic tool for OSA, as it provides information about oxygen desaturation during sleep, and can be analyzed quickly and easily. Kaditis et al.[35] performed a systematic review of 25 original articles to

summarize reference values of nocturnal oximetry parameters in healthy children, as well as to identify abnormal oximetry patterns that predict OSA in snoring children and treatment response or potential complications. They found that nocturnal pulse oximetry (SpO_2) drops below 90% and clusters of more than two desaturation events ≥ 4% were unusual in children without OSA. At least three clusters of desaturation events and three SpO_2 drops below 90% are indicative of moderate-to-severe OSA. Children without clusters of desaturation events have low risk of major respiratory complications following T&A. Pulse oximetry can therefore be a valuable tool for evaluating OSA when PSG is not available.

Lamm et al.[36] explored the use of home audiotaping—15 minute recordings of the child's sleep and breathing to distinguish between PS and OSA. Seven independent observers analyzed the recordings for the presence of a struggle sound and respiratory pauses. The mean sensitivity was 0.71, and the mean specificity was 0.80. Home audiotapes are suggestive of OSA, but not sufficiently specific to reliably distinguish PS from OSA.

Jacob et al.[37] investigated the accuracy and practicality of performing home sleep studies in 21 children. These home sleep studies included cardiorespiratory recordings of SpO_2, pulse rate, pulse wave form, electrocardiogram, and respiratory inductive plethysmography (movement of the chest wall) in addition to an 8-hour recording of the child sleeping. These home sleep studies were validated against PSG. They found saturation, respiratory, and video data were obtained 96.4%, 99.4%, and

90.0% of the time, respectively. They also found that the sleep efficiency was 5% greater at home, while study duration, AHI, spontaneous movement/arousal index, and oxygen saturation were similar for home and laboratory studies. Home sleep studies can be accurate and practical in evaluating more severe OSA, but may fail to identify the presence of associated sleep disorders and partially obstructed breaths. Negative results do not rule out the diagnosis of OSA and should be followed with a PSG.[38] Home PSG studies in children also lack standardization, making it difficult to study their effectiveness on a large scale.

Due to insufficient data for validation, the American Academy of Sleep Medicine (AASM) has released position statements that assert that HSAT can provide valuable information for the diagnostic assessment of adult patients with suspected OSA,[39] but the use of an HSAT is not recommended for the diagnosis of OSA in children.[40]

Polysomnography

Polysomnography is the electrographic recording of simultaneous physiologic variables during sleep. There is currently no consensus on when children with clinical suspicion for OSA are recommended to have PSG. The AAO-HNS recommends that children be referred for PSG prior to T&A if they exhibit complex medical conditions such as obesity, Down syndrome, craniofacial abnormalities, neuromuscular disorders, sickle cell disease, or mucopolysaccharidoses. A PSG should also be sought if surgery is uncertain, or if there is discordance between tonsillar size on physical exam and the reported severity of the SDB.[19] The American Academy of Pediatrics (AAP) recommends all children/adolescents with snoring and symptoms or

signs of OSA be referred for PSG. If no PSG is available, an alternative diagnostic test should be performed, or the patient should be referred to a sleep specialist or otolaryngologist for further evaluation.[11]

Polysomnography is performed at certified sleep centers where trained technicians administer the sleep studies and generate reports. Trained and certified sleep technicians score the reports, and a chief sleep physician provides quality review. Polysomnography measures a large number of variables, including: sleep stage, muscle tone, thoracoabdominal movement, airflow and nasal pressure, oxygen saturation, leg and chin movement, body position, and cardiac rhythm. A sample PSG is provided in Fig. 2, and explanations of measured variables as well as those interpreted by sleep technicians are provided in Tables 2 and 3.

An apneic event is complete closure of the airway or cessation of breathing, while a hypopneic event is the partial closure of the airway or cessation of breathing. AHI is defined as the sum of all obstructive and mixed apneas and hypopneas divided by the total hours of sleep time. As mentioned previously, AHI is the primary parameter used to diagnose and determine the severity of OSA, as well as the treatment course, with mild disease more likely to be treated with intranasal corticosteroids[7] and severe disease treated with surgery.

Sleep studies are not universally available or accessible to patients due to shortages of the necessary certified sleep centers and trained personnel. Additionally, Horwood et

al.[41] estimated the cost for testing 1,000 patients by PSG to be approximately $1,000,000. Thus, there is a significant financial barrier to access. Providing every child suspected of SDB with a PSG is financially untenable. As a result of these factors, only 10% of all children suspected of SDB undergo preoperative PSGs prior to T&A.[20] Thousands of children at risk for potential PRAEs are therefore not stratified prior to surgery. Therefore the diagnostic methodology for pediatric OSA is in need of improvement.

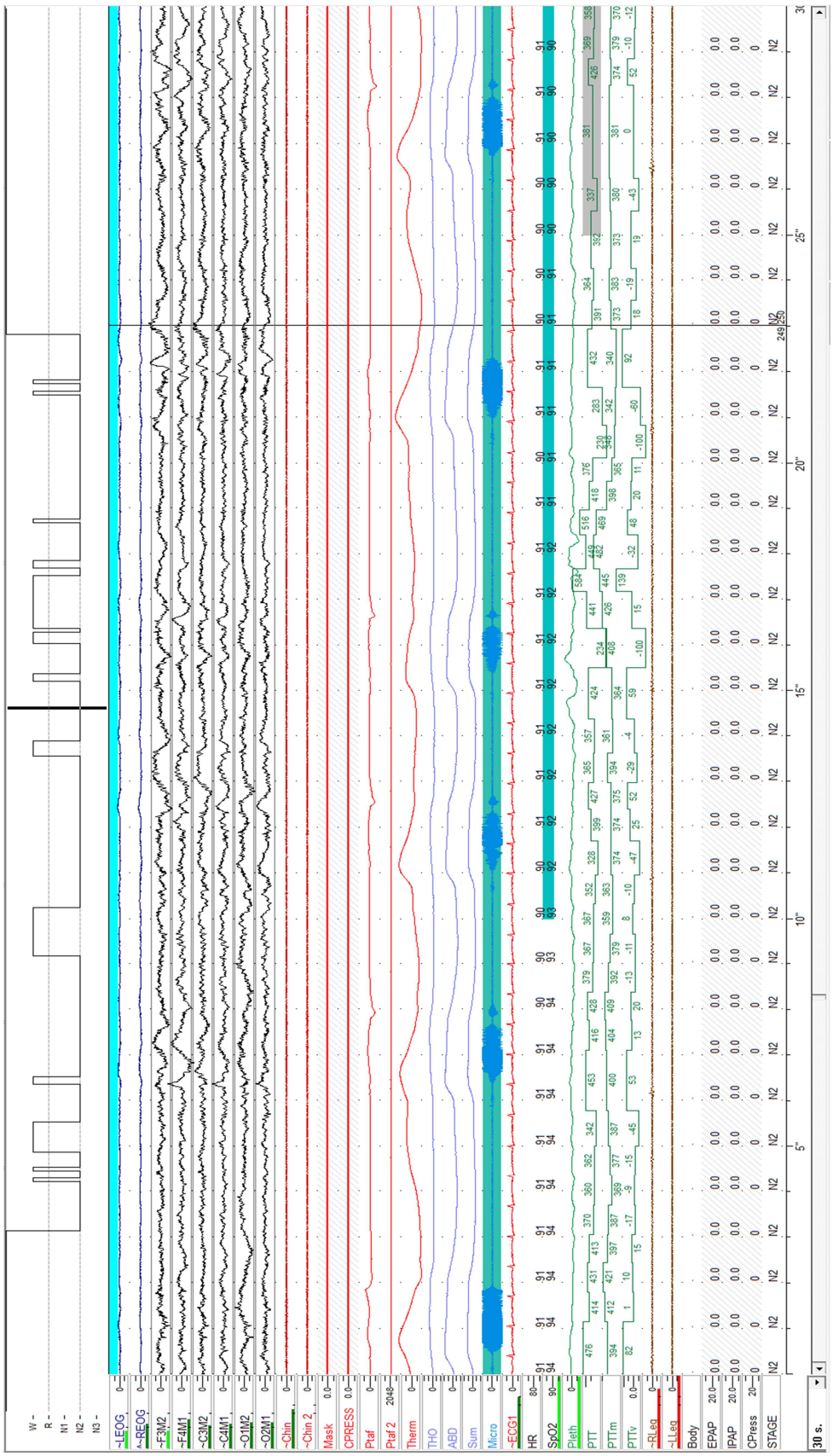

Figure 2: Interpretation of a PSG. See Table 2 for variable definitions.

Parameter	Variable Name	Purpose
Sleep stage	W, R, N1, N2, N3, STAGE	Wake, REM, Stage 1, Stage 2, Stage 3
Electrooculography	LEOG, REOG	Electrodes near eye help determine sleep stage
Electroencephalography	F3M2, F4M1, C3M2, C4M1, O1M2, O2M1	Electrodes on the scalp record brain activity and help determine sleep stage
Electromyography	Chin, Chin 2, RLEG, LLEG	Electrodes on chin and legs (or arms) measure muscle tensions/movement and help determine sleep stage
Electrocardiography	ECG1	Electrodes on chest record heart beats
Heart rate	HR	Heart rate
Pulse transit time	PTT, PTTm, PTTv	Indicator of arterial stiffness and blood pressure variability
Oximetry	SpO2	Probe measures oxygen levels in the blood stream
Microphone	Micro	Snoring and sleep sounds
Respiratory effort	THO, ABD, Sum	Thoracoabdominal movement
Positive airway pressure	Mask, CPRESS, EPAP, IPAP	Mask flow, pressure, expiratory/inspiratory PAP
Pressure transducer airflow	Ptaf, Ptaf 2, Therm	Nasal and oral airflow
Body position	Body	Supine, left, right

Table 2: Parameters measured during overnight PSG. Variable names listed as seen on PSG in Fig. 2 and their purpose.

Parameter	Definition
Total sleep time (TST)	Total time (in minutes) spent asleep
Time in bed (TIB)	Total time (in minutes) spent in bed
Sleep efficiency (SE)	Percentage of time spent asleep while in bed
Sleep latency (SL)	Time (in minutes) spent in bed before onset of sleep
Wake after sleep onset (WASO)	Total time spent awake in bed after the initial onset of sleep
Obstructive apneas	Number of total airway occlusions caused by obstruction (continued respiratory effort)
Central apneas	Number of total airway occlusions not caused by obstruction (no respiratory effort)
Mixed apneas	Total number of apneas
Hypopneas	Number of partial airway occlusions
Apnea index	Number of apneas per hour of sleep
Hypopnea index	Number of hypopneas per hour of sleep
Apnea-hypopnea index (AHI)	Number of apneas and hypopneas per hour of sleep
Arousal index	Number of arousals/awakenings per hour of sleep
Oxygen desaturation index (ODI)	Number of times per hour the blood oxygen level drops by a certain degree (typically 4%) from baseline
SpO_2 Nadir	The lowest oxygen saturation value the patient drops to

Table 3: PSG parameters interpreted by sleep technicians according to the AASM (2012) manual[42] and their definitions.

Actigraphy

Description

Actigraphy testing is the use of a small, portable device (actigraph) that uses an accelerometer to produce a signal in response to movement. The magnitude and duration of the signal depends on the magnitude of movement. These signals are digitized and stored as actigraphy counts, which can subsequently be analyzed. They can then be expressed graphically as actograms (Fig. 3) or numerically as total actigraphy counts per epoch (time interval of measurement, typically 30 seconds for studying sleep). An actiwatch is wrist-worn actigraph that may also be worn around the ankle or on the trunk. Actigraphy is used for the study of sleep patterns, based on the assumption that movement is reduced during sleep relative to being awake. Activity levels as measured by actigraphy can be used as a clinically-validated, non-invasive, and relatively inexpensive method for studying sleep at home. Actigraphs estimate parameters such as TST, WASO and SE. Current clinical practice guidelines suggest the use of actigraphy in normal, healthy adult populations, or in patients suspected of specific sleep disorders such as advanced sleep phase syndrome, delayed sleep phase syndrome, shift work disorder, jet lag disorder, and non-24-hour sleep/wake syndrome. When PSG is not available, actigraphy may be used to characterize circadian rhythm disorders and insomnia, and can also be used to measure TST in patients with OSA.[43] Due to lack of validation, actigraphy is currently not approved for use in children. Meltzer et al. found that actigraphy shows good sensitivity (0.88) and accuracy (0.84), but specificity was not as strong (0.46) in a study of school-aged children.[44] More research is warranted in the area specifically

by combining actigraphy with other physiological data that do not need sleep staging, such as pulse oximetry.

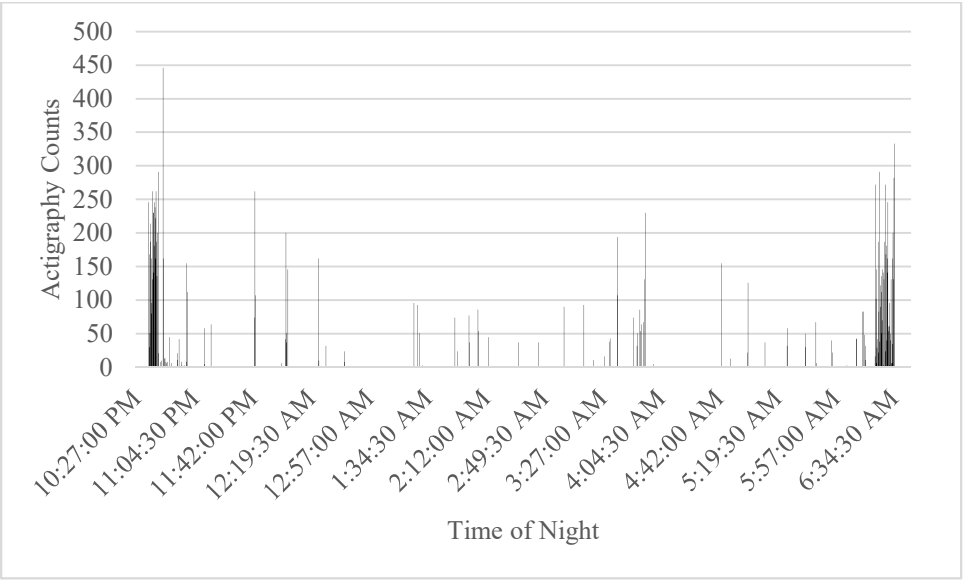

Figure 3: Sample actigram demonstrating a graphical display of actigraphy counts over the time of night.

Current Models

There are a number of different activity trackers and actigraphs currently on the market, such as the Motionlogger® Sleep Watch (Ambulatory Monitoring, Inc., Ardsley, NY), Actiwatch-2® (Philips Respironics, Amsterdam, NL), and commercial devices like the Fitbit Ultra® (Fitbit, San Francisco, CA) and UP® (Jawbone, San Francisco, CA). There are also more sophisticated devices, like the Actiwatch Spectrum Plus®, which has a watch-face display, longer battery life, and more memory for data storage. This device uses the same activity monitoring technology as the Actiwatch-2®. A more detailed description of different clinically validated devices and their specifications is provided in Table 4.

In a different study by Meltzer et al.,[45] the Motionlogger® Sleep Watch and Actiwatch-2® were compared for sensitivity, specificity, and accuracy in measuring TST, WASO, and SE validated against PSG. They found that both devices significantly overestimated WASO, while the Motionlogger® Sleep Watch device significantly underestimated TST. Agreement with PSG differed depending on the scoring algorithm used for the Motionlogger® Sleep Watch and the sensitivity setting for the Actiwatch-2®. Another study by Meltzer et al.[46] compared the Fitbit Ultra® with PSG, as well as the Motionlogger® Sleep Watch and Actiwatch-2®. Both the 'Normal' and 'Sensitive' sleep-recording modes of the Fitbit® were tested. The normal mode showed good sensitivity (0.86) and accuracy (0.84), but poor specificity (0.52), while the sensitive mode had adequate specificity (0.79), but inadequate sensitivity (0.70) and accuracy (0.71). Neither sleep-recording mode provided clinically comparable results to PSG. A similar study by Toon et al.[47] compared the commercial device UP®, as well as the MotionX 24/7® smartphone application (Fullpower, Santa Cruz, CA) to PSG and Actiwatch-2®. UP® showed good sensitivity (0.92) and accuracy (0.86), but poor specificity (0.66) when compared to PSG and was relatively analogous to the Actiwatch-2®. The MotionX 24/7® application did not accurately reflect sleep or wake.

Company	Actigraph	Advanced Brain Monitoring Inc.	Ambulatory Monitoring Inc.	CamNtech
Product	ActiGraph Link 	Night Shift 	Micro Motionlogger Watch 	MotionWatch 8 
Website	www.actigraphcorp.com	www.advancedbrainmonitoring.com	www.ambulatory-monitoring.com	www.camntech.com
Accelerometer Technology	3-axis solid state accelerometer with digital filtering	Minaturized poly-silicone-surface micromachined sensor (MEMS) triaxial digital output motion sensor	Solid state triaxial	3-axis MEMS
Dimensions	3.5 x 3.5 x 1.5	5.6 x 3.8 x 1.3	3.6 x 3.6 x 1.2	3.81 x 2.54 x 1.016
Light Sensor	N/A	N/A	Photodiode	Digital
Battery Life	14 days (wireless & gyro disabled, 30Hz sample rate)	Up to 5 nights between charges	30+ days	120 days
Memory Storage	4 GB nonvolatile	256 KB nonvolatile	2 MB nonvolatile	4 MB nonvolatile
Peer-reviewed Sleep Validation Study	Cellini N., et al. 2013[48]	Levendowski DJ., et al. 2014[49]	Cole R, et al. 1992[50]	Stevens A, et al. 2014[51]

Company	Condor Instruments	Philips Respironics	SOMNOmedics America Inc.
Product	ActTrust 	Actiwatch-2 	SOMNOwatch 
Website	www.condorinstcom	www.actigraphy.com	www.somnomedics-diagnostics.com
Accelerometer Technology	3-axis MEMS	Solid-state Piezo-electric accelerometer	MEMS
Dimensions	4.7 x 3.1 x 1.2	4.3 x 2.3 x 1.0	4.5 x 4.5 x 1.6
Light Sensor	RGB-IR Color light sensor	Silicon photodiode	NPN-phototransistor
Battery Life	90 days	30 days at 1-minute epoch length	26+ days
Memory Storage	4 MB nonvolatile	1 MB nonvolatile	8 MB internal storage card
Peer-reviewed Sleep Validation Study	Bellone, GJ., et al. 2016[52]	Shin, M., et al. 2015[53]	Dick R, et al. 2010[54]

Table 4: Examples of clinically validated actigraphs, their manufacturers, and some specifications. Note: this is not an exhaustive list, but a representative sample of available technologies.[55]

Machine Learning (ML)

Description

Machine learning involves statistical modeling of complex data, facilitating identification of 'rules' from a dataset to improve predictive capabilities of any given

statistical model. A dataset is partitioned into a test and training portions randomly (i.e., 20% and 80% of the original dataset, respectively). A chosen algorithm is trained on features (variables) associated with the training dataset. Subsequently, the rules generated by the training portion of the dataset are applied to the test dataset, and the metrics associated with classification determined.[56] Machine learning techniques are used in the technology industry for applications such as facial recognition, targeted advertising, and reverse image searching. These methods have also been applied to address questions in the medical field, such as predicting protein folding[57] or outcomes in radiation therapy.[58] Algorithms such as support vector machine (SVM) have been used to analyze electroencephalogram (EEG), electrooculogram (EOG), and electromyogram (EMG) signals during a PSG in order to predict sleep stage. Sleep stage is normally scored manually by a trained technician as defined by the AASM 2012 clinical staging manual.[59] The overall agreement between scoring performed by experts and the results of the ML algorithm was 82.53%, demonstrating that these techniques can be used to predict manually scored sleep annotations. Other ML algorithms have been applied to sleep medicine to study the sleep-wake transition in narcolepsy,[60] genes associated with circadian rhythms,[61] and sleep quality.[62]

Machine learning is broadly classified into two categories: supervised and unsupervised learning. Supervised learning methods require the value of the output variable to be known for each training sample. These algorithms train a model that predicts the value of output variables from the input variables, using the defined

features in the process. If the output variables are continuous, then the model is called a "regression function," whereas if the output variables are discrete, the model is called a "classifier." Because the output is known, the performance of these models can be tested using a separate test dataset. Unsupervised learning techniques require only the input feature values in the training data. Clustering techniques that partition data into groups, such as K-means clustering, hierarchical clustering, and spectral clustering are examples of unsupervised learning.[56]

Present Study

Pediatric OSA is a common condition that may lead to long-term complications if left untreated. Objective characterization is necessary prior to surgery due to the risk of PRAEs. Unfortunately, the PSG is expensive and inaccessible to many patients. Due to the ease of use, we investigated the utility of wrist-based actigraphy as a surrogate for sleep fragmentation and respiratory event annotations obtained from a PSG, either alone or in combination with other easily obtainable physiological parameters (e.g., oximetry). We hypothesized that analysis of our results with machine learning algorithms may support a framework for risk stratification in pediatric OSA. We also hypothesized that the majority of the sleep annotations could be successfully identified. If successful, this project ultimately lays the groundwork for the development of budget-conscious actigraphs, using accelerometers and/or gyroscopes, to streamline diagnosis and treatment of pediatric OSA.

Chapter 2: Methods

Subjects

Children were included in this study based on (i) age (between 2 and 17 years old) and (ii) referral for a PSG based on clinical symptoms suspicious for OSA. Exclusion criteria included (i) presence of a neuromuscular disorder, (ii) presence of a craniofacial syndrome, (iii) previous history of upper airway surgery, or (iv) developmental delay. If a new diagnosis was made (e.g., severe asthma) during the course of testing, subjects were withdrawn.

Data Collection

The patient and caregiver were counseled on the study during the course of clinical evaluation in the outpatient facility at the University of Maryland Medical Center (UMMC). Consent was obtained from the caregiver and assent was obtained from appropriately aged children (age 12 and above). Demographic and clinical data, including age, ethnicity, insurance status, height, weight, tonsil size (1-4 scale), and presence of asthma, allergies, and gastroesophageal reflux disease (GERD) were also obtained during these clinical visits. The BMI was calculated from height and weight data and converted to percentiles and BMI z-scores using data charts from the Center for Disease Control and Prevention (CDC) for children aged 2 to 18 years old.[63] All patients underwent full-night, in-laboratory PSG at the University of Maryland Pediatric Sleep Disorders Center, Baltimore, according to the AASM practice parameters.[42] When children arrived at the sleep lab, an Actiwatch-2® was secured around their wrists and adjusted for comfort.

Variables obtained from PSG included TST, SE, SL, WASO, AHI, SpO_2 nadir, arousal index, percentage of rapid eye movement (REM) and non-REM (NREM) sleep, and percentage of time spent in each stage of NREM sleep (N1, N2 and N3). Respiratory parameters such as apneas and hypopneas were obtained in 30-second epochs. Actigraphy data was exported to a computer-based software (Actiware, Philips Respironics, Amsterdam, NL) for analysis. Actigraphy counts were measured from Lights-Off time to Lights-On time, as well as from the start of Stage 2 sleep to the end of sleep time (estimated by adding TST, SL, and WASO). The activity during WASO was also estimated by identifying actigraphy counts during wake states. Wake states were indicated by the hypnogram obtained from the PSG recording. A total asleep actigraphy count was calculated by subtracting the WASO activity from the actigraphy count determined by sleep stage. A summary of respiratory events pertinent to our analysis and how they were measured is provided in Table 5.

Parameter	Definition	How it is measured
Actigraphy counts	Activity movement over threshold[a]	Actiwatch-2® (30s epoch)
Apnea	≥ 2 missed breaths, airflow reduction by > 90%, rib cage and/or abdominal effort continues, associated with desaturations	Oronasal thermistor, nasal air pressure and respiratory inductance plethysmography (30s epoch)
Hypopnea	≥ 2 missed breaths, airflow reduction by > 50%, rib cage and/or abdominal effort continues, associated with arousals or awakenings	Oronasal thermistor, nasal air pressure and respiratory inductance plethysmography (30s epoch)
Respiratory event with desaturation	Drop from baseline SpO_2 preceding respiratory event, to the nadir in SpO_2 following the event (> 3%)	Pulse oximeter (30s epoch)
Arousal	Abrupt shift of EEG/EMG frequency including alpha, theta and/or frequencies > 16Hz lasting at least 3s and preceded by at least 10s of stable sleep	EEG and EMG channels monitor sleep stage (30s epoch)
Total AHI	Sum of apnea and hypopnea events (per hour)	

Table 5: Summary of variables pertinent to analysis, their definitions, and how they were measured.[42]
Abbreviations: EEG, electroencephalography; EMG, electromyography; AHI, apnea-hypopnea index.
[a]Proprietary threshold defined by Philips Respironics.

Statistical Analysis

All data were entered into a Microsoft Excel (2016; Microsoft, Redmond, WA) spreadsheet. Each demographic variable was reported using descriptive statistics (mean and percentage). Continuous variables were described as means with 95% confidence intervals (CIs). Total actigraphy counts and time-based counts (per 30

second epoch) were obtained from each patient. Respiratory event annotations were also obtained for each of the epochs—including apneas, hypopneas, respiratory events, arousals, and total AHI.

The intrinsic relationships between individual predictors thus collected were assessed first by Pearson correlation coefficients and associated *P* values. A correlation matrix was graphically constructed for visualization of the relationships. Statistically significant pairwise correlations were identified.

To identify the differences between four principal sleep-related states (awake, apnea, hypopnea, and asleep), an empiric cumulative distribution function (CDF) was plotted. The CDF calculates the cumulative probability for a given value of the x-axis. A CDF provides probability that a data value is less than or equal to a certain value, higher than a certain value, or between two values.

After plotting a CDF, a Q-Q plot was used to demonstrate the non-parametric nature of this dataset. A Q-Q plot is a probability plot that compares two probability distributions by plotting their quantiles. The statistical differences between actigraphy counts plotted as a function of the sleep state were identified by a Kruskal-Wallis rank sum test. The Kruskal-Wallis rank sum test is used to determine statistically significant differences between 3 or more groups with a non-parametric dataset. A post-hoc Dunn's test was performed to determine statistically significant pairwise

differences between groups. A boxplot was constructed to identify the differences between the distributions of the data.

Two ML algorithms were selected for modeling the data collected in this experiment: Linear discriminant analysis (LDA) and conditional inference tree. Both of these algorithms are used when the output variable is known (supervised) and both have discrete output variables (classifiers). Conditional inference tree is a type of classification and regression tree (CART), in which the predicted outcome is the class to which the data belongs. Conditional inference trees are structured to provide separation between classes, with multiple test procedures applied to determine whether there is no significant association between the predictor and the response variables.[64]

The utility of actigraphy counts with oximetry to predict the severity of OSA was investigated using LDA. Linear discriminant analysis takes a data set of patients and their characteristics that include (i) severity of OSA diagnosed as mild, moderate, or severe, (ii) actigraphy counts from the entire duration of sleep, and (iii) number of respiratory events with clinically significant desaturations (> 3% from baseline). Linear discriminant analysis uses the input parameters to derive the coefficients of a scoring function for each severity class of OSA. Each function takes the numeric variables as arguments, described by (ii) and (iii), and obtained from a patient. Linear discriminant analysis scales each variable according to its OSA severity-specific coefficients and outputs a score. The LDA model uses the highest score to allocate a

patient to a category (OSA severity). Results were plotted with a decision boundary. Misclassifications were also identified.

To predict the annotation obtained from a PSG (apnea, hypopnea, awake, and asleep) an ML model utilizing conditional inference was identified. The algorithm includes a 5-fold cross-validation that included training using 80% of the data chosen randomly (training), and subsequently applied to the rest of the data (20%, test). Classification ability of the ML algorithm ascertained from the test dataset was assessed using a receiver operating characteristic (ROC) analysis, plotting the true positive rate vs. the false positive rate (or sensitivity vs. 1-specificity). The overall prediction was assessed using area under the curve (AUC). Each pairwise ROC analysis was performed separately (e.g., apneas vs. awake, hypopneas vs. apneas, etc.) All statistical testing was performed using the R statistical system (v4.3, www.r-project.org). A P value < 0.05 was considered significant.

Chapter 3: Results

Demographics and Clinical Characteristics

The demographic and clinical characteristics of the children are summarized in Table 6. Most children (70%) were younger than 6 years old. The male to female ratio was 3:2. The majority were Hispanic or Black (7, 70%). Allergic rhinitis was seen in 7 children (70%), of whom 3 were on intranasal corticosteroids. Most children (6, 60%) had tonsillar hypertrophy (grade 3 or 4). Mild asthma was present in 20% of patients, of whom only one was managed with a bronchodilator. Snoring was present in 100% of children, while nocturnal enuresis was only present in 10% of children.

Characteristic	**Value**
Age, mean (95% CI)	5.8 (3.1 to 8.5)
Gender, n (%)	
Male	6 (60)
Female	4 (40)
Ethnicity, n (%)	
Hispanic	1 (10)
Black	6 (60)
White	2 (20)
Other	1 (10)
Snoring, n (%)	10 (100)
Allergic rhinitis, n (%)	7 (70)
Asthma, n (%)	1 (10)
Gastroesophageal reflux disease, n (%)	0 (0)
Nocturnal enuresis, n (%)	1 (10)
Body mass index z-score[a], mean (95% CI)	0.69 (-0.15 to 1.47)
Tonsil size[b], mean (95% CI)	2.6 (2.3 to 2.9)

Table 6: Demographics and Clinical Characteristics of the Study Population (n = 10).

Abbreviation: CI, confidence interval

[a]The body mass index z-score was calculated using Centers for Disease Control and Prevention data charts.[63]

[b]Tonsillar size is defined on a 1 to 4 scale by the Brodsky scale.

PSG Results

PSGs were successfully completed in all 10 patients. The sleep efficiency was excellent (> 80% in all patients). The mean AHI was 16.2 (9.8 to 22.7). Moderate OSA (AHI ≥ 5) was seen in 30% of patients, and severe OSA (AHI ≥ 10) in 60% of patients. The average SpO_2 nadir was 82.2 (74.1 to 90.3). A summary of PSG parameters obtained from the study population is provided in Table 7.

PSG Parameter	**Value**
Total sleep time, min	400.6 (386.5 to 414.7)
Sleep latency, min	30.8 (18.8 to 42.8)
Sleep efficiency, %	87.9 (84.8 to 91.0)
REM latency, min	201.3 (154.7 to 247.9)
WASO, min	23.6 (16.8 to 30.4)
REM duration, %	11.1 (9.9 to 12.3)
Sleep stage N1 duration, %	0.08 (-0.002 to 0.2)
Sleep stage N2 duration, %	52.2 (48.1 to 56.3)
Sleep stage N3 duration, %	24.6 (21.4 to 27.9)
Obstructive apneas	54.0 (24.9 to 83.1)
Obstructive hypopneas	49.7 (23.3 to 76.1)
Apnea index	8.9 (4.7 to 13.1)
Hypopnea index	7.3 (3.6 to 11.0)
Mixed apneas	0.9 (-0.1 to 1.9)
Central apneas	5.8 (2.7 to 8.9)
Snoring episodes	167 (51.7 to 282.3)
Total number of respiratory events	110.4 (64.4 to 156.4)
AHI	16.2 (9.8 to 22.7)
Basal SpO_2	97.9 (97.4 to 98.6)
SpO_2 nadir, %	82.2 (74.1 to 90.3)
Average SpO_2 during REM	97.6 (96.7 to 98.5)
Average SpO_2 during NREM	98.0 (97.4 to 98.6)
Sleep time with SpO_2 <90%, min	2.2 (0.4 to 3.9)
Awake HR	101.1 (93.6 to 108.5)
Arousal index	14.2 (9.2 to 19.1)
Spontaneous arousals	19.4 (14.8 to 24.0)
Respiratory events with desaturations	39.7 (9.0 to 70.4)

Table 7: PSG Parameters interpreted by sleep technicians according to the AASM 2012 manual,[42] and their values, mean (95% CI).
Abbreviations: CI, confidence interval; REM, rapid eye movement; AHI, apnea-hypopnea index; NREM, non-rapid eye movement; HR, heart rate.

Actigraphy Results

The actigraph stayed on the wrist for the entire duration of sleep in all patients. No adverse events were noted. There were no instances of device malfunction or data loss. Discretized actigraphy counts were obtained successfully in all 10 patients. Table 8 summarizes the total counts obtained as a function of wake-sleep stage from all patients.

Actigraphy Parameter	**Value**
Total actigraphy count (lights on to off)	9936.4 (5891.8 to 13981.0)
Total actigraphy count (stage 2 to wake)	5947.4 (3632.0 to 8262.8)
WASO activity acount	2708.2 (869.6 to 4546.8)
Sleep activity (total activity from sleep stages minus WASO)	3239.2 (2004.9 to 4473.5)

Table 8: Actigraphy parameters measured by the Actiwatch-2®, and their values, mean (95% CI).
Abbreviations: WASO, wake after sleep onset; CI, confidence interval.

Statistical Analysis

In the current study, 6 variables (actigraph counts, respiratory events with desaturation, arousals, apneas, hypopneas, and the AHI) were obtained from the PSGs and the actigraphy recordings. The relationships among these 6 parameters were assessed using Pearson correlation coefficients and their associated *P* values, as shown in Fig. 4. The correlation matrix demonstrates significant pairwise relationships among all 6 parameters. The range of correlation coefficients between actigraph counts and PSG parameters were from 0.23 for apneas ($P = .09$) to 0.67 for respiratory events ($P = .002$).

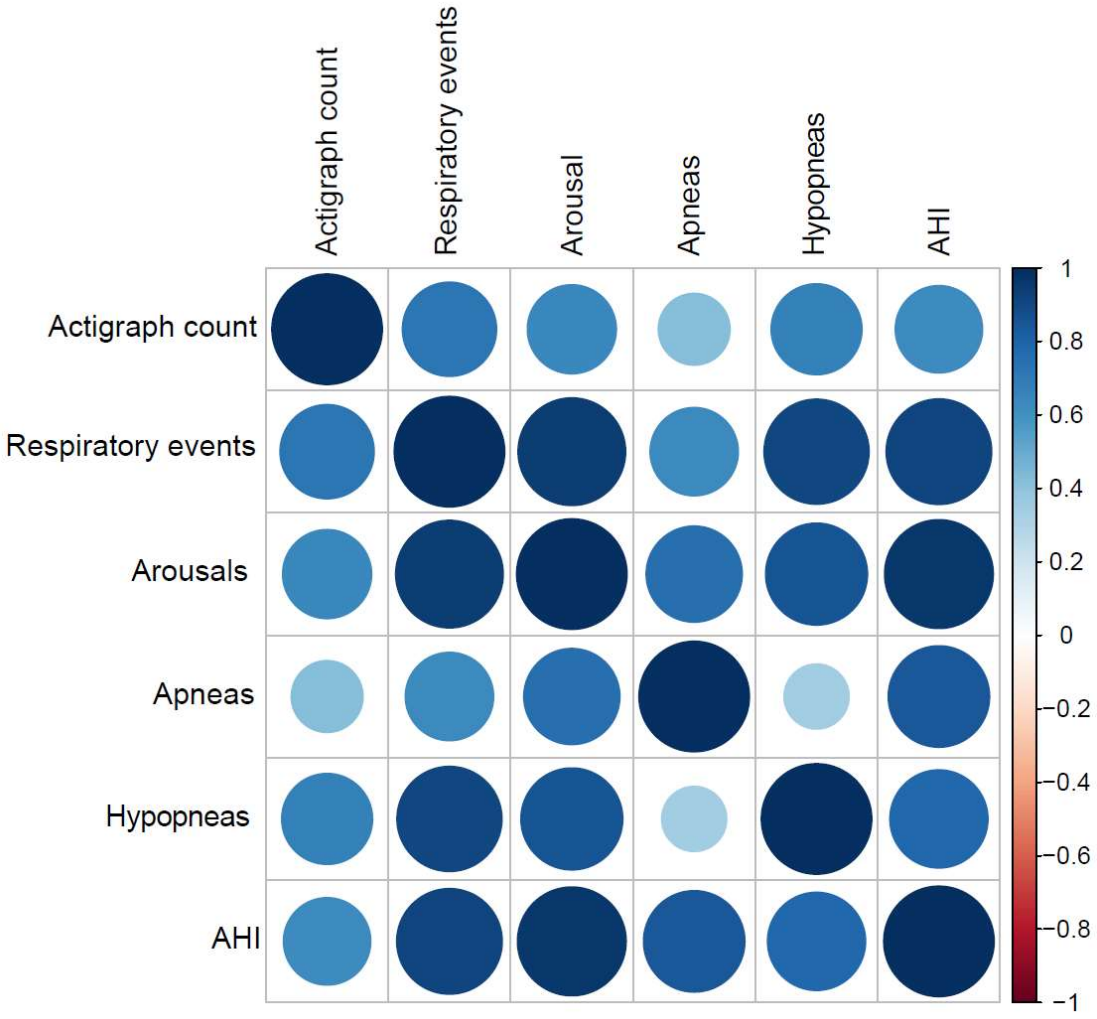

Figure 4: Correlation matrix of actigraphy counts and indices representing respiratory events. The color and the size of each circle indicates the magnitude of the pairwise correlation. Total counts correlated well with each of the respiratory events except apneas. The legend is provided to the right of the matrix.

Actigraphy counts as a function of 4 PSG-annotated respiratory epochs (awake, asleep, apnea, and hypopnea events) were plotted using an empiric cumulative distribution function (CDF) in Fig. 5. Here, the CDFs show the probability level associated with actigraphy counts either greater than or less than a specific value. The plot demonstrates the lack of overlap between the CDFs associated with awake and

hypopnea counts. The majority of counts obtained from the asleep and apnea annotations were convergent.

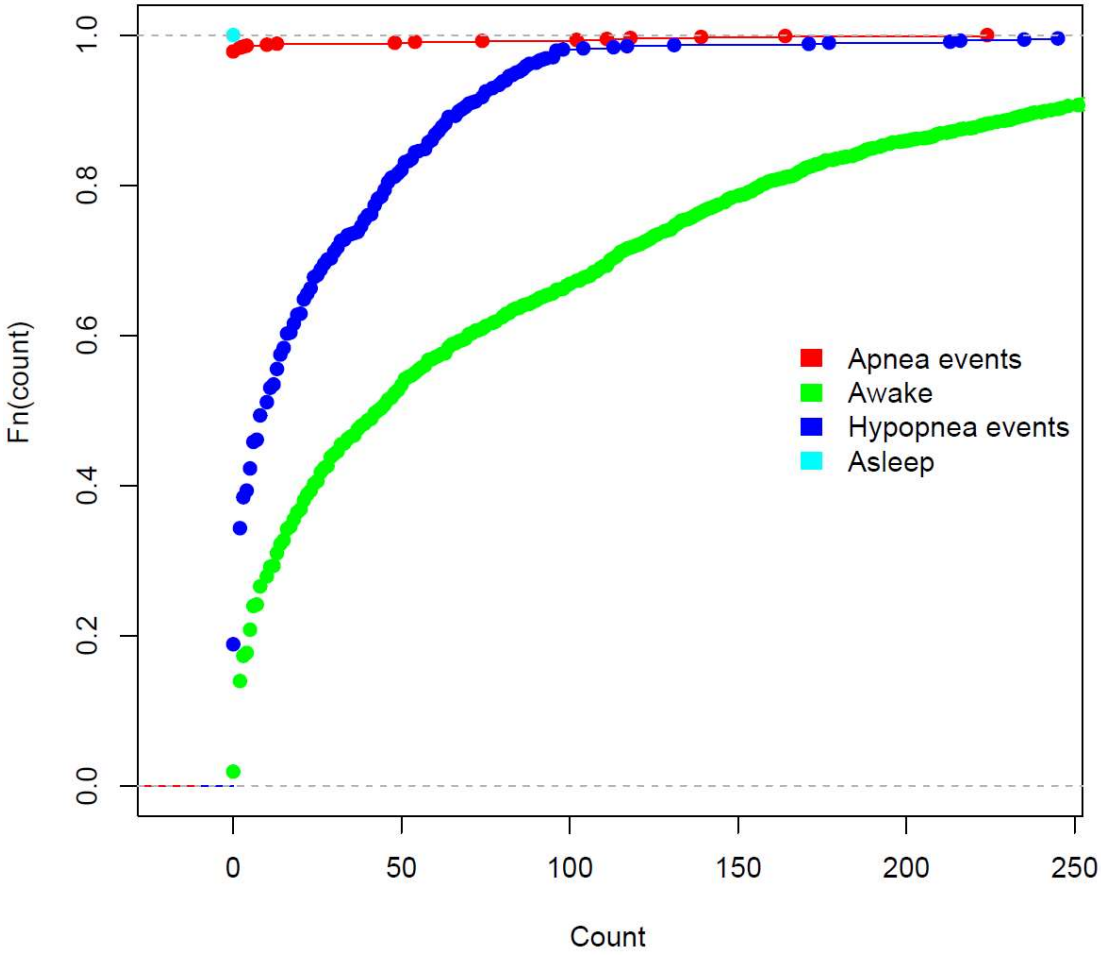

Figure 5: Empirical cumulative distribution function showing distribution of actigraphy counts as a function of type of respiratory event. The plot demonstrates the proportion of the total distribution at a certain level of actigraphy counts grouped by annotated respiratory events (legend to the right). Higher overall levels of activity were seen both in the awake and hypopnea events, but there was little overlap between the two groups.

The mean actigraphy counts were highest in the awake state (94.16 events/epoch; 91.16 to 97.48) and lowest in the asleep state (1.37 events/epoch; 0.89 to 1.85). Next, the counts obtained from each annotation were compared using the Kruskal-Wallis

rank sum test, due to the non-normal distribution of the data. The differences between the groups was statistically significant ($\chi^2 = 9935.3$, $P = 10^{-15}$). Multiple pairwise comparisons between the different groups of actigraphy counts using Dunn's test demonstrated statistically significant differences (range of P values: 10^{-30} to 10^{-5}) except one comparison (asleep vs. apnea events, $P = .14$).

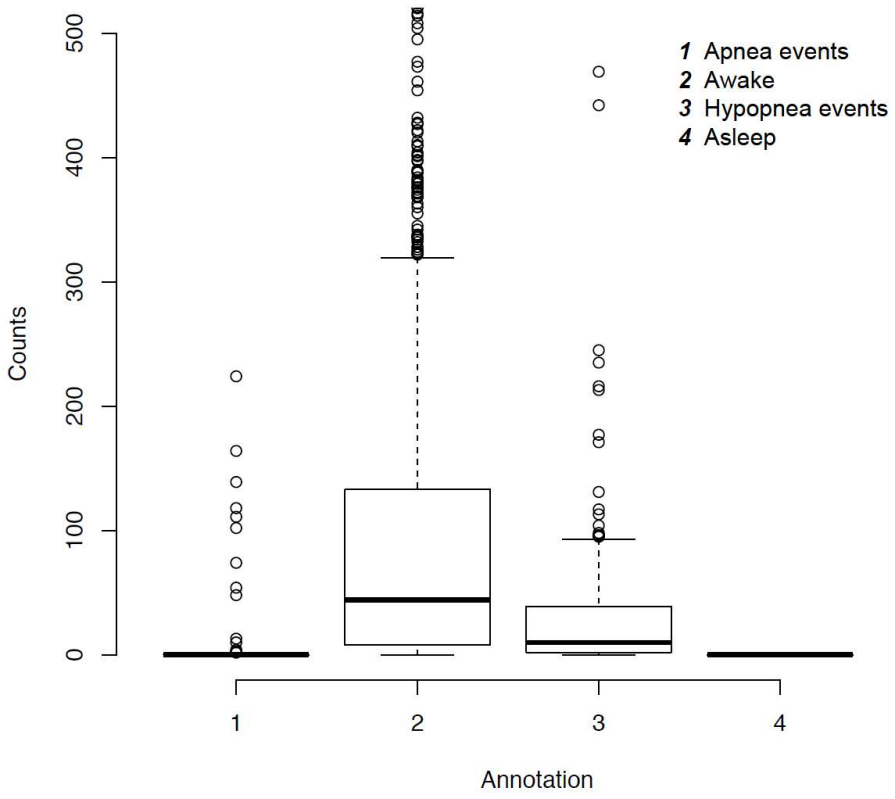

Figure 6: Boxplot of overall counts grouped by annotation of respiratory events. Each boxplot is centered around the median with the 25th-75th percentile shown by the extent of the box and the whiskers spanning the range. The actigraphy counts were significantly higher for both awake and hypopnea annotations. Statistical comparisons are provided in text.

Machine Learning

Results from ML algorithms used in the current study are summarized in Fig. 7 and 8. First, results from LDA analysis was used to identify a linear combination of two

features (mean actigraphy counts and respiratory events in each patient) to predict the severity class of OSA by their labels known *a priori*. Two separate classes emerged as shown by a classification boundary that divides the figure into two regions (pink and white). Patients labeled as 3 had AHI greater than 10 (severe OSA). Labels 1 and 2 represented mild and moderate OSA, respectively. As is shown in Fig. 7, all but one patient was correctly classified as severe OSA.

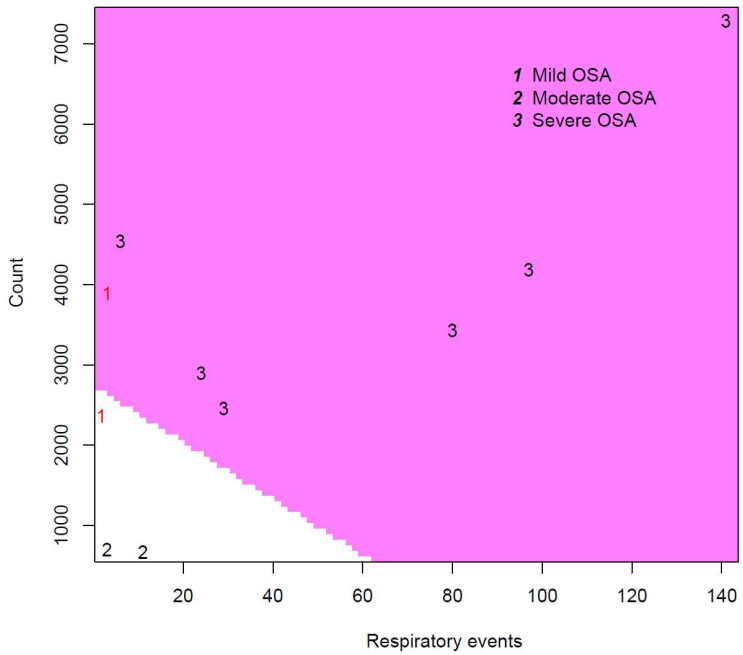

Figure 7: Plot demonstrating classification of observations based on linear discriminant analysis for every combination of two variables (actigraphy counts and respiratory events) for each patient. In addition, the classification border is drawn, and the misclassified patients shown in red. Respiratory events denote events where clinically significant deterioration in oximetry recording (> 3% drop) was identified.

Next, results from the conditional inference tree are shown in Fig. 8 and 9. The model obtained from the training portion of dataset (80%) is demonstrated in Fig. 8. The tree demonstrates a successful split based on two separate levels of actigraphy counts.

Here, nodes 2 and 5 represent classifications that contain mostly asleep and awake annotations. Node 4 shows that the algorithm was not as effective in separating awake and hypopnea events. The overall structure of the tree demonstrates that the range of actigraphy counts between 0 and 98 was not fully resolved.

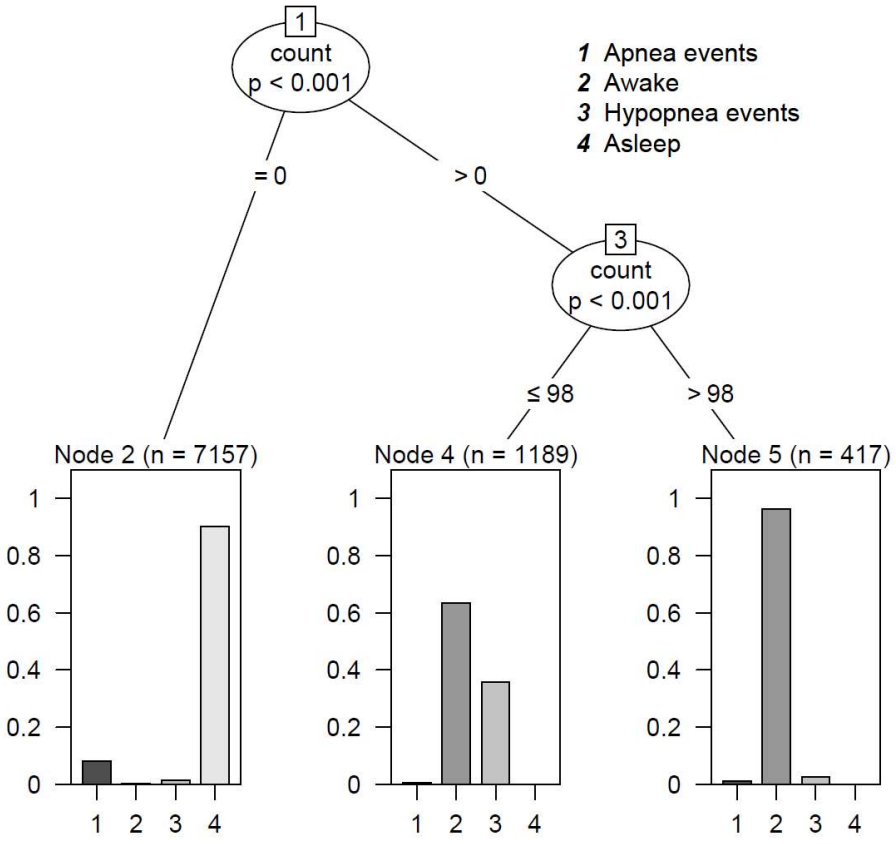

Figure 8: Results from a machine learning algorithm (conditional inference tree) as applied to the training dataset for various classes of respiratory events. Visualization of this interpretable classification algorithm demonstrates the separation of the classification tree into "nodes." Node 1 represents classification based on counts obtained from actigraphy. Node 2 shows classification results as indicated by proportion of patients classified correctly. For example, the majority of actigraphy counts related to an annotation indicating 'asleep' were zero. Node 3 demonstrates a second threshold of actigraphy counts. Based on this threshold, most patients were classified to be awake when their actigraphy counts exceeded 98. Associated *P* values are also listed. Legend for annotation is provided in the figure.

The classification tree created by the conditional inference algorithm was subsequently applied to the training dataset, which constituted 20%. As the prediction is a multiclass problem (four distinct classes of respiratory event annotations), an ROC curve was constructed for each pairwise comparison. The AUC ranged from 0.52 (apneas vs. asleep) to 0.88 (awake vs. asleep). A sample ROC is shown for separation of hypopneas from asleep annotations.

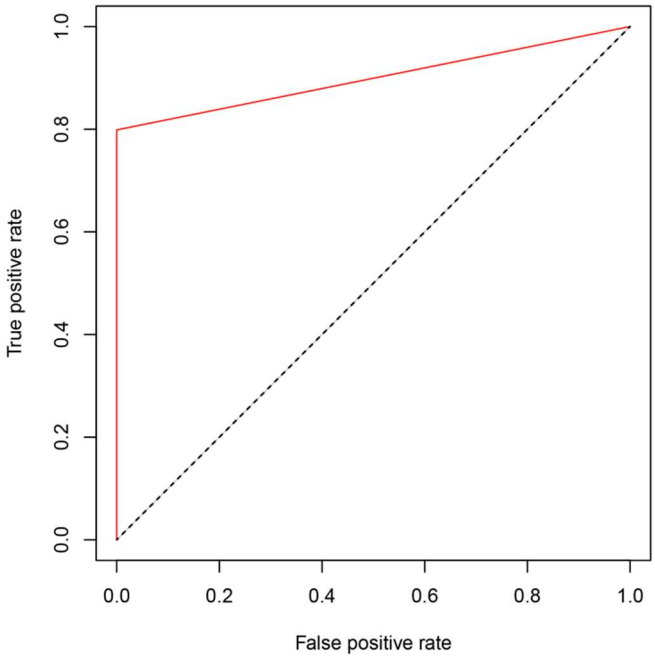

Figure 9: Receiver operating characteristic curves of machine learning techniques to predict respiratory events (hypopneas events vs. asleep events). The figure plots false positive rate (x-axis) as a function of true positive rate (y-axis). Performance of the ML method is summarized by area under the curve (AUC). An ideal classifier has an AUC of 1, while a chance-based classifier, positioned diagonally (broken line), has an AUC of 0.5. In the current ROC, an AUC of 0.88 was obtained.

Chapter 4: Discussion

The current study used a non-invasive wrist-based movement sensor with the aim of capturing sleep fragmentation and arousals in children undergoing PSG. The preliminary results demonstrate that actigraphy counts can be reliably obtained during a PSG, and no adverse events were identified. Actigraphy counts correlated well with the respiratory event annotations from the sleep study. Linear discriminant analysis showed that 100% of children with severe OSA were classified correctly using actigraphy counts and respiratory events alone. A conditional inference model demonstrated that hypopnea events can be classified with high accuracy, while the distinction between apnea and asleep events could not be achieved effectively. These preliminary results support further investigation of the use of actigraphy in combination with pulse oximetry.

The male predominance in this study population is consistent with previous studies that report a 3:2 male-to-female ratio.[65] The presence of allergic rhinitis in most children is consistent with the inflammatory etiology of OSA. The mean AHI was 16.2 (9.8 to 22.7), which would be classified as severe. Every child tested was diagnosed with some degree of OSA, confirming the clinician's suspicion based on clinical parameters.

Actigraphy counts were found to correlate strongly with all respiratory events, except apneas. Apneas are defined by complete cessation of airflow, and can be associated with desaturations. One of the defining parameters of hypopneas is arousal resulting

in movement.[42] Arousals would be expected to correlate more with wrist movement than desaturations would, so these results are consistent. Plotting actigraphy count distributions as a function of respiratory event annotations showed that the highest activity levels were seen in the awake and hypopneic events, with little overlap between these two groups. However, there was no significant distinction between asleep and apneic events. These results show that actigraphy counts alone were not sufficient to detect apneic events.

The combination of total actigraphy and respiratory events with LDA generated a decision boundary for classification of patients. Linear discriminant analysis uses several predictor variables to separate categorical variables with decision boundaries. This ML strategy projects data points into regions characterized by their vectors of means, or "centroids."[66] These centroids are labelled by categories, and have distinct linear boundaries. Linear discriminant analysis achieves this by deriving coefficients of a scoring algorithm (called a discriminant function) using the input data for each category and using the predictor variables as arguments. When evaluating an unknown sample, the discriminant functions output a score, and the LDA uses the highest score from each function to determine the nearest centroid. Then the sample is allocated to a category.[56] Linear discriminant analysis was chosen over other classification ML algorithms as the results are generalizable in a clinical population, and are potentially more interpretable to a clinician. In this experiment, prediction is based on a patient being projected above or below the classification boundary. A

patient above the line would be predicted to have severe OSA, and a patient below the line would be predicted to have mild or moderate OSA.

When this analysis attempts to stratify patients into the three categories of mild, moderate, and severe OSA, 20% of the patients are misclassified (those with mild OSA are misclassified as moderate or severe). If instead, only the two categories of mild/moderate and severe are separated by the decision boundary, only 10% of patients are misclassified (those with mild/moderate are misclassified as severe). Regardless, 100% of the severe OSA cases were correctly identified. A higher false positive rate is preferred to a higher false negative rate, as a patient with mild OSA misclassified as severe will simply be referred for PSG or monitored post-operatively. A patient with severe OSA misclassified as mild/moderate would be at risk for PRAEs if they were not screened further with PSG or monitored post-operatively.

A decision tree resembles a flowchart, where each internal node denotes a test, each branch represents an outcome, and each leaf holds a class label. Conditional inference trees determine which variable to split based on a measure of association between each covariate and the outcome of interest. Then, the best split point for a variable is determined using statistical inference procedures. The association between each covariate and the outcome is quantified using a regression model, and a node is chosen to be split if there is sufficient evidence to reject the global null hypothesis (the hypothesis that none of the covariates has a univariate association with the outcome). Once a node is chosen, the optimal split point for that variable is

determined based on maximizing the two-sample association test. Stopping rules may be put into place to prevent overfitting.[67] A conditional inference tree was selected over other ML algorithms, such as SVM, because it can be easily clinically interpreted. Other more complicated algorithms, like SVM, may be more accurate but are not as easy to interpret and may require larger training datasets to function.

Conditional inference has been used in the past to model intimate partner violence,[68] bioaccumulation of organic chemicals,[69] and gene regulation.[70] In this experiment, the variable of interest was actigraphy counts. The first node splits the presence or absence of actigraphy—when activity is 0, the patient is predicted to be asleep or apneic. Node 3 represents a second threshold for activity, and when the actigraphy count is greater than 98, the patient is predicted to be awake. When the actigraphy count is non-zero but less than 98, this algorithm cannot reliably distinguish between awake and hypopnea events. Increase in sample size is expected to improve the performance of the algorithm used.

The principal limitation for this study was the patient sample size and lack of adequate statistical power. As data collection was performed over the winter, attrition due to illness and weather prevented recruitment of a larger population. Machine learning typically requires much larger sample sizes to optimize validation. Therefore, the results presented here are preliminary and lay the foundation for larger studies, which will hopefully improve the current ML algorithms and allow for

exploration of other, more complex algorithms. Improvements in the recruitment process have been made through partnership with an additional sleep laboratory.

When developing a model for a given dataset with an appropriate predictive accuracy, it is necessary to avoid overfitting the data. Increasingly complex models may be able to fit the data exceedingly well, and eliminate residual error between the model and the training data, but will be unable to successfully make predictions about new data.[56] This concept of determining the proper fit is demonstrated in Fig. 10 below. We attempted to prevent overfitting by refraining from the potentially more accurate and complex ML algorithms, in favor of simpler and more clinically interpretable models.

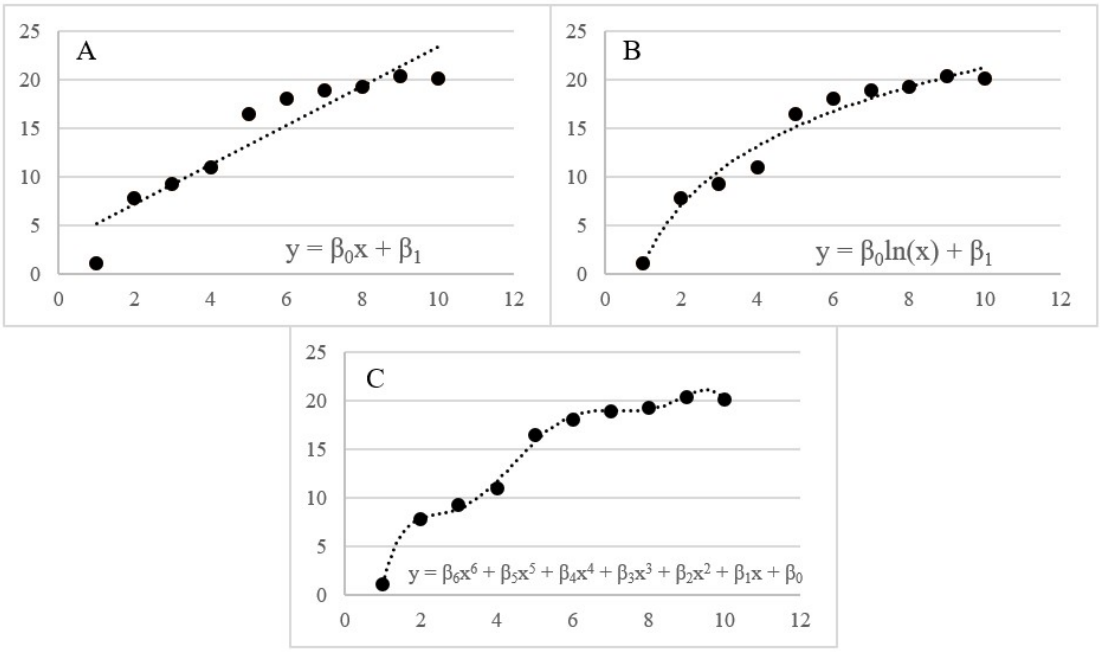

Figure 10: **Demonstration of model fit.** The sample data is logarithmic: (A) shows underfitting (linear), (B) shows appropriate fit (logarithmic), (C) shows overfitting (6^{th} degree polynomial).[56]

Although the utility of actigraphy was demonstrated for identifying hypopneas and other respiratory events, actigraphy alone cannot capture apneas. Complete cessation of airflow does not necessarily imply arousal. Actigraphy cannot be used on its own to evaluate the severity of OSA. As apneas are physiologically coupled usually with reduction in oxygen saturation, our future efforts will aim to identify events with a combination of oximetry and actigraphy for capturing a greater proportion of respiratory events.

Using the Actiwatch-2® has limitations as well. This proprietary technology utilizes a solid-state piezo-electric accelerometer with a certain, unknown threshold for detecting movement. These actigraphs are also expensive, costing around $700. This estimated cost does not include the required docking-station, and software necessary for data retrieval. The ultimate goal of this project is to design and build a budget-conscious actigraph, but further experiments would be warranted to optimize the sensitivity to movement data and achieve results similar to that of the Actiwatch-2®.

Chapter 5: Conclusions and Future Directions

The preliminary results presented here provide a framework for home-based risk stratification of pediatric OSA. Actigraphy and home oximetry could be simple, cost-effective and reliable instruments for screening for OSA. Our results and future directions may provide a logical outline for two sleep centers—one to screen, and the other to further validate higher risk children. This model can hopefully reduce the burden on current sleep centers, assist clinicians in decision making, decrease resource utilization, and more importantly, reduce the discomfort to children and their families in the current model of obtaining PSGs in all children.

Combining real-time oximetry data with actigraphy would theoretically provide information about both apneas (desaturations) and hypopneas (arousals). Adding this data could enhance the machine learning algorithms and improve their predictive capability. An actigraph and a finger-based pulse-oximeter could easily and conveniently be implemented as an HSAT if further research can validate their combined use.

The current classification border is able to classify patients with mild/moderate OSA and severe OSA. Children classified as severe using this analysis could either be referred for follow-up PSG, or simply monitored for PRAEs post-T&A. Further research could be performed to evaluate the effectiveness, cost, and safety of these treatment strategies.

A budgeted regression analysis should be performed to determine the cost-effectiveness of screening with actigraphy vs. PSG alone. Previous studies have estimated 1,000 children can be evaluated by PSG with $1,000,000;[41] the number of children that could be effectively screened using actigraphy (with or without oximetry) with $1,000,000 should be determined. Ultimately, a cost-conscious device should be built and evaluated using the same budgeted regression analysis.

This cost-conscious device could be achieved by using linear accelerometers to capture 3 degrees of linear acceleration, and rotational accelerometers or gyroscopes to capture 3 degrees of rotational movement. In total, this actigraph would have 6 degrees of freedom. Early prototypes have been developed using an Arduino® board (Arduino, Somerville, MA), an open source hardware and software company that manufactures computer boards. Such an actigraph would provide a continuous distribution of the accelerometry data. Other components could be added to the board, such as a simple heart rate monitor. Heart rate increases during arousals, and this could be used in combination with the actigraphy data to improve detection of apneas and hypopneas. Further considerations for device design include: battery life, memory (data storage), size of the watch, and cost for both development and production. Ultimately, this device could provide a cost-conscious solution to current issues with diagnostic methodologies for pediatric OSA.

Bibliography

1. Kaditis A., Kheirandish-Gozal L. & Gozal D. Algorithm for the diagnosis and treatment of pediatric OSA: A proposal of two pediatric sleep centers. *Sleep Med.* **13,** 217–227 (2012).
2. Lumeng, J. C. & Chervin, R. D. Epidemiology of pediatric obstructive sleep apnea. *Proc. Am. Thorac. Soc.* **5,** 242–252 (2008).
3. Redline, S. *et al.* Risk factors for sleep-disordered breathing in children. Associations with obesity, race, and respiratory problems. *Am. J. Respir. Crit. Care Med.* **159,** 1527–1532 (1999).
4. Rosen, C. L. *et al.* Prevalence and risk factors for sleep-disordered breathing in 8- to 11-year-old children: association with race and prematurity. *J. Pediatr.* **142,** 383–389 (2003).
5. Montgomery-Downs, H. E. & Gozal, D. Snore-associated sleep fragmentation in infancy: mental development effects and contribution of secondhand cigarette smoke exposure. *Pediatrics* **117,** e496-502 (2006).
6. Guilleminault C., Ji H.L. & Chan A. Pediatric obstructive sleep apnea syndrome. *Arch. Adolesc. Med.* **159,** 775–785 (2005).
7. Kheirandish-Gozal, L., Bhattacharjee, R., Bandla, H. P. R. & Gozal, D. Antiinflammatory therapy outcomes for mild OSA in children. *Chest* **146,** 88–95 (2014).
8. Marcus C.L. *et al.* Use of nasal continuous positive airway pressure as treatment of childhood obstructive sleep apnea. *J. Of* **127,** 88–94 (1995).
9. Pirelli, P., Saponara, M. & Guilleminault, C. Rapid maxillary expansion in children with obstructive sleep apnea syndrome. *Sleep* **27,** 761–766 (2004).

10. Marcus, C. L. & Loughlin, G. M. Obstructive sleep apnea in children. *Semin. Pediatr. Neurol.* **3,** 23–28 (1996).
11. Marcus C.L. *et al.* Diagnosis and management of childhood obstructive sleep apnea syndrome. **130,** e714–e755 (2012).
12. O'Brien, L. M. *et al.* Neurobehavioral implications of habitual snoring in children. *Pediatrics* **114,** 44–49 (2004).
13. Gottlieb, D. J. *et al.* Sleep-disordered breathing symptoms are associated with poorer cognitive function in 5-year-old children. *J. Pediatr.* **145,** 458–464
14. Schechter, M. S. & Section on Pediatric Pulmonology, S. on O. S. A. S. Technical report: diagnosis and management of childhood obstructive sleep apnea syndrome. *Pediatrics* **109,** e69 (2002).
15. Marcus C.L. *et al.* A randomized trial of adenotonsillectomy for childhood sleep apnea. *N. Engl. J. Med.* **368,** 2366–2376 (2013).
16. Mitchell, R. B. & Kelly, J. Outcome of adenotonsillectomy for obstructive sleep apnea in obese and normal-weight children. *Otolaryngol.--Head Neck Surg. Off. J. Am. Acad. Otolaryngol.-Head Neck Surg.* **137,** 43–8 (2007).
17. Cullen, K. A., Hall, M. J. & Golosinskiy, A. Ambulatory surgery in the United States, 2006. *Natl. Health Stat. Rep.* 1–25 (2009).
18. Isaiah A & Pereira KD. Outcomes after adenotonsillectomy using a fixed anesthesia protocol in children with obstructive sleep apnea. *Int. J. Pediatr. Otorhinolaryngol.* **79,** 638–43 (2015).

19. Roland, P. S. *et al.* Clinical practice guideline: Polysomnography for sleep-disordered breathing prior to tonsillectomy in children. *Otolaryngol.--Head Neck Surg. Off. J. Am. Acad. Otolaryngol.-Head Neck Surg.* **145,** S1-15 (2011).
20. Mitchell, R. B., Pereira, K. D. & Friedman, N. R. Sleep-disordered breathing in children: survey of current practice. *The Laryngoscope* **116,** 956–8 (2006).
21. Isaiah A., Hamdan H., Johnson R.F., Naqvi K. & Mitchell R.B. Very Severe Obstructive Sleep Apnea in Children: Outcomes of Adenotonsillectomy and Risk Factors for Persistence. *Otolaryngol. - Head Neck Surg. U. S.* **157,** 128–134 (2017).
22. Cha, J. *et al.* The Effects of Obstructive Sleep Apnea Syndrome on the Dentate Gyrus and Learning and Memory in Children. *J. Neurosci. Off. J. Soc. Neurosci.* **37,** 4280–4288 (2017).
23. Gatica, D. *et al.* Association between sleep-related breathing disorders and academic performance among children from Concepcion, Chile. *Arch. Argent. Pediatr.* **115,** 497–500 (2017).
24. Landau, Y. E. *et al.* Impaired behavioral and neurocognitive function in preschool children with obstructive sleep apnea. *Pediatr. Pulmonol.* **47,** 180–8 (2012).
25. Nieminen, P. *et al.* Growth and biochemical markers of growth in children with snoring and obstructive sleep apnea. *Pediatrics* **109,** e55 (2002).
26. Pena-Zarza, J. A. *et al.* Glycated hemoglobin and sleep apnea syndrome in children: beyond the apnea-hypopnea index. *Sleep Breath. Schlaf Atm.* (2017). doi:10.1007/s11325-017-1509-2

27. Kothare, S. V. *et al.* Quality measures for the care of pediatric patients with obstructive sleep apnea. *J. Clin. Sleep Med. JCSM Off. Publ. Am. Acad. Sleep Med.* **11,** 385–404 (2015).
28. Dr Robert C. Wang, Dr Tina P. Elkins, Dr Daniel Keech, Dr Albert Wauquier & Dr Douglas Hubbard. Accuracy of Clinical Evaluation in Pediatric Obstructive Sleep Apnea. *Otolaryngol. Neck Surg.* **118,** 69–73 (1998).
29. Chervin, R. D. *et al.* Pediatric sleep questionnaire: prediction of sleep apnea and outcomes. *Arch. Otolaryngol. Head Neck Surg.* **133,** 216–22 (2007).
30. Brouilette, R. *et al.* A diagnostic approach to suspected obstructive sleep apnea in children. *J. Pediatr.* **105,** 10–14 (1984).
31. Carroll, J. L., McColley, S. A., Marcus, C. L., Curtis, S. & Loughlin, G. M. Inability of clinical history to distinguish primary snoring from obstructive sleep apnea syndrome in children. *Chest* **108,** 610–618 (1995).
32. Chervin, R. D., Hedger, K., Dillon, J. E. & Pituch, K. J. Pediatric sleep questionnaire (PSQ): validity and reliability of scales for sleep-disordered breathing, snoring, sleepiness, and behavioral problems. *Sleep Med.* **1,** 21–32
33. Franco, R. A. J., Rosenfeld, R. M. & Rao, M. First place--resident clinical science award 1999. Quality of life for children with obstructive sleep apnea. *Otolaryngol.--Head Neck Surg. Off. J. Am. Acad. Otolaryngol.-Head Neck Surg.* **123,** 9–16 (2000).
34. Constantin, E., Tewfik, T. L. & Brouillette, R. T. Can the OSA-18 quality-of-life questionnaire detect obstructive sleep apnea in children? *Pediatrics* **125,** e162-168 (2010).

35. Kaditis A., Kheirandish-Gozal L. & Gozal D. Pediatric OSAS: Oximetry can provide answers when polysomnography is not available. *Sleep Med. Rev.* **27,** 96–105 (2016).

36. Lamm Carin, Mandeli John & Kattan Meyer. Evaluation of home audiotapes as an abbreviated test for obstructive sleep apnea syndrome (OSAS) in children. *Pediatr. Pulmonol.* **27,** 267–272 (1999).

37. Jacob S. V. *et al.* Home testing for pediatric obstructive sleep apnea syndrome secondary to adenotonsillar hypertrophy. *Pediatr. Pulmonol.* **20,** 241–252 (2005).

38. Flemons, W. W. *et al.* Home diagnosis of sleep apnea: a systematic review of the literature. An evidence review cosponsored by the American Academy of Sleep Medicine, the American College of Chest Physicians, and the American Thoracic Society. *Chest* **124,** 1543–1579 (2003).

39. Kapur, V. K. *et al.* Clinical Practice Guideline for Diagnostic Testing for Adult Obstructive Sleep Apnea: An American Academy of Sleep Medicine Clinical Practice Guideline. *J. Clin. Sleep Med. JCSM Off. Publ. Am. Acad. Sleep Med.* **13,** 479–504 (2017).

40. Kirk, V. *et al.* American Academy of Sleep Medicine Position Paper for the Use of a Home Sleep Apnea Test for the Diagnosis of OSA in Children. *J. Clin. Sleep Med. JCSM Off. Publ. Am. Acad. Sleep Med.* **13,** 1199–1203 (2017).

41. Horwood, L., Brouillette, R. T., McGregor, C. D., Manoukian, J. J. & Constantin, E. Testing for pediatric obstructive sleep apnea when health care resources are rationed. *JAMA Otolaryngol.-- Head Neck Surg.* **140,** 616–23 (2014).

42. Berry, R. B. *et al.* Rules for scoring respiratory events in sleep: update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events. Deliberations of the Sleep Apnea Definitions Task Force of the American Academy of Sleep Medicine. *J. Clin. Sleep Med. JCSM Off. Publ. Am. Acad. Sleep Med.* **8,** 597–619 (2012).
43. Morgenthaler, T. *et al.* Practice parameters for the use of actigraphy in the assessment of sleep and sleep disorders: an update for 2007. *Sleep* **30,** 519–529 (2007).
44. Meltzer, L. J. *et al.* Validation of Actigraphy in Middle Childhood. *Sleep* **39,** 1219–1224 (2016).
45. Meltzer, L. J., Walsh, C. M., Traylor, J. & Westin, A. M. L. Direct comparison of two new actigraphs and polysomnography in children and adolescents. *Sleep* **35,** 159–166 (2012).
46. Meltzer, L. J., Hiruma, L. S., Avis, K., Montgomery-Downs, H. & Valentin, J. Comparison of a Commercial Accelerometer with Polysomnography and Actigraphy in Children and Adolescents. *Sleep* **38,** 1323–1330 (2015).
47. Toon, E. *et al.* Comparison of Commercial Wrist-Based and Smartphone Accelerometers, Actigraphy, and PSG in a Clinical Cohort of Children and Adolescents. *J. Clin. Sleep Med. JCSM Off. Publ. Am. Acad. Sleep Med.* **12,** 343–350 (2016).
48. Cellini, N., Buman, M. P., McDevitt, E. A., Ricker, A. A. & Mednick, S. C. Direct comparison of two actigraphy devices with polysomnographically recorded naps in healthy young adults. *Chronobiol. Int.* **30,** 691–698 (2013).

49. Levendowski, D. J., Seagraves, S., Popovic, D. & Westbrook, P. R. Assessment of a neck-based treatment and monitoring device for positional obstructive sleep apnea. *J. Clin. Sleep Med. JCSM Off. Publ. Am. Acad. Sleep Med.* **10,** 863–871 (2014).

50. Cole, R. J., Kripke, D. F., Gruen, W., Mullaney, D. J. & Gillin, J. C. Automatic sleep/wake identification from wrist activity. *Sleep* **15,** 461–469 (1992).

51. Stevens, A. The Effect of Sleep Disturbance during Pregnancy and Perinatal Period on Postpartum Psychopathology in Women with Bipolar Disorder. *J. Womens Health Care* **03,** (2014).

52. Bellone, G. J. *et al.* Comparative analysis of actigraphy performance in healthy young subjects. *Sleep Sci. Sao Paulo Braz.* **9,** 272–279 (2016).

53. Shin, M., Swan, P. & Chow, C. M. The validity of Actiwatch2 and SenseWear armband compared against polysomnography at different ambient temperature conditions. *Sleep Sci.* **8,** 9–15 (2015).

54. Dick, R. *et al.* AASM standards of practice compliant validation of actigraphic sleep analysis from SOMNOwatch versus polysomnographic sleep diagnostics shows high conformity also among subjects with sleep disordered breathing. *Physiol. Meas.* **31,** 1623–1633 (2010).

55. Roy, S. Actigraphy Comparison Guide (March 2015). *Sleep Review* Available at: http://www.sleepreviewmag.com/2015/03/actigraphy-comparison-guide-march-2015/. (Accessed: 8th April 2018)

56. Bastanlar, Y. & Ozuysal, M. Introduction to machine learning. *Methods Mol. Biol. Clifton NJ* **1107,** 105–128 (2014).

57. Corrales, M. *et al.* Machine Learning: How Much Does It Tell about Protein Folding Rates? *PloS One* **10,** e0143166 (2015).

58. Kang, J., Schwartz, R., Flickinger, J. & Beriwal, S. Machine Learning Approaches for Predicting Radiation Therapy Outcomes: A Clinician's Perspective. *Int. J. Radiat. Oncol. Biol. Phys.* **93,** 1127–1135 (2015).

59. Zhang, Y. *et al.* Automatic sleep staging using multi-dimensional feature extraction and multi-kernel fuzzy support vector machine. *J. Healthc. Eng.* **5,** 505–520 (2014).

60. Jensen, J. B. *et al.* Sleep-wake transition in narcolepsy and healthy controls using a support vector machine. *J. Clin. Neurophysiol. Off. Publ. Am. Electroencephalogr. Soc.* **31,** 397–401 (2014).

61. Anafi, R. C. *et al.* Machine learning helps identify CHRONO as a circadian clock component. *PLoS Biol.* **12,** e1001840 (2014).

62. Kaplan, K. A., Hardas, P. P., Redline, S. & Zeitzer, J. M. Correlates of sleep quality in midlife and beyond: a machine learning analysis. *Sleep Med.* **34,** 162–167 (2017).

63. Kuczmarski, R. CDC growth charts: United States. *Adv Data* **314,** (2000).

64. Hothorn, T., Hornik, K. & Zeileis, A. Unbiased Recursive Partitioning: A Conditional Inference Framework. *J. Comput. Graph. Stat.* **15,** 651–674 (2006).

65. Bhattacharjee, R. *et al.* Adenotonsillectomy outcomes in treatment of obstructive sleep apnea in children: a multicenter retrospective study. *Am. J. Respir. Crit. Care Med.* **182,** 676–83 (2010).

66. Dabney, A. R. Classification of microarrays to nearest centroids. *Bioinformatics* **21,** 4148–4154 (2005).

67. Venkatasubramaniam, A. *et al.* Decision trees in epidemiological research. *Emerg. Themes Epidemiol.* **14,** 11 (2017).

68. Salis, K. L., Kliem, S. & O'Leary, K. D. Conditional inference trees: a method for predicting intimate partner violence. *J. Marital Fam. Ther.* **40,** 430–441 (2014).

69. Strempel, S., Nendza, M., Scheringer, M. & Hungerbuhler, K. Using conditional inference trees and random forests to predict the bioaccumulation potential of organic chemicals. *Environ. Toxicol. Chem.* **32,** 1187–1195 (2013).

70. Bessonov, K. & Van Steen, K. Practical aspects of gene regulatory inference via conditional inference forests from expression data. *Genet. Epidemiol.* **40,** 767–778 (2016).