THE UNIVERSITY OF QUEENSLAND

AUSTRALIA

Movement characterisation with accelerometry for improved paediatric sleep assessment

Marnie L. Lamprecht

A thesis submitted for the degree of Doctor of Philosophy at
The University of Queensland in 2016

School of Information Technology and Electrical Engineering

Abstract

Actigraphy is increasingly used to non-invasively estimate sleep quality in children with a suspected sleep disorder. Commercial actigraphs summarise wrist movement, conventionally measured with a uni-axial accelerometer, within a fixed epoch (typically 30s). Wake is subsequently identified as epochs of increased activity and sleep is identified as epochs of inactivity. This classification framework has some distinct limitations: actigraphy misclassifies activity during sleep as wake, and inactivity during wake (i.e. quiet rest) as sleep. In this thesis we will address these limitations by investigating three hypotheses. Firstly, uni-axial accelerometry measured solely at the wrist restricts prediction accuracy, since movements orthogonal to the measurement axis, or occurring elsewhere on the body, cannot be detected. Utilising multisite tri-axial accelerometry may consequently improve sleep and wake prediction. Secondly, there are movement characteristics that can differentiate sleep from wake because the physiological nature of these movements differ. Identifying these characteristics may reduce false wake detections. Finally, physiological and pathological events such as apnoeas, hypopneas and transient arousals may be associated with sleep movements that contribute to false wake detections. Exploring this association may consequently explain the presence of some sleep movements.

In order to address the hypotheses, 38 participants (27 male, aged $5 - 16$ years) were recruited from children attending the sleep laboratory for suspected sleep-disordered breathing. These children were studied concurrently with polysomnography and a custom system (synchronised to within 0.1s) that recorded raw tri-axial accelerometry data ($8-$bit, 100Hz, $\pm$2G) simultaneously at the left index fingertip, left wrist, upper thorax, left ankle and left great toe.

The first analysis compared the accuracy of predicting sleep and wake epochs with uni-axial, tri-axial, and multisite accelerometry. Tri-axial versions of the conventional 30s epoch summaries were derived and compared to conventional uni-axial accelerometry. Multisite data were explored and verified using two feature selection algorithms with the tri-axial summaries for each accelerometer. Classification performance was significantly improved when incorporating additional accelerometers, and measuring movement with tri-axial accelerometry (Kappa agreement for multisite, tri-axial and uni-axial accelerometry: 0.565(0.231), 0.402(0.141) and 0.268(0.210), $p < 0.05$). Tri-axial accelerometry has clear benefits with no increase in cost or invasiveness. Although multisite accelerometry provides additional performance benefits, these benefits come at the expense of system complexity and patient discomfort.

Moving away from epoch-by-epoch predictions, the second analysis assessed wake detection on a movement-by-movement basis. Localised spectral characteristics of raw segmented wrist movements were identified using the discrete wavelet transform. Characteristics that significantly differed between sleep and wake movements were then used to predict wake on a movement-by-movement basis. In general, short-duration wake movements had regions of increased spectral energy, were more vigorous, and consistently had spectral content characteristic of limb positional changes. However, predicting wake on a movement-by-movement basis had similar performance to the 30s activity counts (area under the receiver operating characteristics curve: 63.9(6.7)% vs. 69.7(7.9)% respectively). The similar performance of these distinctly different approaches, together with the consistently average predictive performance seen throughout the analyses, shows that movement information cannot accurately estimate sleep in a generalised classification model.

The final analysis explored possible causes of the confounding sleep movements by analysing the temporal association with transient arousals, apnoeas and hypopneas manually scored from polysomnography. On average, 21.4% of apnoeas, 40.8% of hypopneas and 67.5% of arousals coincided with wrist movement. However, the prevalence and corresponding associations varied considerably across the cohort. Arousals during sleep that were associated with movement were generally longer than other arousals (12.2s vs. 7.9s, $p < 0.01$). Similarly, movements that occurred during an arousal were longer than other sleep movements (9.56s vs. 2.35s, $p < 0.01$). The association between lengthy arousals and lengthy sleep movements suggests that these longer arousals contribute to false wake detections. Although actigraphy cannot predict all arousals, it can likely predict the lengthier arousals that disrupt sleep.

We can conclude from the analyses in this thesis that multisite tri-axial accelerometry offers distinct performance benefits for sleep assessment; however, the associated practical compromise from additional accelerometers may not be appealing for a device targeted at home-based sleep assessment. Transient arousals are strongly associated with the lengthier sleep movements that confound sleep estimates with commercial actigraphy. Considering that these arousals are characteristic of sleep disturbance, and actigraphy likely detects these events, incorporating the detection of these longer sleep movements into the actigraphy scoring routine may capture the severity of sleep disturbance associated with a sleep disorder. Existing actigraphy systems only estimate sleep quality; however, future actigraphy systems will likely benefit from identifying signs indicative of sleep disorder severity.

Declaration by author

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my research higher degree candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the policy and procedures of The University of Queensland, the thesis be made available for research and study in accordance with the Copyright Act 1968 unless a period of embargo has been approved by the Dean of the Graduate School.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate I have obtained copyright permission from the copyright holder to reproduce material in this thesis.

Publications during candidature

Peer-reviewed Publications

M. L. Lamprecht, P. I. Terrill, C. L. Parsley, and A. P. Bradley, "Characterization of Movements during Restless Sleep in Children: A Pilot Study", in *Engineering in Medicine and Biology Computing, 2014. EMBC 2014, 36th Annual International Conference of the IEEE, 2014*, pp 274 -277.

M. L. Lamprecht, A. P. Bradley, T. Tran, A. Boynton, and P. I. Terrill, "Multisite accelerometry for sleep and wake classification in children", *Physiological Measurement*, *vol. 36*, *no. 1*, pp. 133-147, Jan. 2015.

M. L. Lamprecht, A. P. Bradley, G. Williams, and P. I. Terrill, "Temporal associations between arousal and body/limb movement in children with suspected obstructed sleep apnoea", *Physiological Measurement*, *vol. 37*, *no. 1*, pp. 115-127, Dec. 2015.

Conference Abstracts

M. L. Lamprecht, T. Tran, C. Parsley, J. Greenhill, C. Freakley, G. Williams, S. Suresh, C. Dakin, S. Wilson, A. P. Bradley, P. I. Terrill, "A novel actigraphy system detects sleep and wakefulness better than a commercial actigraph", *Australian Biomedical Engineering Conference, ABEC2012.*

M. L. Lamprecht, T. Tran, J. Greenhill, C. Parsley, C. Freakley, G. Williams, S. Suresh, C. Dakin, S. Wilson, A. P. Bradley, P. I. Terrill, "A novel continuous multisite accelerometry system discriminates sleep from wake better than a commercial actigraph", *Australian Sleep Association, Sleep DownUnder 2012.*

M. L. Lamprecht, T. Tran, G. Williams, S. Suresh, A. P. Bradley, P. I. Terrill, "Combining multiple actigraph placements for sleep assessment in children", *Australian Sleep Association, Sleep DownUnder 2013.*

M. L. Lamprecht, P. I. Terrill, S. Suresh, C. L. Parsley, and A. P. Bradley, "Sleep and wake assessment with a movement detector: improving actigraphy for children with fragmented sleep", *Australian Sleep Association, Sleep DownUnder 2014.*

M. L. Lamprecht, P. I. Terrill, G. Williams, and A. P. Bradley, "Impact of transient arousals on sleep staging with actigraphy", *Australian Sleep Association, Sleep DownUnder 2015.*

Publications included in this thesis

M. L. Lamprecht, P. I. Terrill, C. L. Parsley, and A. P. Bradley, "Characterization of Movements during Restless Sleep in Children: A Pilot Study", in *Engineering in Medicine and Biology Computing, 2014. EMBC 2014, 36th Annual International Conference of the IEEE, 2014*, pp 274 -277.

Incorporated as most of Chapter 5.

Contributor	*Statement of Contribution*
Marnie L. Lamprecht (Candidate)	Designed experiment (90%) Configured equipment and recruited patients (50%) Implemented analysis (100%) Interpreted results (90%) Wrote and edited paper (90%)
Philip I. Terrill	Designed experiment (10%) Interpreted results (5%) Wrote and edited paper (5%)
Chloe L. Parsley	Configured equipment and recruited patients (50%)
Andrew P. Bradley	Interpreted results (5%) Wrote and edited paper (5%)

M. L. Lamprecht, A. P. Bradley, T. Tran, A. Boynton, and P. I. Terrill, "Multisite accelerometry for sleep and wake classification in children", *Physiological Measurement*, *vol. 36*, *no. 1*, pp. 133-147, Jan. 2015.

Incorporated as part of Chapter 4.

Contributor	*Statement of Contribution*
Marnie L. Lamprecht (Candidate)	Designed experiment (80%) Configured equipment and recruited patients (50%) Implemented analysis (100%) Interpreted results (90%) Wrote and edited paper (85%)
Andrew P. Bradley	Interpreted results (5%) Wrote and edited paper (5%)
Tommy Tran	Designed experiment (15%)
Alison Boynton	Configured equipment and recruited patients (50%)
Philip I. Terrill	Designed experiment (5%) Interpreted results (5%) Wrote and edited paper (10%)

M. L. Lamprecht, A. P. Bradley, G. Williams, and P. I. Terrill, "Temporal associations between arousal and body/limb movement in children with suspected obstructed sleep apnoea", *Physiological Measurement*, *vol. 37*, *no. 1*, pp. 115-127, Dec. 2015.

Incorporated as part of Chapter 6.

Contributor	*Statement of Contribution*
Marnie L. Lamprecht (Candidate)	Designed experiment (80%) Configured equipment and recruited patients (50%) Implemented analysis (100%) Interpreted results (80%) Wrote and edited paper (85%)
Andrew P. Bradley	Designed experiment (5%) Interpreted results (10%) Wrote and edited paper (5%)
Gordon Williams	Configured equipment and recruited patients (50%)
Philip I. Terrill	Designed experiment (15%) Interpreted results (10%) Wrote and edited paper (10%)

Contributions by others to the thesis

No additional contributions by others.

Statement of parts of the thesis submitted to qualify for the award of another degree

None.

Acknowledgements

I would like to thank Dr. Philip Terrill for his enthusiasm and Prof. Andrew Bradley for his steady reassurance. I am grateful to both for offering an open door and sharing their knowledge. Without their invaluable guidance, enduring patience and implacable confidence, I would not have completed this thesis.

I would also like to thank my family and friends for their patience, encouragement and support. Thank you for the happy distractions.

Keywords

Actigraphy, sleep disturbance, sleep fragmentation, automated sleep scoring, time-series signal analysis, wavelet analysis, sleep disorder detection

Australian and New Zealand Standard Research Classification (ANZSRC)

ANZSRC code: 090609, Signal Processing, 50%
ANZSRC code: 090303, Biomedical Instrumentation, 30%
ANZSRC code: 090304, Medical Devices, 20%

Fields of Research (FoR) Classification

FoR code: 0903, Biomedical Engineering, 60%
FoR code: 0906, Electrical and Electronic Engineering, 40%

Contents

List of Figures

List of Tables

Acronyms & Abbreviations

AASM American Academy of Sleep Medicine
AHI apnoea hypopnea index, *Glossary:* AHI
AIC Akaike information criterion
ANOVA analysis of variance, *Glossary:* ANOVA
ASA Australian Sleep Association
ASDA American Sleep Disorders Association
ASTA Australian Sleep Technologists Association
AUC area under the receiver operating characteristics curve

BPF band-pass filter

CMAS Continuous Multisite Accelerometry System

DC mean value of the signal
DI digital integration movement representation technique

EEG electroencephalogram, *Glossary:* electroencephalography
EMG electromyogram, *Glossary:* electromyography
EOG electroculogram, *Glossary:* electrooculography

GUI graphical user interface

HSD Tukey's honest significant difference test, *Glossary:* HSD

IQR interquartile range

LPF low-pass filter

MAXACT maximum magnitude of acceleration
MEMS micro electro-mechanical systems
mRMR minimum Redundancy Maximum Relevancy

N1 non-REM sleep stage 1
N2 non-REM sleep stage 2

N3 non-REM sleep stage 3
N4 non-REM sleep stage 4
NLMS normalised least mean squares
NREM non-rapid eye movement

OCDWT over complete discrete wavelet transform
OSA obstructive sleep apnoea syndrome, *Glossary:* sleep apnoea

PAUC partial receiver operating characteristic area under curve
PC personal computer
PSG polysomnography, *Glossary:* polysomnography

QDA quadratic discriminant analysis

RDI respiratory disturbance index
REM rapid eye movement
RERA respiratory effort-related arousal, *Glossary:* RERA
ROC receiver operating characteristics

SNR signal to noise ratio, *Glossary:* signal to noise ratio
SUMPST integrated angle of posture change

TAT time above threshold movement representation technique

WASO wake after sleep onset

ZC zero crossing movement representation technique

1 Introduction

"Sleep exists for the preservation of animals, and the waking state is its final cause and purpose."

— Parva Naturalia, Aristoteles, *Philosopher, 384 - 322 BC* [1]

Sleep is an essential process for maintaining general health and facilitating daily function. It is considered vital for memory consolidation, muscle restoration, and general quality of life [2, 3]. Among many other detrimental factors, disrupted sleep often manifests as poor temperament, fatigue, decreased motor function and decreased cognitive ability. Sleep disorders are disruptive to sleep because they impair vital physiological processes, incite unwanted movements, or incite involuntary speech [4, 5]. One such sleep disorder is obstructive sleep apnoea syndrome (OSA), which affects approximately $2 - 4\%$ of children [6, 7]. These children experience a complete cessation (apnoea) or partial restriction (hypopnea) of breathing throughout the night. Airway obstructions are often, but not always, caused by enlarged tonsils and/or adenoids [7]. These respiratory obstructions are cleared by subconscious physiological processes that often cause either abnormal periods of wake or transient arousals. In addition to indirectly disrupting sleep, respiratory obstructions reduce blood oxygen levels [8], which inhibits brain development and limits general cognitive ability [9, 10]. If left untreated, OSA persists into adulthood, where the health consequences can be more severe [11]. Therefore, it is essential for development and general well-being that children are diagnosed and treated early.

[1] Philosophy on sleep and dreaming detailed in Aristoteles and Hammond [1].

The advancement of technology in the 1900s facilitated the development of sleep disorder diagnostic systems. Since Smith [12] first observed distinct changes in brain activity during the transition from wake to sleep in the 1930s, the electroencephalogram (EEG) was increasingly used to assess sleep architecture. EEG was progressively refined and combined with other sensors that monitor physiological processes during sleep: including, but not limited to, heart rhythm, blood oxygenation, respiratory effort, and muscle activation. This collection of sensors has evolved into polysomnography: the current 'gold standard' for sleep assessment. The various sensors ensure that polysomnography accurately monitors the relevant physiological processes [13, 14]. But, this results in an invasive system that can cause poor sleep or abnormal sleep behaviour. Therefore, the estimated sleep behaviour may not accurately represent the child's actual sleep behaviour. Polysomnography also requires a trained sleep technician to configure and continually monitor the system in a dedicated sleep laboratory. These facilities are commonly located in major cities, which is particularly problematic in sparsely populated countries such as Australia. The limitations of polysomnography motivate the development of alternative sleep disorder diagnostic systems that are ideally non-invasive, inexpensive and user-friendly.

Alternative diagnostic methods typically measure a subset of the physiological signals available with polysomnography. These methods include qualitative techniques such as sleep diaries and questionnaires, as well as quantitative techniques such as blood oxygenation, respiratory effort, oxygen desaturation, and body movement. Unfortunately, the most effective systems are typically the most invasive and complex. The non-invasive alternatives generally only describe one aspect of OSA: respiratory cessations or sleep quality. Actigraphy, an existing method for estimating sleep quality using body movement, has the potential to describe both aspects. Actigraphy currently detects wake on the basis of increased activity. In addition to this, it may be possible to detect respiratory cessations or physiological events characteristic of OSA if they are associated with movement. Actigraphy is also non-invasive, inexpensive, and user-friendly. However, actigraphy has distinct technical limitations that impact the accuracy of predicted sleep and wake regions.

Sleep actigraphy systems conventionally detect wrist movement using a uni-axial accelerometer, and summarise fixed epochs (typically 30s) of movement as 'activity counts' [15]. Essentially, epochs of activity are identified as 'wake' and epochs of inactivity are identified as 'sleep'. This framework misidentifies sleep epochs with high activity as 'wake', and wake epochs with low activity as 'sleep'; consequently, existing actigraphs have distinct limitations when monitoring patients with disorders that cause movement (i.e. OSA), or extended regions of quiet wake (i.e. insomnia). Several studies have attempted to address these confounding regions [16, 17]. Despite known relationships between specific movements and sleep stages [18, 19], these attempts have focused on optimising filter coefficients for post-processing activity counts (documented in Fig. 1.2). The filter combines movement information from surrounding epochs,

effectively smoothing large epochs into surrounding epochs. This smoothing process has two main effects: 'peaky' movements during sleep are attenuated, increasing the likelihood of isolated high-activity sleep epochs correctly scoring as sleep; and extended movements during wake are distributed into surrounding epochs, increasing the likelihood of isolated low-activity wake epochs correctly scoring as wake. However, passively smoothing epochs introduces errors by reducing activity in wake epochs with neighbouring epochs of low-activity, and increasing activity in sleep epochs with neighbouring epochs of high-activity. Another limitation of activity counts is the inadequate temporal resolution for detecting specific movements or isolating specific characteristics.

1.1 Research rationale

Movement during sleep and wake is conventionally represented using a summary of movement within a fixed epoch, termed 'activity counts'. Regions of wake are classified as epochs of high activity, and regions of sleep are classified as epochs of low activity. The major limitation of this framework is summarised in Fig. 1.1, which shows the distribution of activity counts during sleep and wake: low-activity epochs occur during wake, and high-activity epochs occur during sleep. This results in overlapping distributions that cause classification errors when the standard threshold-classifier is applied. In order to address the limitations of actigraphy in sleep assessment (detailed in Fig. 1.2), we will consider the two key error types (highlighted in Fig. 1.1 and summarised in Fig. 1.3):

Error 1. False negatives: wake epochs with no observed movement are incorrectly identified as 'sleep'.

False sleep detections may be caused by suboptimal sensor configuration: existing actigraphy systems use a single accelerometer (most often located on the dominant or non-dominant wrist) to measure movement. A single accelerometer cannot detect movements in other limbs. Incorporating additional accelerometers to detect these movements may reduce the risk that wake movements are missed. In addition to this, existing systems likely still employ uni-axial movement representation techniques. Monitoring movement with a single axis would increase the risk that genuine movements are not detected. Therefore, representing movement with techniques that summarise tri-axial motion may further reduce false sleep detections.

It is possible for actual regions of no movement to occur during wake (i.e. where the patient is awake but laying still). Since these epochs are likely surrounded by epochs of activity, the conventional smoothing filter will reduce these false sleep detections.

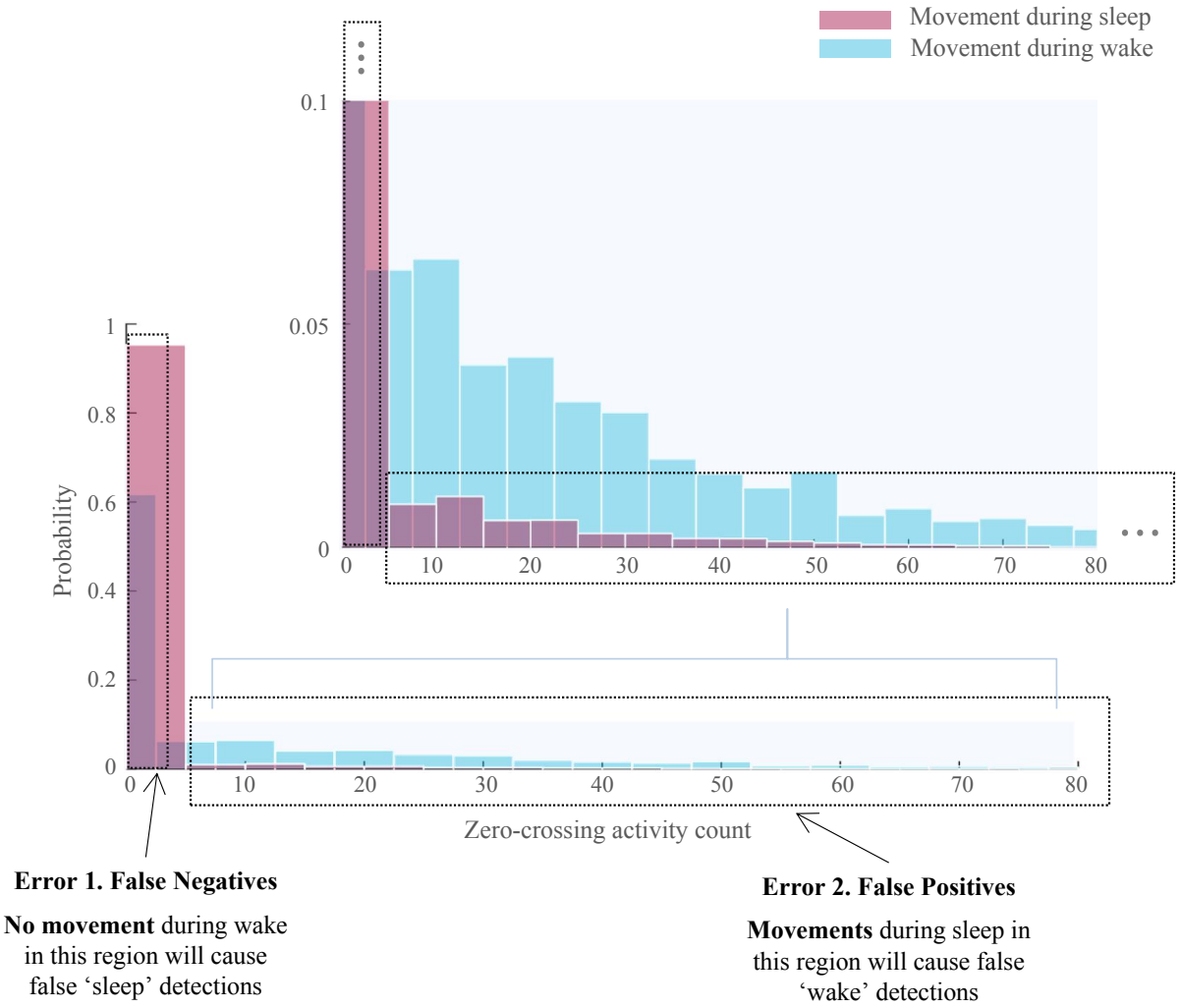


Figure 1.1: Histogram of activity during sleep (red) and wake (blue) for 24 patients using the conventional activity counts derived using the zero-crossing method.

Error 2. False positives: sleep epochs with observed movement are incorrectly identified as 'wake'.

False wake detections may be caused by inadequate movement representations: wrist activity is typically summarised within fixed 30s epochs using a time-series technique with uni-axial data. Representing movement in this way limits the ability to discriminate between sleep and wake movements because physiologically different movements can result in a similar activity value, and some movements cannot be detected altogether. Therefore, measuring movement with tri-axial multisite accelerometry, and identifying characteristic differences between sleep and wake movements, may reduce false wake detections by providing a better movement descriptor than the conventional techniques.

Movements during physical activity or associated with physiological disorders are well defined [20, 21]; however, descriptions of movements during sleep are currently lacking. The

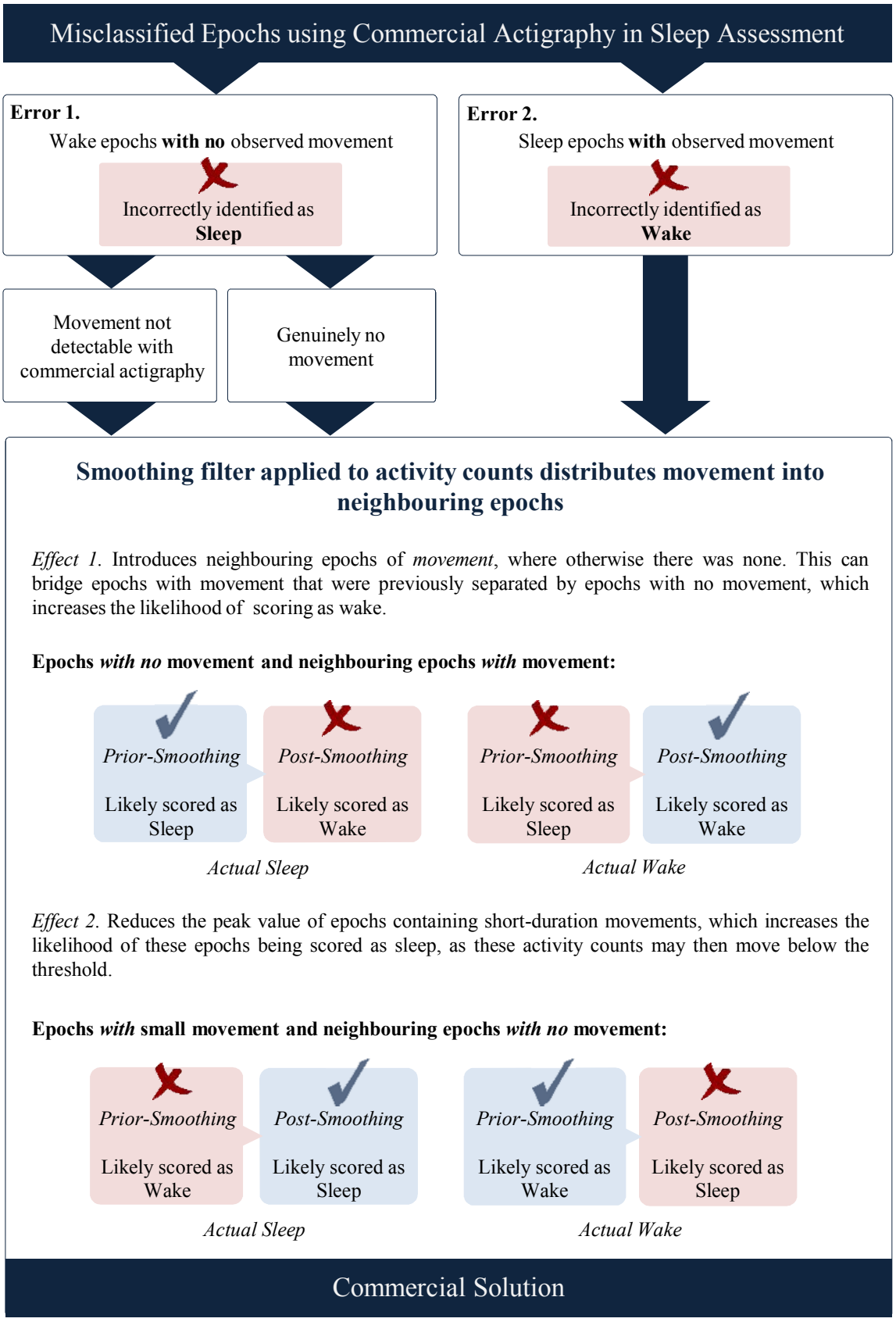


Figure 1.2: Typical process used by commercial actigraphy systems to address the limitations of actigraphy in sleep assessment.

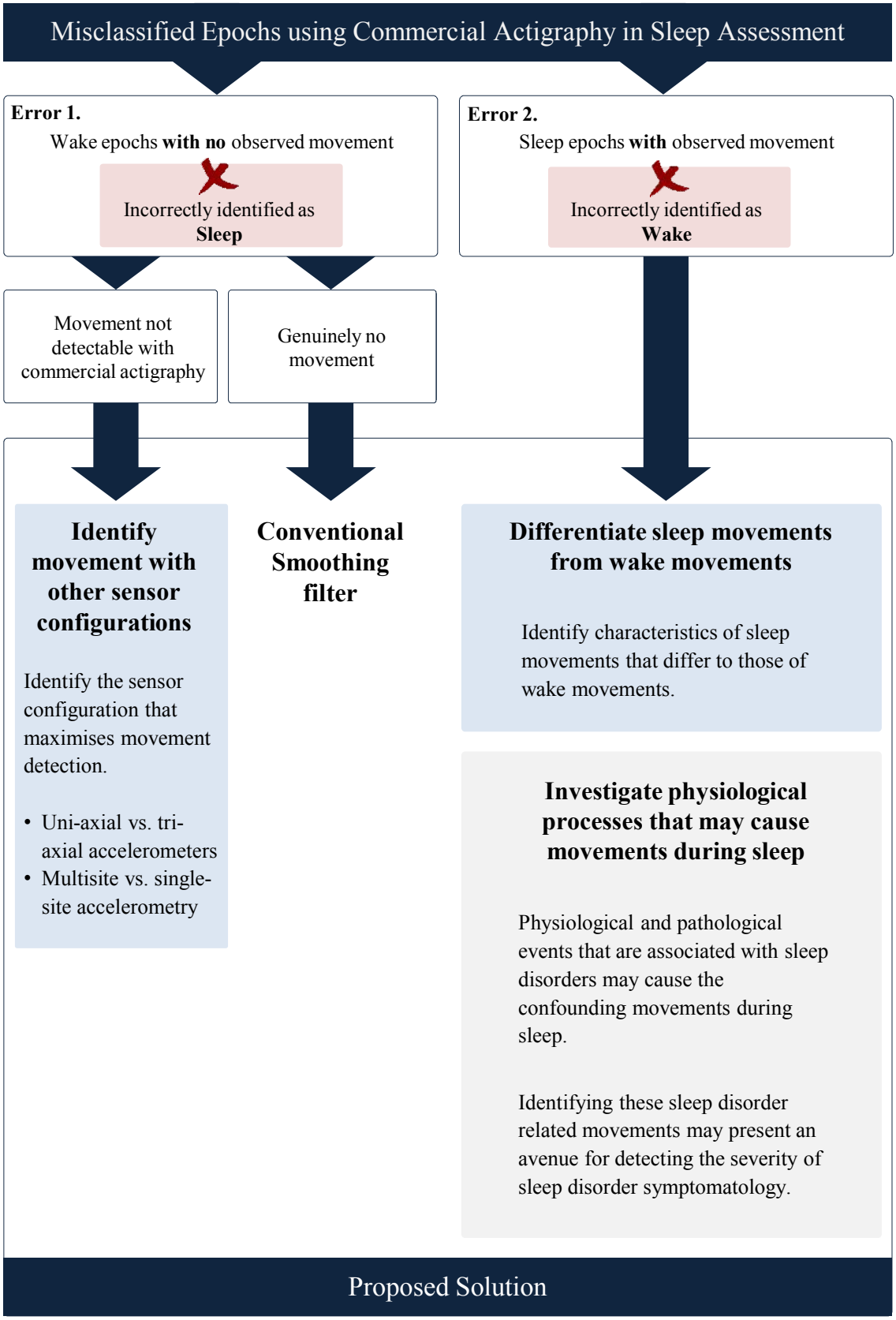


Figure 1.3: Proposed solution for addressing the limitations of actigraphy in sleep assessment.

conventional summaries of movement persist throughout sleep literature, despite evidence of movement characteristics specific to certain sleep stages [22], differences in muscle activation effort between sleep and wake [23], and differences in movement coherence [18]. Considering that literature highlights these differences between movements during sleep and wake, there are likely technically differentiable characteristics. The conventional activity summaries are simple to implement and are used throughout commercial systems. Since existing systems likely still employ uni-axial movement representations, despite incorporating tri-axial accelerometers, the range of detectable movements is limited. Incorporating tri-axial movement representations and identifying differentiable characteristics using the common summary techniques will provide the greatest short-term gain with negligible changes to existing systems.

The conventional movement summaries are unable to identify short-duration movements, or identify specific characteristics because they have low temporal resolution (in the order of 30s). Although using these summaries is ideal for integration into existing systems, analysing characteristics on a movement-basis with raw high-resolution data is likely to provide greater performance benefits by facilitating the discrimination of sleep and wake movements of varying durations. Raw data also allows time-varying spectral characteristics to be considered. Localised spectral techniques may isolate physiologically different characteristics of sleep and wake movements, which consequently may reduce false wake detections.

Actigraphy in sleep assessment is used to estimate sleep quality by identifying sleep and wake epochs. Literature questions the accuracy of these estimates, partly because of the confounding sleep movements. However, it is possible that these confounding movements correlate with physiological or pathological events (such as transient arousal or apnoea/hypopnea) associated with sleep disorders. Instead of confounding wake estimates, these movements may provide an avenue for identifying signs indicative of sleep disorder severity.

1.2 Overall aims and outline of thesis

The overarching goal of this thesis is to explore the utility of an improved actigraphy system for assessing paediatric sleep disorders in the home environment. To achieve this, the two primary limitations of actigraphy in sleep assessment will be addressed. Chapter 4 addresses the first limitation (wake epochs with no observed movement) by investigating if adopting tri-axial/multisite accelerometry reduces the number of wake epochs with no movement. Chapter 5 and Chapter 6 address the second limitation (sleep epochs with movement resulting in activity values similar to wake epochs). Chapter 5 identifies movement characteristics that differ

between sleep and wake, and predicts sleep and wake on a movement-by-movement basis. Chapter 6 identifies causes of some of the confounding sleep movements by analysing the correlation with physiological and pathological events associated with sleep disorders. The overview of each chapter is outlined below.

Chapter 2

Overview of sleep assessment

- Introduces obstructive sleep apnoea syndrome and the relevant diagnostic techniques;
- Introduces the clinical and technical concepts required for data analysis; and
- Discusses the technical limitations of actigraphy for identifying sleep issues associated with sleep apnoea.

Chapter 3

Experimental methodology and materials

- Outlines the patient characteristics;
- Outlines the clinical study design;
- Documents the post-processing procedure for the accelerometry data and the statistical analysis procedures;
- Outlines the final data set formats for analysis; and
- Describes the general outcome measures of the thesis.

Chapter 4

Conventional representations of tri-axial multisite accelerometry

- Compares the effectiveness of predicting sleep and wake using the conventional representations of movement with a custom accelerometry system;
- Compares the performance of uni-axial and tri-axial accelerometry for summarising movement associated with sleep and wake; and
- Compares the sleep and wake discrimination ability and the sleep and wake predictive performance of multiple accelerometer placements across the body using time-series representations.

Chapter 5

Differentiating sleep and wake movements

- Identifies and explores the effect of movements specific to restless sleep on conventional activity counts;
- Identifies discriminatory localised spectral characteristics of sleep movements and wake movements;
- Determines the sleep and wake predictive performance of the identified time-frequency characteristics; and
- Explores the effectiveness of detecting regions of sleep and wake on a movement-by-movement basis, rather than the conventional 30s activity summaries.

Chapter 6

Physiological associations with sleep movements

- Explores the association between transient arousal and movement, and thereby identifies if arousals confound conventional actigraphy-derived wake scores and if actigraphy can predict arousal events; and
- Explores the association between apnoeic events and movement and thereby identifies if sleep movements predict apnoeic events.

Chapter 7

Summary

- Discusses the thesis contributions to the field of actigraphy in sleep assessment;
- Discusses the clinical implications of the results in this thesis;
- Identifies the limitation of the analyses; and
- Identifies areas for future research.

2

Overview of sleep assessment

Approximately 400 BC, Plato considered sleep to be caused by rising digestive stomach vapours that block the sensory pores in the brain, disconnecting it from the body, and thereby inducing sleep. Despite discussing why or how we sleep, Plato thought it neither necessary for soul nor body; 'indeed, asleep we have no more value than a dead person'.

— Plato, *Philosopher, 428 - 348 BC* [1]

Sleep is an essential process that contributes to quality of life [2]. Unfortunately, for many people sleep is often disrupted throughout the night by one of many sleep-related disorders. These disorders range from narcolepsy (the seemingly spontaneous onset of sleep) and insomnia (the inability to sleep) to parasomnias and breathing disorders [8]. Although all of these disorders impact health and daily functioning, we will focus on the issues surrounding diagnosing obstructive sleep apnoea syndrome (OSA).

This chapter introduces the clinical and technical background behind sleep assessment for diagnosing OSA in children. We will start by introducing the theory and clinical practice behind sleep scoring and OSA. We will then discuss the techniques currently used to assess OSA and their associated limitations. Finally, we will identify techniques that may improve these limitations.

[1] Historical information gathered from literature studies detailed in Wiedemann and Dowden [24].

2.1 Introduction to sleep

2.1.1 Sleep and scoring

Sleep is a complex process that is essential for healthy functioning [25]. Sleep is characterised, relative to wake, by an altered consciousness and a restriction of sensory intake and voluntary muscular control. Since Aserinsky and Kleitman [26] first observed rapid eye movements during sleep in 1953, researchers have sought to determine the role of sleep. This observation launched sleep research because, before then, it was considered that no great physiological process happened while our bodies were asleep; sleep was thought as merely a cessation of brain activity [23]. There are many theories as to why we sleep; however, this is still an on-going discussion. Currently, it is thought that sleep may:

- aid with memory consolidation [27, 28, 29];
- allow brain cells to repair [23, 27];
- aid with maintaining the integrity of the brain region responsible for sustained attention [27]; and/or
- aid in body restoration and energy conservation [30].

Since each of these processes are required for maintaining healthy functioning, it is essential that sleep is unimpaired and regularly achieved.

The required amount of sleep differs from infancy through to adulthood and old-age. Approximately 60% of each day (14.5 hours) is required for sleep during infancy, 40% (9.5 hours) during adolescence and 33% (8 hours) during adulthood [31]. The required sleep duration steadily reduces throughout adulthood. Sleep is greatly influenced by the environment, particularly light and darkness (i.e. circadian rhythm). Indeed, sleep is commonly impaired when changing between daylight savings and standard time [32, 33]. An extreme example of this influence is the significant phase shift in circadian rhythm experienced by those living in the Arctic region, who are exposed to full days of sunlight in summer, and next to no sunlight in winter [34]. This influence is also often experienced as jet-lag by world-wide travellers, flight attendants and pilots [8, 33].

Sleep is comprised of six recognised stages (as defined by the American Academy of Sleep Medicine (AASM) [38]): wake, rapid eye movement (REM), non-REM sleep stage 1 (N1), non-REM sleep stage 2 (N2), non-REM sleep stage 3 (N3), and non-REM sleep stage 4 (N4). In healthy adults and children, these stages cycle sequentially through N1, N2, N3, N4 and REM every $90 - 110$ minutes throughout the night [35, 39], with the occasional arousal to the wake stage (as illustrated by an example of a child's typical sleep cycle in Fig 2.1). Each sleep stage exhibits different physiological characteristics, ranging from brain activity, eye movements, and muscle tone to dream intensity and frequency. The proportion and characteristics of each stage are specified in Fig. 2.2.

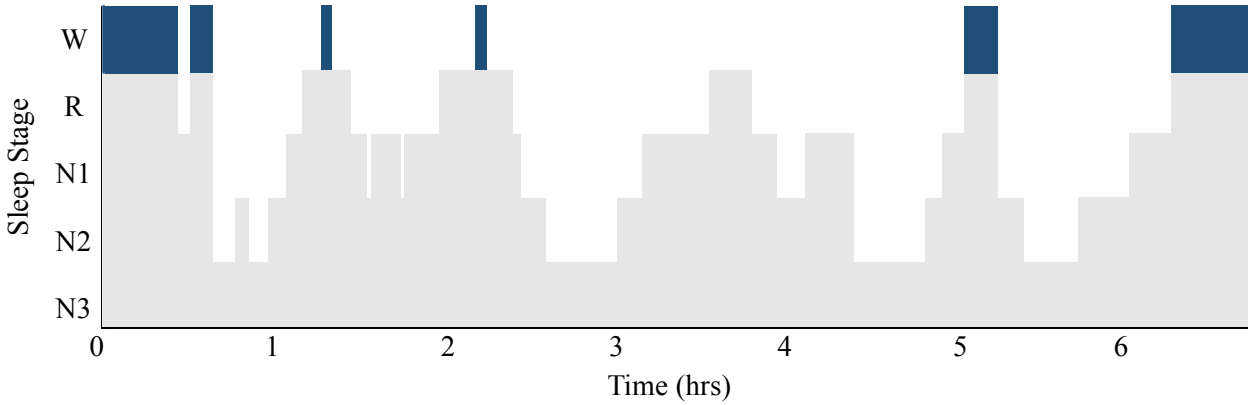


Figure 2.1: Example hypnogram (sleep architecture) of 'normal' sleep cycles across a single night (shown as hours since lights off).

The physiological differences in the sleep stages led to characterisation and consequently the development of scoring rules for each stage. These rules were first defined by Rechtschaffen and Kales [40] in 1968, commonly termed the *R&K scoring guidelines*. The scoring criteria differs for infants (< 2 months of age), children (2 months to 18 years) and adults. Table 2.1 outlines the scoring criteria for each sleep stage for children [40, 13, 41]. The R&K scoring guidelines have been updated by the AASM in 2007 [38] and again in 2013 [42] to address and reflect changes in the sleep research field within the 40 years since the release of the original R&K guidelines. The Australian Sleep Technologists Association (ASTA) and Australian Sleep Association (ASA) have also provided suggestions for amendments to the AASM scoring criteria [43, 44]. These suggestions are also noted in Table 2.1.

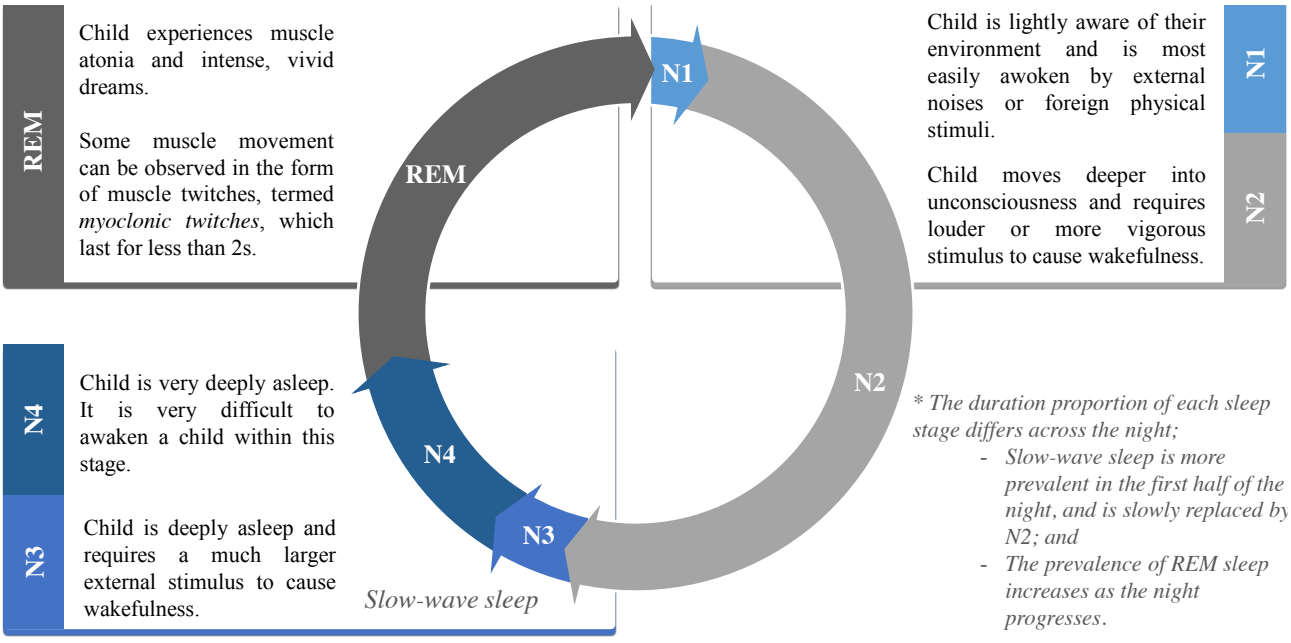


Figure 2.2: Proportion, sequence and physiological characteristics of each sleep stage across the night [8, 35, 36, 37, 22].

Table 2.1: ASA/ASTA recommendations for scoring paediatric sleep [40, 43, 44]

Sleep Stage	Description
Wake, W	Electroencephalogram (EEG) characterised by low amplitude, high frequency waves. When preparing for sleep, EEG alpha activity ($8-13$Hz) is prominent. Electroculogram (EOG) shows voluntary eye movement. Electromyogram (EMG) shows tonic muscle activity related to voluntary movements.
Non-REM sleep, N1	Loss of alpha activity, replaced by mixed-frequency EEG activity with mostly theta activity ($3-7$Hz). EOG shows slow and rolling eye movements. EMG shows relaxed muscle tone. Sudden muscle contractions can occur.
Non-REM sleep, N2	Appearance of K-complexes (high-amplitude negative sharp wave followed by a slow wave) and sleep spindles (oscillations of $12-15$Hz of 1 second in duration that occur $5-10$ times a minute) in the EEG recording. *ASTA/ASA amendment:* the end of NREM Stage 2 is identified as occurring when there is either a transition to a different sleep stage or if the majority of the epoch meets N1 criteria, and a cortical arousal and/or major body movement occurs.
Non-REM sleep, N3	Slow delta waves ($<$ 2Hz) in the EEG recording. No eye movements and reduced EMG activity.
Non-REM sleep, N4	Similar to N3.
REM sleep, REM	EEG activity similar to wake or N1. Loss of muscle tone and increased myoclonic twitches.

Each sleep stage is important for effective daily functioning. If deprived of one of the sleep stages, or if awakened at set intervals throughout a normal sleep period, daytime sleepiness, fatigue and increased emotional instability will often be reported the following day [45]. Physiologically, sleep deprivation impairs temperature control, dietary metabolism and immune function [46]. Prolonged sleep deprivation can, in extreme cases, lead to death [25]. In children, sleep deprivation negatively impacts cognitive and emotional behaviour and can cause hyperactivity or aggression [47, 48]. If not addressed early, the effects of prolonged sleep impairment can lead to permanent deficiencies [27]. As well as external stimulus, sleep deprivation is commonly instigated by some sleep-related disorder or by an abnormally high number of arousal events.

2.1.2 Transient arousals

Arousals can occur as a response to unwanted sleep abnormalities: cortical arousals can occur from external stimuli (e.g. unfamiliar noise, environment disruptions and pain) and respiratory-

Table 2.2: Arousal scoring criteria for children [42, 43, 44]

Arousal

According to the AASM, to score an arousal event,

- there must be at least 10 continuous seconds of any stage of sleep prior to an arousal;
- the EEG frequency shift must be greater than 3 seconds in duration;
- NREM: arousals may occur *without* EMG amplitude increase; and
- REM: arousals must be accompanied by EMG amplitude increase.

ASTA/ASA amendments:

In addition to the AASM criteria,

- a movement-related arousal is scored if there is $0.5s$ between the cessation of limb movement and an AASM arousal event; and
- a respiratory-related arousal is scored if it occurs less than $5s$ or two respiratory cycles after the termination of a respiratory event.

related events (e.g. after a certain level of inspiratory effort or airway pressure is achieved) [49, 4]. Arousals transitioning to wake, or to a lighter sleep stage, can either occur briefly (commonly termed 'transient arousal') or completely, leading to extended periods of wake. There are two distinct arousal types: spontaneous and respiratory-related arousals. Respiratory effort-related arousal (RERA) is defined as an arousal that occurs closely after airflow restrictions (i.e. apnoea or hypopnea). Spontaneous arousals are those that are not caused by any stimuli such as respiratory events, limb movements, snoring, etc. These arousals occur spontaneously, lending themselves as a natural process of sleep.

Similar to sleep staging, arousal events are scored using the AASM manual [38, 42]. In 1968 Rechtschaffen and Kales [40] outlined movement arousals in the R&K Scoring Manual as an intended aid to scoring stages. However, there was no mention of concurrent EEG scoring. An increased interest in the correlation of arousals and daytime sleepiness motivated the need to standardise arousal scoring [50]. In 1992, the ASDA [51] developed criteria for scoring arousals, independent of the Rechtschaffen and Kales [40] sleep scoring criteria. In addition to this, the ASTA/ASA have suggested amendments to the AASM scoring criteria for some arousal events [43, 44]. The AASM scoring rules for arousals and the ASTA/ASA amendments are detailed in Table 2.2.

Frequent arousals contribute to increased sleep disruption throughout the night, and are consequently detrimental to general health [13, 52]. Arousals can coincide with events related

to some sleeping disorders, such as OSA [52]. These disorders also severely affect sleep quality and will be described in the following section.

2.1.3 Sleeping disorders in children

Many children experience difficulties during sleep; from parasomnias (e.g. night terrors and sleep walking) to insomnia or sleep apnoea. These sleep problems disrupt the restorative quality of sleep and often negatively affect the child's daily functioning and general health. Sleep-related disorders in children range from sleep talking, enuresis and night terrors to restless leg syndrome and sleep apnoea [2]. Although all sleeping disorders are detrimental to general health and well-being, sleep apnoea is relatively common in children, requires treatment and requires complex and expensive procedures for diagnosis (discussed in Section 2.2). For these reasons, in this thesis we will focus on aiding the detection of OSA in children. Although OSA will be the focus, the techniques developed in this thesis may be extended to other disorders.

OSA is a sleep-related breathing disorder where breathing is restricted at intervals throughout the night. In addition to the effects of disrupted sleep, sleep-related breathing disorders, such as OSA, can permanently reduce cognitive function [53, 54, 55]. One theory for this impairment is hypoxia (i.e. reduced blood oxygenation) that is caused by a cessation or abnormality in respiration [30]. Other causes may be arousals that coincide with apnoeic events, or considerable expenditure of energy required to regain respiration. In extreme cases, children have presented with permanent brain damage, temporary neurological dysfunction, hypertension, heart failure and respiratory failure [56]. Death can also occur; however, it is uncommon [9]. Unfortunately, parents can dismiss the symptoms of sleep breathing disorders in children as general behavioural problems [57, 7]. In these cases, these disorders will go undetected and the effects of disturbed sleep and/or restricted airflow can worsen and continue into adulthood.

OSA is common in childhood, with $2-4\%$ of children affected (approximately one third that of asthma) [58, 59]. An obstructive event occurs when the upper airways are partially or completely obstructed throughout the night [60]. The leading causes of OSA in children are adenotonsillar hypertrophy and obesity [61, 62]. Enlarged tonsils and adenoids in the upper airways restrict the ability to breathe. An example of an airflow obstruction caused by enlarged tonsils and adenoids is shown in Fig. 2.3. Similar to adenotonsillar hypertrophy, the additional fat deposits in the neck region that are common in obese children can cause respiratory muscle collapse [62]. Airflow can also be restricted by pharyngeal muscle relaxation from atonia during REM sleep [63]. These restrictions can be partial (hypopnea: $> 30-50\%$ fall in respiratory signal amplitude relative to baseline) or complete (apnoea: $> 90\%$ fall in respiratory signal amplitude relative to baseline) [38].

Apnoea and hypopnea events are categorised as obstructive, central or mixed, using the AASM criteria outlined in the 'AASM Manual for the Scoring of Sleep and Associated Events'. The criteria are summarised in Table 2.3 [42]. Children are scored differently to adults because

their respiratory characteristics and physiology differ. The scoring rules for children are used for infants (> 2 months of age) and children less than 18 years of age. The adult scoring rules can be used, at the discretion of the sleep technician, for children aged $13 - 18$ years of age. The severity of the disorder is described by the number of apnoea and/or hypopnea events per hour: the apnoea hypopnea index (AHI) or respiratory disturbance index (RDI).

Other than the consequences of disturbed sleep, the most recognised symptom for OSA in children is snoring [48, 60]; however, OSA can occur without the presence of snoring [56], and snoring can occur in patients that do not have OSA. Indeed, snoring occurs in approximately $10 - 20\%$ of children [60], where OSA is present in only approximately 2% [59]. Therefore, the presence of snoring cannot be used to identify the presence of OSA in children. OSA is characterised by blood de-oxygenation, a decrease in nasal pressure, increased respiratory effort and/or restlessness during sleep. Therefore, a subset of these signals could be used to identify the disorder. There are devices with varying degrees of effectiveness that use these biological signals to detect OSA in children. These are outlined below.

2.2 Diagnostic methods for obstructive sleep apnoea

Polysomnography, the recognised gold standard for diagnosing sleeping disorders [65], has limitations that motivate diagnostic methods that can be conducted in the home environment. These methods include both qualitative questionnaires and quantitative devices. Many questionnaires have been developed and used to identify sleep-related breathing disorders. Questionnaires are often subjective and require patient integrity to be effective. As such, these techniques are difficult to validate and standardise [60]. There are quantitative methods currently used

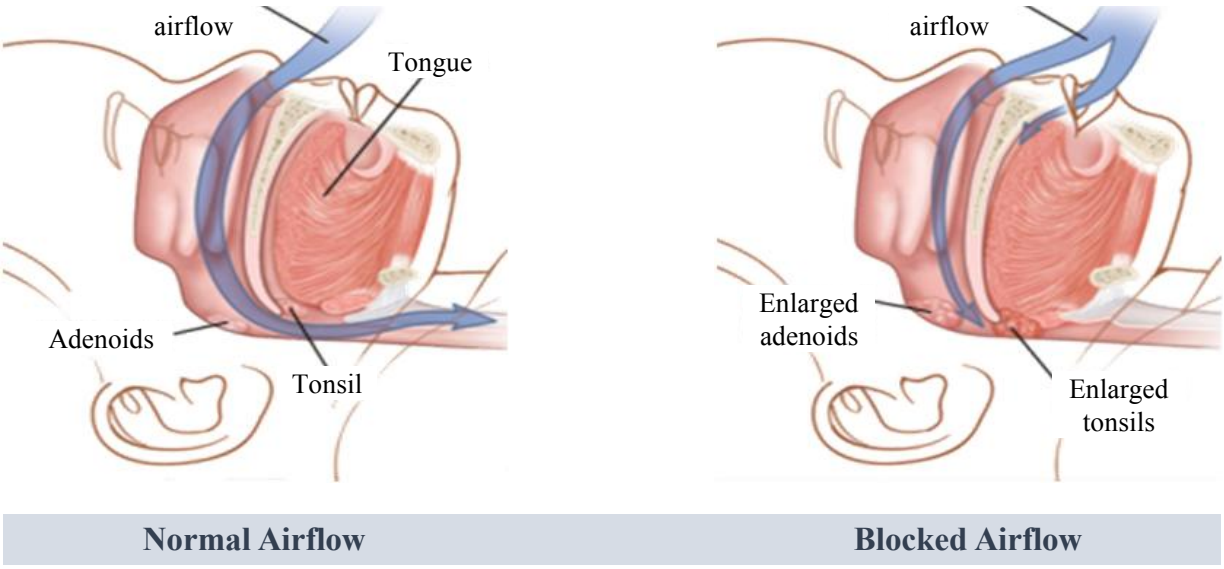


Figure 2.3: Normal airway anatomy prior to inflammation of the adenoids and/or tonsils (left), and post-inflammation (right). Image courtesy of [64].

Table 2.3: Apnoea and hypopnea scoring criteria for children [42]

Apnoea	
	Drop in peak flow signal excursion of $\geq$ 90% from pre-event baseline and meets any of the following duration and respiratory criteria.
Obstructive	*Additionally:* duration of at least 2 breaths and there is respiratory effort during the period of absent airflow.
Central	*Additionally:* absent inspiratory effort and one or more of; • duration of at least 20 seconds; or • at least 2 breaths with an arousal or $\geq$ 3% arterial desaturation; or • decrease in heart rate to $<$ 50 beats per minute for at least 5 seconds.
Mixed	*Additionally:* duration of at least 2 breaths and associated with absent respiratory effort in a portion of the event with the presence of inspiratory effort in another portion.
Hypopnea	
	Drop in peak flow signal by $\geq$ 30% of pre-event baseline, duration of drop lasts for $\geq$ 2 breaths, $\geq$ 3% oxygen desaturation from pre-event baseline of association with an arousal, and meets any of the following criteria.
Obstructive	*Additionally:* one or more of; • snoring; or • increased respiratory flattening of nasal pressure or flow signal compared to baseline breathing; or • occurrence of thoracoabdominal paradox during event but not during pre-event breathing.
Central	*Additionally:* none of the obstructive criteria.

to identify sleep in a non-laboratory setting, which will be discussed in Section 2.2.2. Since these devices are developed to be non-invasive and relatively simple to operate without clinical guidance, they only measure a small subset of relevant biological signals. Polysomnography and these home-based techniques will be discussed in this section.

2.2.1 Polysomnography

Polysomnography (PSG) monitors various physiological changes during sleep, such as brain activity (electroencephalography), heart rhythm (electrocardiography), airflow, blood oxygenation, eye movements (electrooculography) and muscle activity (electromyography) [71, 13, 30].

The polysomnography signals are analysed post-study and sleep stages (as defined by the paediatric criteria outlined in the *AASM Manual for Scoring Sleep and Associated Events* [38] and detailed in Section 2.1.1) are manually scored on a 30 second epoch basis. Polysomnography identifies OSA-related respiratory events by changes in chest and abdominal motion, airflow and arterial oxygen saturation [14]. Polysomnography also provides a summary of the child's sleep architecture. From this, the severity of sleep disturbance can be estimated.

There are some common criticisms for the effectiveness of polysomnography, particularly concerning the clinical significance of abnormalities, the impact of polysomnography on the child's natural sleep behaviour and the validity of performing full sleep studies on each child symptomatic of OSA [14]. These criticisms stem from the cost, availability and complexity of polysomnography. Fig. 2.5 illustrates the limitations of polysomnography in Australia, particularly with respect to geographical availability. The location of sleep laboratories (circles represent those in capital cities and triangles represent those in regional areas) is scattered throughout high-density areas (represented by the shading within each state), and the majority of sleep laboratories are located in capital cities. In Australia, this means that approximately 4% of the population (920,000 out of 23,000,000 people) are further than 1000 km from the nearest laboratory, and approximately 50% of the population (12,000,000 people) are further than 200 km from a sleep laboratory (illustrated by the population per m^2 shaded regions in Fig. 2.5). A diagnostic polysomnogram is also expensive, representing approximately 0.32% (approximately $AUD59.3 million) of the Australian government Medicare budget [68]. Polysomnography is considered effective for diagnosis, because it acquires very detailed measurements. However, this supposed advantage may be detrimental to the accuracy of the measured sleep architecture. That is, the numerous sensors required for gathering detailed measurements may cause the pa-

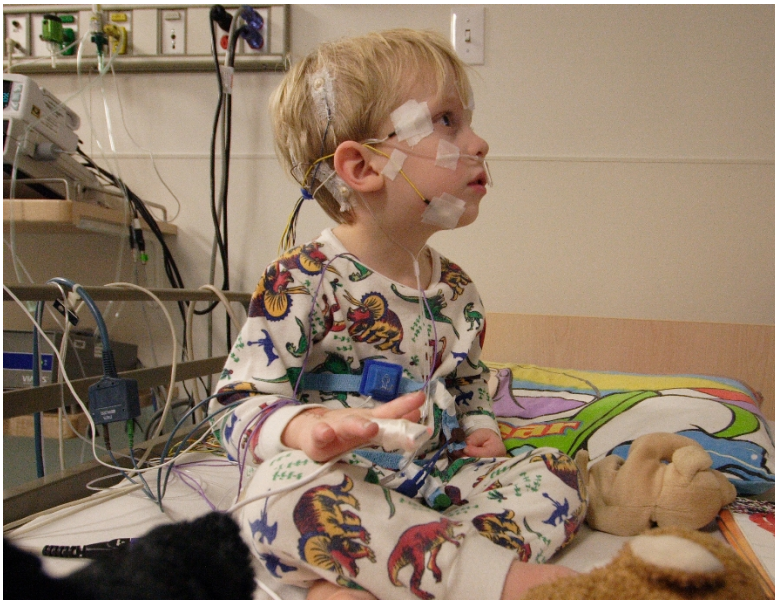


Figure 2.4: Example polysomnogram on child. Image courtesy of [66].

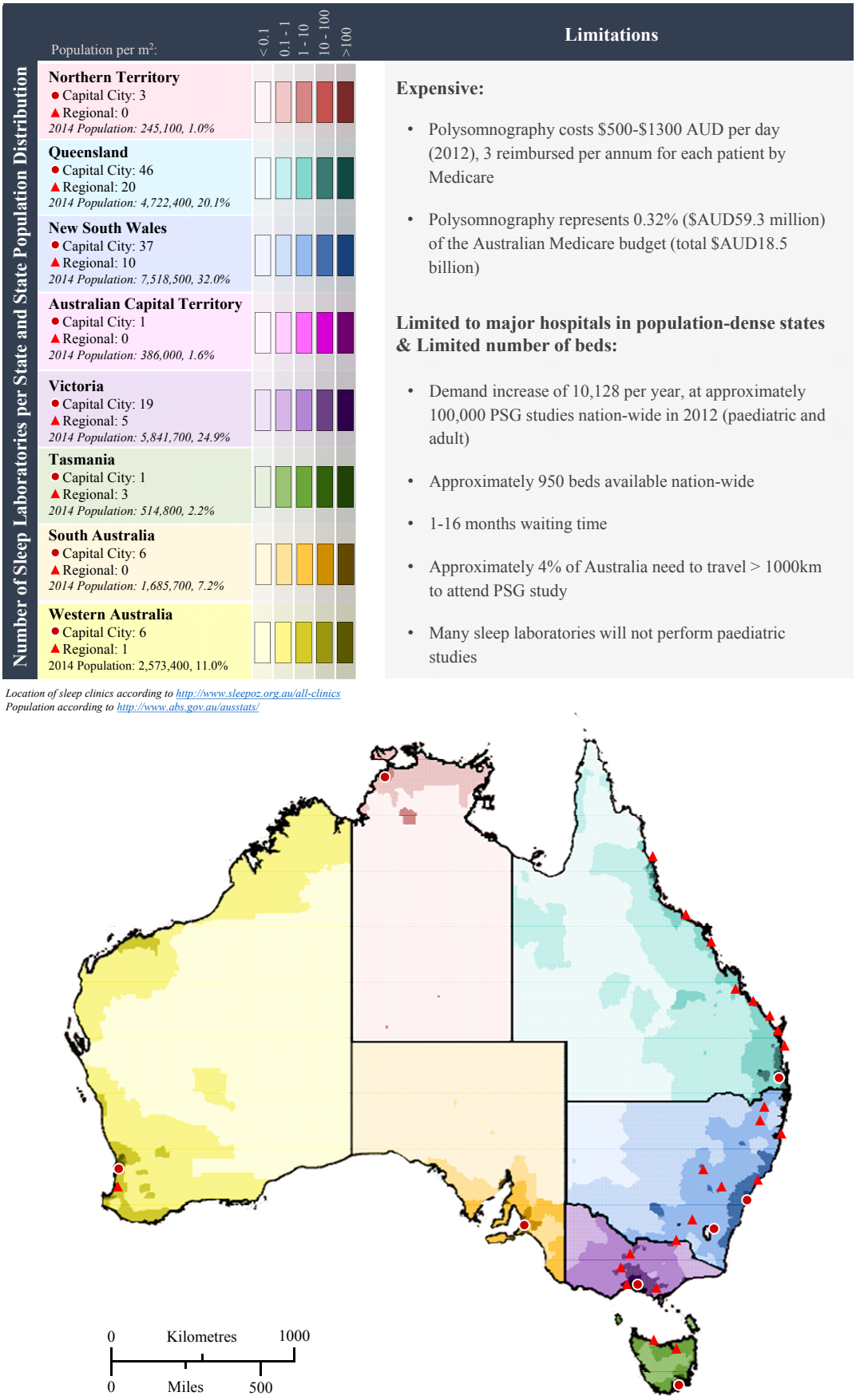


Figure 2.5: Limitations of full diagnostic sleep studies with polysomnography in Australia [67, 68, 69, 70].

tient to sleep poorly or differently to their normal behaviour. For this reason, the *estimated* sleep behaviour may not be indicative of the patients *actual* sleep behaviour.

In addition to the intrusive effect of the sensors, the laboratory environment may hinder sleep quality because it is unfamiliar and intimidating for children. Indeed, the accuracy of home polysomnography has been compared to laboratory polysomnography and the measures of sleep behaviour did differ [72, 73]. This difference may be due to patient discomfort and/or patient awareness of the study [74]. This is a well known trait of polysomnography, termed the 'first-night effect' [75], whereby the emotional and physical discomfort of polysomnography alters sleep behaviour [76]. The altered sleep behaviour, as well as the cost and limited availability of polysomnography, have motivated the development of diagnostic techniques that are less invasive and can be used in a non-clinical setting.

2.2.2 Home-based diagnostic techniques

Many relatively non-invasive devices have been developed to remotely measure sleep behaviour. The most commonly used for OSA are respiratory polygraphy, pulse oximetry and actigraphy.

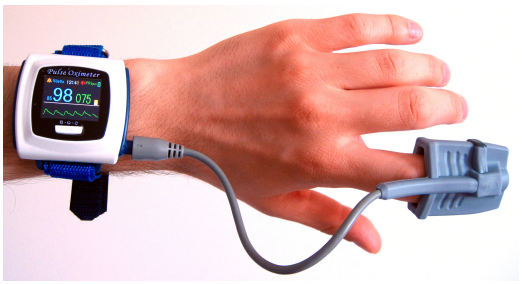
Respiratory Polygraphy

Respiratory polygraphy is a scaled-down polysomnogram that is developed for use in a non-laboratory setting. Respiratory polygraphy combines sensors that monitor cardiorespiratory signals. In addition to this, a camera is occasionally used for recording sleep behaviour [73]. The cardiorespiratory signals summarise pulse rate and waveform, oxygen saturation, heart rhythm and thoracic and abdominal excursions. Unlike in adults, monitoring facial signals (such as EEG or EOG) is avoided in children so as to minimise strangulation risk and reduce invasiveness, thereby mitigating any alteration of the child's natural sleeping behaviour. Although home polygraphy has been found to have high concordance with laboratory polysomnography [73], it suffers from complexity. This is a great limitation in the home environment where parents are likely to be responsible for configuring and/or monitoring the system. For this reason, it is important to consider techniques that are both simple to use and accurately estimate sleep architecture.

Pulse Oximetry

Pulse oximetry is one of the most common non-invasive quantitative methods used to measure blood oxygenation [79, 80, 81, 82] (illustrated in Fig. 2.6). The cessation of respiration caused by an OSA event can cause blood oxygen desaturation. Pulse oximetry attempts to identify these desaturations and thereby detect apnoeic events. Although pulse oximetry has been found an effective tool (validated positive predictive rate of 97% [80]), a major constraint is that it can only identify patients with more severe OSA. One reason for this is that arousals from sleep

(a)



(b)

Figure 2.6: Example pulse oximeter setup on child (a) and device (b). Images courtesy of [77, 78].

can occur without significant oxygen desaturation [6]. These arousals contribute to disturbed sleep, but are unidentifiable with pulse oximetry.

Actigraphy

Actigraphy conventionally estimates sleep and wake by measuring wrist activity. Identified activity during the night likely corresponds with periods of wake. Actigraphy uses this principle to identify the severity of sleep disturbance across the night, which in turn aids with identification of a particular sleep disorder [15]. Although actigraphy has good agreement with the polysomnography (approximately 90% [83, 84]), it has limitations that impact its effectiveness at identifying disorders where the patient moves while asleep, or remains still while awake. In particular, activity that occurs while the patient is asleep will often be misidentified as occurring during wake. Similarly, regions of no activity during wake will be misidentified as regions of sleep. Although there are currently technical limitations that affect the accuracy

Figure 2.7: Example actigraph: MiniMitter Philips Actiwatch 2. Image courtesy of [85].

of estimating sleep architecture with actigraphy, there are distinct advantages that make it an attractive choice for assessing sleep in a non-laboratory setting. Actigraphy is durable, relatively cost-effective, and easily configured without clinical guidance; wearing an actigraph is as invasive and complex as wearing a wrist watch (illustrated in Fig. 2.7).

Respiratory polygraphy, pulse oximetry and actigraphy monitor physiological signals that allow identification of events related to obstructive sleep apnoea. Respiratory polygraphy offers high concordance with polysomnography; however, it is still greatly invasive, costly and complex to configure and/or monitor. Although pulse oximetry offers a relatively non-invasive and effective method for identifying oxygen desaturation caused by a respiratory-related event (common in OSA), it is unable to identify arousals from sleep that are not respiratory-related. These arousals may be important for identifying the severity of sleep disturbance. Furthermore, pulse oximetry cannot estimate sleep architecture and thereby any measure of sleep efficiency or disturbance.

In contrast to pulse oximetry, actigraphy is able to noninvasively identify any event that causes movement, such as an arousal, apnoea, or disturbed sleep. Actigraphy is currently used to assess sleep quality; however, the technical limitations hinder sleep estimation accuracy. Considering that actigraphy appears an ideal candidate for detecting the severity of sleep disturbance, there is considerable motivation to address the specific technical limitations. For this reason, this thesis will focus on improving the efficacy of actigraphy for identifying sleep disturbance in children.

2.3 Technical overview of commercial actigraphy

Commercial sleep actigraphy systems generally consist of a uni-, bi- or tri-axial accelerometer that quantifies movement as a summarised 'activity count' per time period or epoch [83], commonly at 30 seconds. Actigraphy in sleep assessment typically classifies epochs with movement as wake and epochs with no movement as sleep. These movements are quantified using one of several common time-series techniques. Time-series techniques are favoured because they are easily interpretable and realisable with simple hardware. However, this simplicity comes with a trade-off against accuracy, which will be discussed in Section 2.3.1.

Despite achieving relatively high agreement with polysomnography, actigraphy currently suffers from poor sleep specificity; actigraphy often misidentifies movement during sleep as 'wake'; and regions of quiet wake as 'sleep'. This misidentification may be due to technical limitations of commercial systems, such as the sensor type and placement, the movement representation and/or the sleep/wake classification method. Actigraph placement and movement representations differ between commercial systems and within published literature. Unsurprisingly, there is much discussion concerning the choice of single accelerometer placement and activity count derivation technique. There have been many validation studies with actigraphy against sleep questionnaires, sleep diaries and polysomnography. Table 2.4 summarises the concordance of actigraphy with the various validation devices for paediatric sleep assessment. Where possible, the sensitivity and specificity have been noted, otherwise the agreement or concordance with the validation metric is noted.

Table 2.4: Actigraphy validation history for children[a]

Year	Author/s	Subjects	Actigraph	AC Method[b]	Location[c]	Validation Method[d]	Reported Performance[e]
1989	Sadeh et al. [86]	13, 3-13 years	unspecified, Ambulatory Monitoring	ZC	Wrist	PSG	ρ: 0.813
1991	Sadeh et al. [87]	11, 12-48 mo, 4M	unspecified, Ambulatory Monitoring	ZC	Left leg	PSG	Se: 87.7% Sp: 76.9% Ag: 85.3%
1994	Sadeh [88]	50, 9-24 mo, 28M	unspecified, Ambulatory Monitoring	n.d.	n.d.	SD	TST: 585(60) mins vs. 598(49) mins SE: 81.2(6.2)% vs. 92.4(6.8)%
1995	Sadeh et al. [89]	41, 3-12 mo	AMA-32, Ambulatory Monitoring	ZC	Left ankle	CBO	Ag: 95.6%
1996	Sadeh [90]	66, 7-26 mo, 46M	AMA-16, Ambulatory Monitoring	n.d.	Leg	SD	TST: ρ 0.74 SE: ρ 0.41
2002	Gnidovec et al. [91]	10, 1-6 mo, 6M	Z80, Gaehwiler Electronics	n.d.	Left leg	CBO	Ag: 87-95%
2004	Sazonov et al. [92]	26, infants (CHIME)	Custom, uni-axial	max, var, ZC	Waist	PSG	Ag: 77-92%
2005	Acebo et al. [93]	169, 1-5 years, 84M	AMA-32 Mini, Ambulatory Monitoring	ZC	Left ankle (<36mo) ND Wrist (> 36mo)	SD	TST: 582(36) mins vs. 624(36) mins
2005	So et al. [94]	22, <6 mo, 4M	Actiwatch AW64, Mini-Mitter Company Inc.	n.d.	Mid-right Calf	PSG	Se: 87.5% Sp: 63.5% Ag: 86.5%
2007	Hyde et al. [95]	45, 1-12 years, 29M	Actiwatch AW64, Mini-Mitter Company Inc.	n.d.	ND Wrist	PSG	Se: 93.9% Sp: 59.0% Ag: 87.3%

Table 2.4: Actigraphy validation history for children[a] (continued)

Year	Author/s	Subjects	Actigraph	AC Method[b]	Location[c]	Validation Method[d]	Reported Performance[e]
2007	Johnson et al. [96]	181, 12-16 years, 910M	Octagonal Sleep Watch 2.01, Ambulatory Monitoring Inc.	ZC, TAT, DI	Wrist	PSG	TST ICC — ZC: 0.32 TAT: 0.41 DI: 0.34
2008	Werner et al. [97]	50, 4-7 years, 28M	Actiwatch + AW4, Cambridge Neurotechnology	n.d.	ND Wrist	SD	TST: 505(22) mins vs. 614(18) mins
2008	Sitnick et al. [98]	58, 4-6 years, 29M	Actiwatch AW64, Mini Mitter Company Inc.	n.d.	ND Ankle	VSG	Se: 97.6% Sp: 24.3% Ag: 94.6%
2009	Sung et al. [99]	10, 6-9 mo, 8M	Actiwatch AW64, Mini Mitter Company Inc.	n.d.	Mid-right Calf	CBO	Se: 88.7% Sp: 55.0% Ag: 85.1%
2010	O'Driscoll et al. [100]	130, 2-18 years, 85M	Actiwatch AW64, Mini Mitter Company Inc.	n.d.	ND Wrist	PSG	Se: 82.2% Sp: 50.9% Ag: 66.9%
2010	Holley et al. [101]	91, 6-12 years, 44M	Basic Mini Motionlogger, Ambulatory Monitoring Inc.	ZC	ND Wrist	CSHQ	n.d.
2010	Insana et al. [102]	22, 13-15 mo, 12M	Actiwatch AW64, Mini Mitter Company Inc.	n.d.	Ankle	PSG	Se: 91.2% Sp: 58.9% Ag: 89.6%
2010	Weiss et al. [103]	30, 16-18 years, 19M	Sleepwatch (SV Ambulatory Monitor Actiwatch (AW), Re: Actical (AC), Resp	n.d.	ND Wrist	PSG	ρ for TST: SW: 0.822 AW: 0.836 AC: 0.722
2011	Gregory et al. [104]	122, 7-17 years, 48M	Octagonal Basic Motionlogger Actigraph, Ambulatory Monitoring	n.d.	ND Wrist	PSG	TWT: ρ -0.18 SE: ρ -0.11
2011	Spruyt et al. [105]	149, 4.1-8.8 years, 62M	Actiwatch AW64, Mini Mitter Company Inc.	n.d.	ND Wrist	PSG	TST: 460(38) mins vs. 492(41) mins WASO: 50(27) mins vs. 24(25) mins
2011	Dayyat et al. [106]	327, 3-10 years, 169M	Actiwatch AW64, Mini Mitter Company Inc.	n.d.	ND Wrist	SD	SOL: 21.8(0.8) mins vs. 23.7(4.5) mins
2012	Short et al. [107]	385, 13-18 years, 231M	MicroMotionlogger, Ambulatory Monitoring	ZC	ND Wrist	SD	TST: 468(28) mins vs. 508(30) mins WASO: 26(20) mins vs. 11(8) mins
2012	Ward et al. [108]	71, 9-11 years, 29M	Actiwatch AW64, Mini Mitter Company Inc.	DI	ND Wrist	PSG	Se: 90% Sp: 77% Ag: 87%

Table 2.4: Actigraphy validation history for children[a] (continued)

Year	Author/s	Subjects	Actigraph	AC Method[b]	Location[c]	Validation Method[d]	Reported Performance[e]
2012	Meltzer et al. [84]	115, 3-18 years, 56M	Motionlogger (ML), Actiwatch-2 (AW Mini Mitter	ML: ZC AW: n.d.	ND wrist	PSG	ML — Se: 89% Sp: 73% Ag: 87% AW — Se: 93% Sp: 69% Ag: 89%
2013	Boyne et al. [109]	10, 19-17 years	Actiwatch 16/64, Respironics	ZC	ND Wrist	PSG	SOL: ρ 0.911 TST: ρ 0.536

[a] *n.d.* signifies that the associated data was not disclosed.

[b] *ZC, DI* and *TAT* refer to the zero crossing, digital integration and time above threshold representation methods.

[c] *ND* refers to the non-dominant limb position.

[d] *PSG* refers to polysomnography, *SD* refers to sleep diary, *CBO* refers to clinical behavioural observations and *CSHQ* refers to Children's Sleep Habits Questionnaire [110].

[e] Values shown as actigraphy vs. validation method. ρ refers to the correlation coefficient, *Se, Sp* and *Ag* refer to sensitivity, specificity and agreement respectively, *TST* refers to total sleep time, *TWT* refers to total time awake, *SOL* refers to sleep onset latency, *SE* refers to sleep efficiency, and *ICC* refers to the inter-correlation coefficient.

2.3.1 Detecting wake in the presence of movement

Commercial actigraphy systems identify wake by first isolating large movements that are assumed to occur solely during wake. They then often apply a smoothing procedure, or a re-scoring algorithm, that remove short periods of quiet wake. The techniques used to quantify movement and then re-score wake are described below.

Quantifying movement

Accelerometers measure both dynamic activity and static positions (i.e. positions with respect to gravity) [111]. Until recently, commercial systems have used uni-axial accelerometers to monitor motion. Although simple, uni-axial accelerometers only measure motion along one axis of movement. The consequence of this is that movement that occurs in other axes will be attenuated or not detected at all, as illustrated by three example vectors of movement in Fig. 2.8: illustrating the ideal case, $\boldsymbol{c}$ is completely along the axis of measurement u, which means that all of the magnitude is measured; $\boldsymbol{a}$ is at a 0.873 rad, or 50 deg, angle to the primary axis of measurement u, which means that only cos(0.873), or 64.3%, of the magnitude of $\boldsymbol{a}$ is measured along axis u; finally, in the least ideal case, $\boldsymbol{b}$ is at a $\frac{\pi}{2}$ rad, or 90 deg, angle to u, which means that none of the magnitude of $\boldsymbol{b}$ is measured along the axis u. For this reason, commercial systems employing tri-axial accelerometry have recently been released.

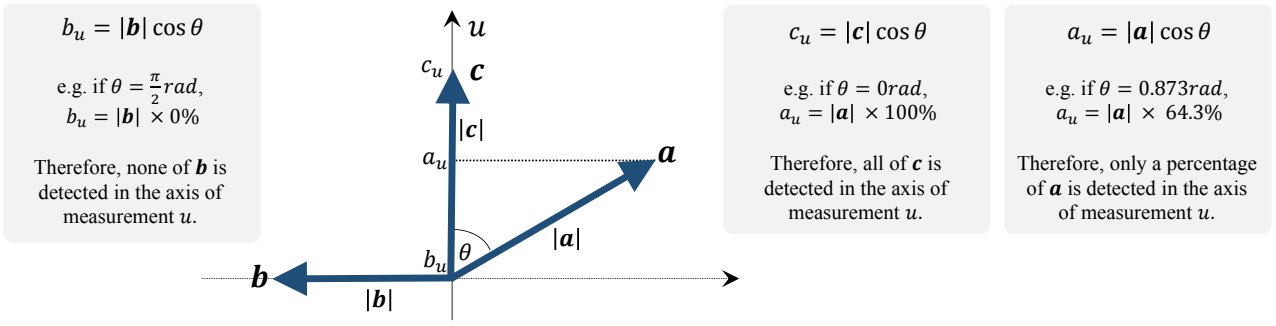


Figure 2.8: Limitations of uni-axial accelerometry for detecting movement.

The standard approach for representing movements for both uni- and tri-axial systems is a time-series summary within a fixed time frame, or epoch (typically 30s) [83]. The most common time-series methods for summarising movement are time above threshold (TAT), digital integration (DI) and zero crossing (ZC). TAT represents movement by summarising the time a signal $\mathbf{a}$ is above a defined threshold T,

$$TAT = \sum_{n=1}^{N \cdot f_s} \mathbb{1}\{\|\mathbf{a}[n]\| > T\}, \tag{2.1}$$

where N represents the epoch length (typically 30s), f_s represents the sampling rate and $\mathbb{1}$ represents the indicator function. DI represents movement by summarising the amount of movement within the time frame,

$$DI = \sum_{n=1}^{N \cdot f_s} \|\mathbf{a}[n]\|. \tag{2.2}$$

ZC represents movement by counting the number of times the movement signal crosses a defined threshold within the time frame,

$$ZC = \sum_{n=1}^{N \cdot f_s} \mathbb{1}\{(\|\mathbf{a}[n]\| - T) \cdot (\|\mathbf{a}[n-1]\| - T) < 0\}. \tag{2.3}$$

As shown in Fig. 2.9, TAT and DI are effective summaries of total movement. However, they are unable to effectively represent high-energy movements of short duration (e.g. limb twitches). TAT is unable to represent oscillatory movements because they will register above the threshold for a short time, regardless of the magnitude, and small movements will likely remain below the threshold and be missed altogether. DI is unable to accurately represent oscillatory movement and short-duration movements because the signal often returns to baseline, restricting the area under the signal. Converse to TAT and DI, ZC is effective at representing oscillatory motion, but is unable to effectively represent high-energy movements that are long in duration and do not oscillate about the threshold. Since ZC counts the number of times the signal crosses the

threshold, movements that do not oscillate will cross the threshold a small number of times, despite being significant movements. These movements will give small summary values, which can be misidentified as 'no movement'. In addition to these limitations, TAT and DI do not describe the spectral nature of any movement because they are solely time-series approaches. ZC only provides a superficial indication of the spectral content of movement because it describes the frequency with which movement crosses a threshold.

Actigraphs, or raw accelerometry, are often used in literature and clinical settings to quantify movement for gait analysis, physical activity [21], seizure detection [112], and fall detection in the elderly [113, 114]. These applications often use spectral representations within each epoch

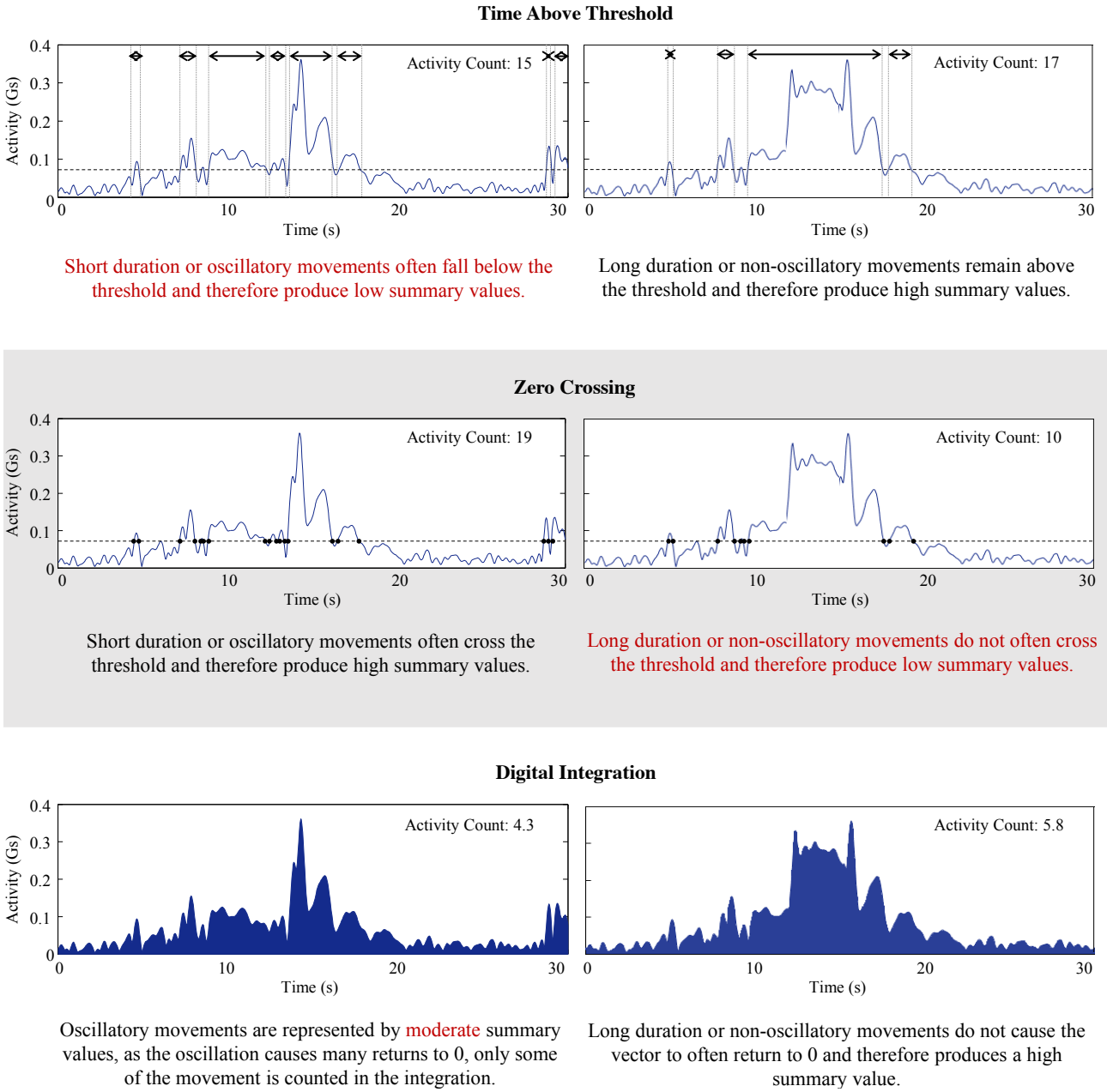


Figure 2.9: Limitations of the conventional time-series representations for detecting movement.

to identify these events [115, 116]. Such windowed spectral representations can be categorised as a time-frequency representation. Following from this, a more suitable approach is likely wavelet analysis because this indicates the prevalent frequency components for *varying* window sizes. Indeed, wavelet analysis has successfully been applied to identify single occurrences of different activities in physical activity [115, 117, 118, 119].

Wavelet analysis identifies localised temporal instances where there is a change in frequency content; for example, the temporal instance when a positional change occurs within an epoch. An advantage of wavelets over other time-frequency representations is that wavelets allow variable frequency and temporal resolution (illustrated in Fig. 2.10), whereas typical spectral representations only allow variable frequency resolution *or* time resolution (represented by the Short-Time Fourier Transform in Fig. 2.10), but not both (commonly known as the 'Heisenberg-Gabor limit' or the 'Gabor limit' [120]). Temporal resolution is useful in this application because the frequency content of movement is likely to change over time. This technique also allows temporal isolation of particular movement types or postural changes. This resolution may aid in improving the poor specificity of actigraphy for sleep assessment by providing an avenue for effectively discriminating sleep/wake movements.

Movement during sleep appears to be more sporadic and less periodic than movement during physical activity (e.g. walking or stair climbing). As such, the windowed nature of wavelet analysis seems particularly suited to analysing movements during sleep. Movements during sleep are also likely to differ in duration, which further supports the viability of the varied window lengths in the wavelet transform. While this approach seems particularly suited to this application, it has not previously been applied to actigraphy in sleep assessment.

Sleep scoring algorithms

One limitation of actigraphy in sleep assessment is the inability to accurately detect when a patient is immobile but awake [15]. Several algorithms have been developed for different commercial actigraphy systems in an attempt to improve this limitation. These algorithms apply a weighted function over the movement representations and score sleep or wake if the algorithm produces a result above or below a certain threshold (detailed in Table 2.5). These algorithms effectively smooth the movement and thereby reduce the rate at which quiet wake is scored as 'sleep'. Various algorithms have been previously developed using regression and/or optimisation techniques (summarised in Table 2.5). Smoothing movement during sleep reduces 'peaky' movement, making it likely to fall below the wake threshold. Similarly, smoothing movement during wake spreads the distribution of movement across epochs, making it more likely that short periods of quiet wake between movements will correctly be scored as wake. These effects somewhat improve the specificity, albeit only by a small amount.

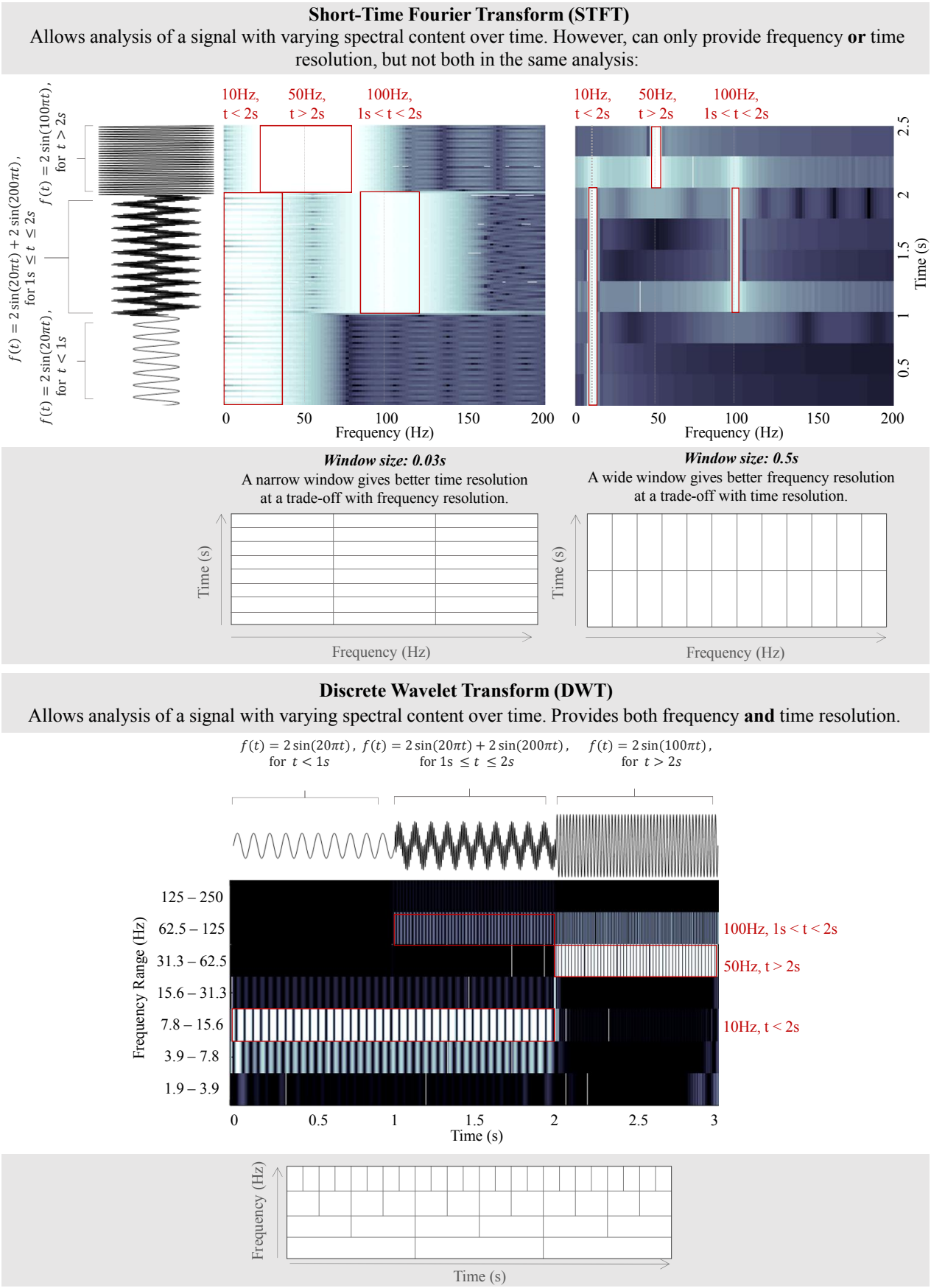


Figure 2.10: Spectral representations (such as the Short-Time Fourier Transform (STFT)) only allow either good frequency-resolution or good time-resolution because they are restricted to a fixed window size. The wavelet transform allows both good frequency- and time-resolution by varying the window size for each scale of the transform.

Table 2.5: Smoothing algorithms developed for commercial actigraphy

Year	Author/s	Actigraph	Algorithm	Reported Performance[a]
1982	Webster et al. [121]	Medilog, Ambulatory Monitoring Inc.	$P_{Wake} = 0.036(0.07A[-5] + 0.08A[-4] + \ldots$ $0.10A[-3] + 0.11A[-2] + 0.12A[-1] + 0.14A[0] + \ldots$ $0.09A[1] + 0.09A[2] + 0.09A[3] + 0.10A[4]) \geq 1,$ where $A[n]$ represents the maximum 2 second epoch value within minute n for dominant wrist movement.	Ag: 94.46%
1989	Sadeh et al. [86]	Unspecified model, Ambulatory Monitoring Inc.	$P_{Sleep} = 4.532 - 0.06828XO - 0.0385\sigma[-5] - \ldots$ $0.0299\sigma[-2] - 0.038\sigma[9] + 0.0298\mu[2] \geq 0,$ where XO represents the number of zero crossings of the current 1 minute epoch for non-dominant wrist movement, and $\mu[n]$ and $\sigma[n]$ represent the mean and standard deviation of epoch n relative to the current epoch.	Nominal — Sp: 63.54% - 76.18% Insomnia or OSA[b] — Sp: 48.48%, 56.47%
1992	Cole et al. [122]	Motionlogger, Ambulatory Monitoring Inc.	$P_{Wake} = 0.00001(404A[-4]+598A[-3]+326A[-2]+$ $441A[-1] + 1408A[0] + 508A[1] + 350A[2]) \geq 1.$ where $A[n]$ represents the average $2s$ zero-crossing activity for minute n for wrist movement.	Ag: 87.05%
1994	Sadeh et al. [89]	AMA-32, Ambulatory Monitoring Inc.	$P_{Sleep} = 7.601 - 0.065Mean_{W5min} - \ldots$ $1.08NAT - 0.056\sigma_{last6min} - 0.703LOG_{Act} > 0.$ where NAT represents the number of epochs with (zero-crossing) activity, A, $50 \geq A \leq 100$ within an 11 minute window ($\pm$5 minutes) , $W5min$ is the average number of activity counts during a 5 minute window surrounding the current epoch, and ACT is the number of activity counts during the scored (and the next) epoch. Epochs were scored in 1 minute intervals.	Ag: 93% - 98%

Table 2.5: Smoothing algorithms developed for commercial actigraphy[a] (continued)

Year	Author/s	Actigraph	Algorithm	Reported Performance[a]
1997	Oakley [123]	Actiwatch, MiniMitter Co., Inc.	$P_{Wake} = 0.04A[-4]+0.04A[-3]+0.20A[-1]+2A[0]+$ $0.20A[1] + 0.20A[2] + 0.04A[3] + 0.04A[4] > T.$ where $A[n]$ represents $30s$ zero-crossing activity for epoch n for non-dominant wrist movement.	Sp: 28% - 48%
2001	Kushida et al. [124]	AW4, MiniMitter Co., Inc.		Varying disorders — Ag: 73% - 84% Se: 92% - 99% Sp: 24% - 63%

[a] *Se, Sp* and *Ag* refer to sensitivity, specificity and agreement respectively.
[b] OSA refers to obstructive sleep apnoea syndrome.

2.3.2 Detecting wake in the absence of movement

The smoothing algorithms in the previous section help detect periods of restful wake that occur between periods of movement. Another approach to this problem is a heuristic developed by Webster et al. [121] in 1982, which assumes that quiet wake will always follow periods of active wake. That is, it dilates identified wake periods in an attempt to correctly identify quiet wake. The following heuristic is applied to the predicted sleep stages:

Data: Epoch scoring from actigraph
Result: Re-scored epochs as sleep or wake
repeat
  if *previous 4 minutes scored as wake* **then**
    re-score next minute as wake;
  else if *previous 10 minutes scored as wake* **then**
    re-score next 3 minutes as wake;
  else if *previous 15 minutes scored as wake* **then**
    re-score next 4 minutes as wake;
  if *6 minutes of sleep is surrounded by 10 minutes of wake* **then**
    re-score 6 minutes of sleep as wake;
  else if *10 minutes of sleep is surrounded by 20 minutes of wake* **then**
    re-score 10 minutes of sleep as wake;
until *all epochs are checked*;
repeat one consecutive epoch loop **until** *stop*;

This heuristic has been used in literature and gave an approximate 1% performance improvement [121, 122]. This improvement, as well as the smoothing algorithms in the previous section, suggests that there is often actual wake within a short time after detected wake periods. Smoothing the regions surrounding wake does indeed improve specificity. Although these approaches successfully identify some regions of quiet wake, they are not without limitations. Each of these methods make the assumption that there will always be a set time before and/or after wake where the patient will *always* remain awake. This assumption will sometimes incorrectly identify regions of actual sleep as wake, thereby contributing to the poor specificity. There are also short periods of quiet wake amongst sleep epochs (e.g. sleep fragmentation) that this heuristic will not identify.

2.4 Problem statement

Obstructive sleep apnoea syndrome (OSA) in children is not as well studied as in adults, even though, as discussed in Section 2.1.3, it is greatly detrimental to the child's general health, behavioural functioning and cognitive ability. OSA in children often goes undetected due to a combination of parental oversight and limited availability of diagnostic methods, specifically detailed in Fig. 2.5. In addition to this, when the child is able to attend a diagnostic sleep study with polysomnography — the current gold standard for diagnosing OSA — their measured sleep behaviour may be inaccurate. As discussed in Section 2.2.1, these inaccuracies are often caused by the large number of sensors required for polysomnography and the unfamiliar laboratory environment. These inaccuracies can be mitigated by performing multiple studies over consecutive nights; however, multiple studies increase both cost and inconvenience. Therefore, there is motivation for developing mobile diagnostic tools that are non-invasive, cost-effective and can easily be configured and monitored by the child's parents in the home environment.

Actigraphy, introduced in Section 2.3, is inexpensive, non-invasive, and easily configurable. Actigraphy provides an estimate of sleep quality and a potential avenue for identifying sleeping disorders (either on its own, or in combination with other tools such as pulse oximetry or respiratory polygraphy). However, actigraphy is not without limitation: as discussed in Section 2.3.1, actigraphy cannot identify when a patient is immobile, but awake, or mobile, but asleep. For this reason, actigraphy often suffers from poor specificity (highlighted in the validation studies in Table 2.4), which limits the ability to accurately assess sleep, particularly for patients with a sleep disorder. In this thesis, it is proposed that the misclassification of sleep stages may be caused, in part, by technical restrictions of early actigraphy models. Actigraphy systems were first developed for sleep assessment in the 1980s. The majority of developmental increments and technical exploration studies were consequently performed within the decade or two following this. However, technology has advanced since these studies, particularly with the introduction of micro electro-mechanical systems (MEMS) technology and the increase of

general processing power. Although the accelerometer configuration, conventional time-series representations and simplistic classification rules were required when processing power was limited, they may be somewhat responsible for the poor specificity of current models.

Actigraphy is conventionally placed on the wrist when assessing sleep quality. Restricting movement detection to a single placement on the body prevents detection of movement from other areas of the body. Improving the ability to detect movements may allow actigraphy to identify regions of wake that would not be detected by wrist movement alone. Furthermore, the incidence of certain movements may differ between sleep and wake. Identifying characteristics that are exclusive to movements during sleep, or wake, would allow actigraphy to correctly identify regions of restless sleep as 'sleep', thereby improving specificity. Despite this, multisite accelerometry has not previously been explored in sleep assessment. Another limitation of historical commercial actigraphs for assessing sleep is the use of uni-axial accelerometers. Uni-axial accelerometers are unable to detect movement that occurs orthogonal to the measurement axis, as illustrated in Fig. 2.8. Although recent commercial systems have employed tri-axial accelerometry, there have been no direct comparisons of the tri-axial movement representation techniques initially developed for uni-axial devices. Therefore, it is unclear if the advantages of tri-axial accelerometry are fully exploited.

The conventional time-series representations provide a summary of movement within a fixed period of time. As discussed in Section 2.3.1, the limitations of these summaries vary between the representations. However, each conventional representation cannot summarise movements with particular characteristics, as illustrated in Fig. 2.9; for example, zero crossing (ZC) cannot provide a suitable representation of non-oscillatory movement, such as postural changes, because these movements will not oscillate about the threshold. These limitations restrict detection of significant movements, which then cannot contribute to wake detection. The conventional time-series representations can only provide a vague summary of the spectral characteristics of movement. However, these characteristics are likely important because the spectral content is unique for different movements. Finally, there may be specific localised spectral characteristics of movements that differ between sleep and wake. These differing characteristics would enable correct identification of sleep state, which would further improve specificity. Despite this, as noted in Section 2.3.1, there has been no attempt to apply spectral techniques to movement information for assessing sleep.

In addition to the conventional movement representations, the procedure for identifying wake has distinct limitations. Each commercial system identifies wake using a simple threshold, i.e. activity above a pre-defined threshold is identified as wake. However, as noted above, this rule incorrectly identifies vigorous movement during periods of sleep as wake, and regions of no movement during periods of wake as sleep. Existing systems attempt to address this by applying a smoothing filter, outlined in Section 2.3.1, or a re-scoring algorithm, outlined in Section 2.3.2, to the identified sleep stages. As highlighted in Figure 1.2, these approaches are

only partially successful because they are generally effective only for extreme cases. Further to this, although these approaches successfully identify regions of quiet wake near identified regions of active wake, they incorrectly assume that wake will *always* be preceded and/or followed by a fixed period of quiet wake. For these reasons, as argued in Section 2.3.2, these approaches can worsen the specificity.

In this thesis, we will be identifying improvements to the utility of actigraphy in paediatric sleep assessment by addressing three hypotheses:

I Uni-axial accelerometry measured solely at the wrist limits sleep and wake prediction accuracy because movements orthogonal to the measurement axis, or occurring elsewhere on the body, cannot be detected.

Exploring multisite tri-axial accelerometry may consequently improve sleep and wake prediction by more accurately representing movement: tri-axial accelerometry will capture more movement information, and multisite accelerometry will capture more movements in general. Incorporating these hardware modifications may detect differentiable movements between sleep and wake, which would reduce false wake detections.

II Movement characteristics can differentiate sleep from wake because the physiological nature of these movements differ.

Identifying physiological characteristics that differ between sleep and wake movements may improve sleep and wake predictions by producing feature distributions that do not overlap, at least as much as the conventional techniques. Exploring localised spectral characteristics of segmented movements may identify differentiable characteristics, which would reduce false wake detections.

Predicting sleep and wake on a movement-by-movement basis may improve sleep and wake predictions by assessing movement-specific information, rather than a summary of activity within a fixed period of time. Estimating sleep on a movement-basis would allow incorporation of specific differentiable movement characteristics into the scoring routine.

III Physiological and pathological events characteristic of sleep disorders (e.g. apnoeas, hypopneas and transient arousals) cause sleep movements that contribute to false wake detections.

Exploring the association between pathological and physiological events characteristic of sleep disorders may consequently explain the presence of some sleep movements. This association may provide an avenue for capturing the severity of 'sleep disturbance' associated with a sleep disorder using actigraphy, rather than an estimate of sleep and wake regions.

It is therefore proposed that actigraphy in sleep assessment may be improved by:

- detecting movement with tri-axial accelerometry,
- incorporating accelerometry from different locations on the body,
- movement representations that adequately summarise temporal and spectral characteristics, and/or
- sleep and wake classification on a movement-by-movement basis.

We will investigate the performance benefits of implementing each of these suggestions on a custom accelerometry system (detailed in Section 3.1.3). The performance will be compared to a commercial system and the conventional epoch-by-epoch time-series representations.

We will improve the utility of actigraphy in sleep assessment by:

1. Reducing false sleep detections.

False sleep detections may be caused by some wake movements going undetected with the conventional actigraphy framework. We will explore techniques to increase the number of detected wake movements (Chapter 4). We will explore:

- Tri-axial movement representations; and
- Incorporation of movements from multiple locations on the body into the scoring routine.

2. Reducing false wake detections.

False wake detections may be caused by the inability of the conventional framework to differentiate sleep and wake movements. We will identify differentiable physiological characteristics between sleep and wake movements (Chapter 5). In particular, we will:

- Identify movements specific to restless sleep and explore the effect of removing these movements from the raw accelerometry data as a pre-processing step to activity count derivations;
- Identify differentiable localised spectral characteristics of sleep/wake movements; and
- Predict sleep and wake on a movement-by-movement basis.

3. Identifying whether sleep movements are indicative of 'sleep disturbance'.

We will explore potential physiological processes associated with sleep disorders that may cause sleep movements (Chapter 6). In particular, we will explore:

- The temporal association between movement and transient arousal; and
- The temporal association between movement and apnoeic events.

3

Experimental methodology and materials

This chapter outlines the experimental methodology and clinical study design that is common to each experiment in this thesis. The general data acquisition and processing framework that were applied consistently throughout the analyses will be described. Finally, the general clinical study design and technical procedures for all experiments in this thesis will be outlined.

3.1 Sleep study cohort, recordings and clinical procedure

The study was approved by the Mater health services HREC approval number ref. 1498C.

3.1.1 Participant recruitment and inclusion/exclusion criteria

Patients were recruited from the Mater Children's Hospital in Brisbane, Australia. Children undertaking full diagnostic polysomnography for a sleep-related breathing disorder were invited to participate. Patients were recruited if they provided verbal assent, and their parental guardian provided written consent.

Inclusion Criteria:
$5 - 16$ years of age;
Symptomatic of sleep-disordered breathing; and
Otherwise healthy.

Exclusion Criteria:
Experiencing any neuromuscular disorder or weakness; and
Presenting with a co-morbid disorder that affects craniofacial structure or breathing control.

3.1.2 Patient characteristics

In total, 38 patients were recruited for analysis. Table 3.1.3 details the gender, age, body mass characteristics, apnoea hypopnea index (AHI) and summary sleep characteristics for each patient. Also shown is whether data is available for Continuous Multisite Accelerometry System (CMAS) and/or the commercial actigraph, Actiwatch, for that patient. Patients shown in darker shading are those that the custom accelerometry system, CMAS, stopped recording or recorded extended regions, or an entire study, of a constant value (indicating sensor malfunction). Of the other patients, one had moderate (patient 5, AHI 5.1) and one had severe (patient 21, AHI 16.9) OSA. All other patients had less than moderate AHI.

3.1.3 Accelerometry devices used in analysis

Two accelerometry systems were used in the analyses to record movements during sleep: a commercial system (Actiwatch Mini,CamNTech for patients 1 to 30 and Actiwatch 2, Philips for patients 30 to 38) and a custom Continuous Multisite Accelerometry System (CMAS). The Actiwatch Mini records uni-axial activity counts in 2s epochs using the ZC method. Fig. 3.1 shows an example of the recorded data from the Actiwatch Mini. After technical limitations and device discontinuation, the Actiwatch 2 replaced the Actiwatch Mini. However, data from the Actiwatch 2 were not used in the analyses because only 8 patients wore the Actiwatch 2, which limits the statistical power of the analysis [125].

The custom accelerometry system was designed and developed at the University of Queensland by the research group prior to the commencement of this thesis. CMAS records raw tri-axial 8−bit accelerometry data (range ±2G) at 100Hz from five locations on the body (modules shown in Fig. 3.2):

- Left wrist with an auxiliary sensor on the left fingertip;
- Upper thorax; and
- Left ankle with an auxiliary sensor on the left great toe.

CMAS records temporal data, a saw-tooth timing signal and the $x-$, $y-$ and $z-$axis accelerometry data in a text file on a personal computer (PC), separate to the polysomnography recording. The CMAS modules wirelessly send a packet of data to the receiver unit at 10Hz (i.e. every 100ms) with each packet containing 10 accelerometry samples, sampled every 10ms. Temporal data was identified using the local time on the computer connected to the CMAS

Table 3.1: Full patient characteristics[a]

#	Gender (M/F)	Age (years)	Height (m)	Weight (kg)	BMI[b]	Weight Category[b]	AHI	SE (%)	TST (mins)	REM (%)	CMAS	Actiwatch[c]
1	M	13	1.53	49.45	21.12	Normal	7.7	79	414.5	27	No	No
2	M	13	-	-	-	-	0.5	71	364.5	22	Yes	Yes - Mini
3	M	9	1.31	27.60	16.08	Normal	1.1	82	429.8	12	Yes	Yes - Mini
4	M	12	1.49	41.40	18.65	Normal	1.5	57	315.5	8	Yes	Yes - Mini
5	F	7	1.30	29.85	17.66	Normal	0.7	95	534.5	26	Yes	Yes - Mini
6	M	11	1.41	42.60	21.43	Overweight	5.1	70	396.5	19	Yes	Yes - Mini
7	M	9	1.56	50.25	20.65	Overweight	1.3	80	457.4	21	Yes	No
8	F	8	1.32	27.30	15.67	Normal	0.4	75	426.5	19	Yes	No
9	M	16	1.78	129.00	40.72	Obese	1.4	80	335.5	18	Yes	No
10	M	16	1.70	55.30	19.14	Normal	0.0	73	362.5	12	Yes	No
11	M	6	1.16	19.85	14.75	Normal	0.7	71	417.0	16	Yes	Yes - Mini
12	M	7	1.28	41.80	25.51	Obese	2.7	91	547.7	21	Yes	Yes - Mini
13	M	9	1.39	31.75	16.43	Normal	0.0	87	460.0	27	No	Yes - Mini
14	F	5	1.16	21.96	16.32	Normal	4.4	91	531.2	35	No	Yes - Mini
15	F	6	1.11	22.85	18.55	Overweight	3.4	96	512.4	26	Yes	Yes - Mini
16	M	6	1.20	20.80	14.44	Normal	0.4	83	502.7	26	No	Yes - Mini
17	M	15	1.73	109.00	36.42	Obese	0.4	87	418.0	16	No	Yes - Mini
18	M	6	1.16	21.20	15.76	Normal	0.5	97	552.2	23	No	Yes - Mini
19	M	15	1.72	58.85	19.89	Normal	1.2	89	403.5	29	Yes	Yes - Mini
20	M	7	1.27	24.75	15.35	Normal	0.8	88	459.5	32	Yes	Yes - Mini
21	F	13	1.57	133.20	54.04	Obese	16.9	81	425.5	35	Yes	No
22	M	12	-	-	-	-	1.3	88	503.0	21	Yes	Yes - Mini
23	M	12	1.50	54.00	24.00	Overweight	0.0	83	474.5	22	Yes	No
24	M	12	1.58	42.10	16.86	Normal	1.6	86	468.8	24	Yes	Yes - Mini
25	M	6	1.28	32.50	19.84	Obese	2.6	70	420.0	22	Yes	Yes - Mini
26	F	6	1.24	25.70	16.71	Normal	0.0	83	480.4	29	No	Yes - Mini
27	M	7	1.34	30.35	16.90	Normal	0.0	70	352.0	23	Yes	Yes - Mini
28	F	9	1.34	32.32	18.00	Normal	1.5	87	468.6	33	Yes	No
29	F	15	1.49	48.95	22.05	Normal	0.2	76	492.6	19	Yes	Yes - 2
30	F	10	1.51	45.50	19.96	Overweight	0.0	64	290.0	14	Yes	Yes - 2
31	F	7	1.30	26.75	15.83	Normal	0.7	94	443.5	21	Yes	Yes - 2
32	M	12	1.59	59.35	23.48	Overweight	0.8	87	444.5	26	Yes	Yes - 2
33	M	6	1.16	20.30	15.09	Normal	0.4	81	415.5	31	Yes	Yes - 2
34	M	7	1.28	27.50	16.79	Normal	0.4	85	407.5	20	Yes	Yes - 2
35	M	16	1.82	69.15	20.88	Normal	3.0	88	444.0	33	Yes	Yes - 2
36	M	12	1.39	45.35	23.47	Overweight	6.0	85	420.5	20	No	Yes - 2
37	F	9	1.36	44.70	24.17	Obese	2.7	75	402.5	19	Yes	Yes - 2
38	M	8	1.43	45.10	22.06	Obese	0.4	93	481.9	28	Yes	Yes - 2
μ	**27M**	**9**	**1.38**	**41.60**	**18.89**	**Normal**	**0.8**	**83**	**436.7**	**22**		

[a]Shaded regions highlight patients that cannot be used in any analysis, due to issues with CMAS

[b]BMI is calculated based on the charts developed by Cole et al. [126] and Cole et al. [127]

[c]Due to technical issues and device discontinuation, the Actiwatch Mini, CamNTech, was replaced with the Actiwatch 2, Philips, mid-way through data gathering

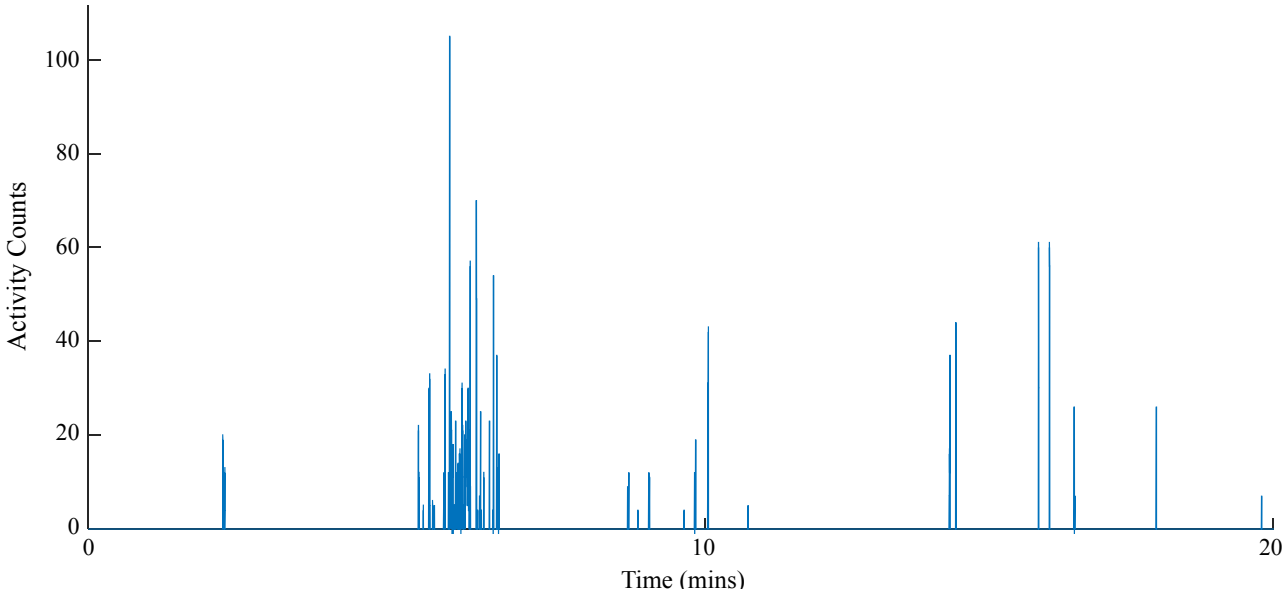


Figure 3.1: Example of 20*mins of an Actiwatch recording.*

receiver unit. Before the sleep study, the CMAS computer was manually synchronised with the computer recording the polysomnography data (illustrated in Fig. 3.5). For greater time-resolution, each CMAS module also recorded a time step (in milliseconds) for each sent packet using the internal oscillator in the micro-controller. The saw-tooth values allow missing packets to be identified post-study. A saw-tooth value is sent per packet of CMAS data (cycling from $1-255$). An example of 80ms and 50s of a CMAS recording is shown in Fig. 3.3. To facilitate temporal synchronisation between CMAS and polysomnography, the raw accelerometry and sawtooth data for both the ankle CMAS module were recorded in the polysomnography montage.

3.1.4 Data collection and clinical study procedure

Three devices were used to record data for each patient: Actiwatch, CMAS and polysomnography. The Actiwatch and CMAS both record accelerometry data (see Section 3.1.3). Polysomnography measures various physiological signals: brain activity, eye movement, muscle tone and movement, respiratory rate, blood oxygenation and others. Fig. 3.4 provides a detailed list of sensors used in the study. The polysomnographic sleep studies in this thesis were conducted using an EMBLA acquisition system (Embla N7000 Bedside Unit, Natus Medical Inc.) and Somnologica software (Somnologica Version 3.3.2 Build 1559). This montage included electroencephalogram (EEG) referential F4-A1, C4-A1 and O2-A1, which sampled at 200Hz with a $16-$bit resolution. All procedures were completed between 2010 and 2014, and were compliant with the Thoracic Society of Australia and New Zealand recommendations for paediatric sleep laboratories [128].

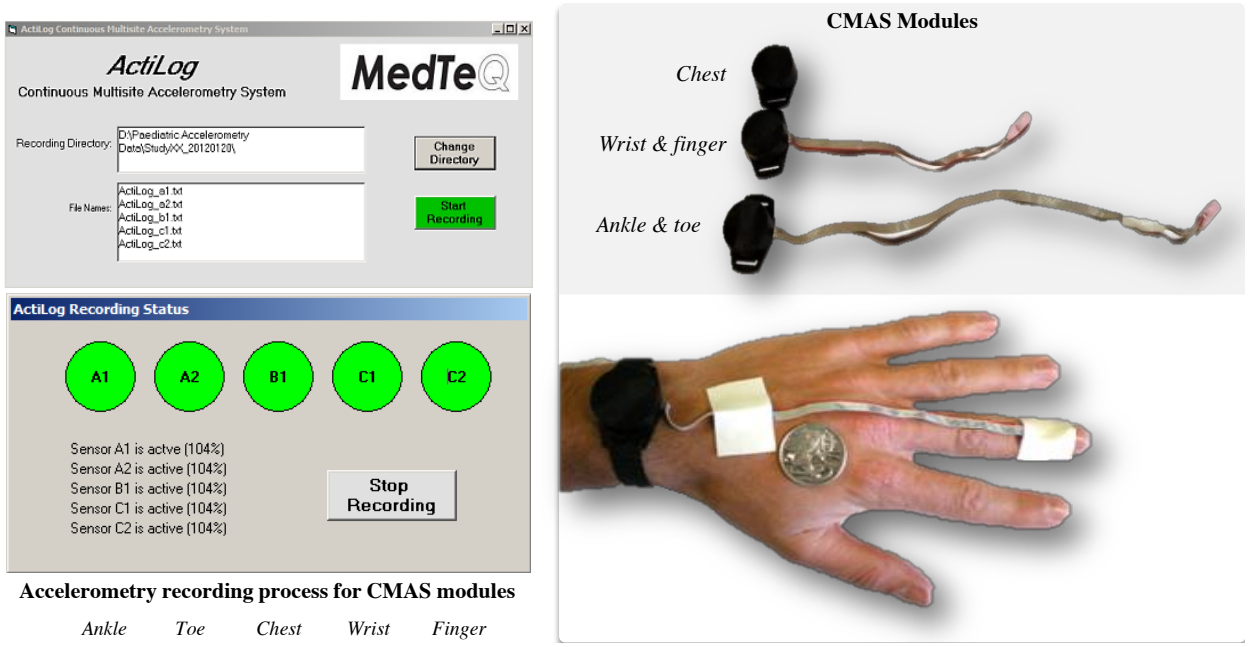


Figure 3.2: CMAS recording software and hardware modules.

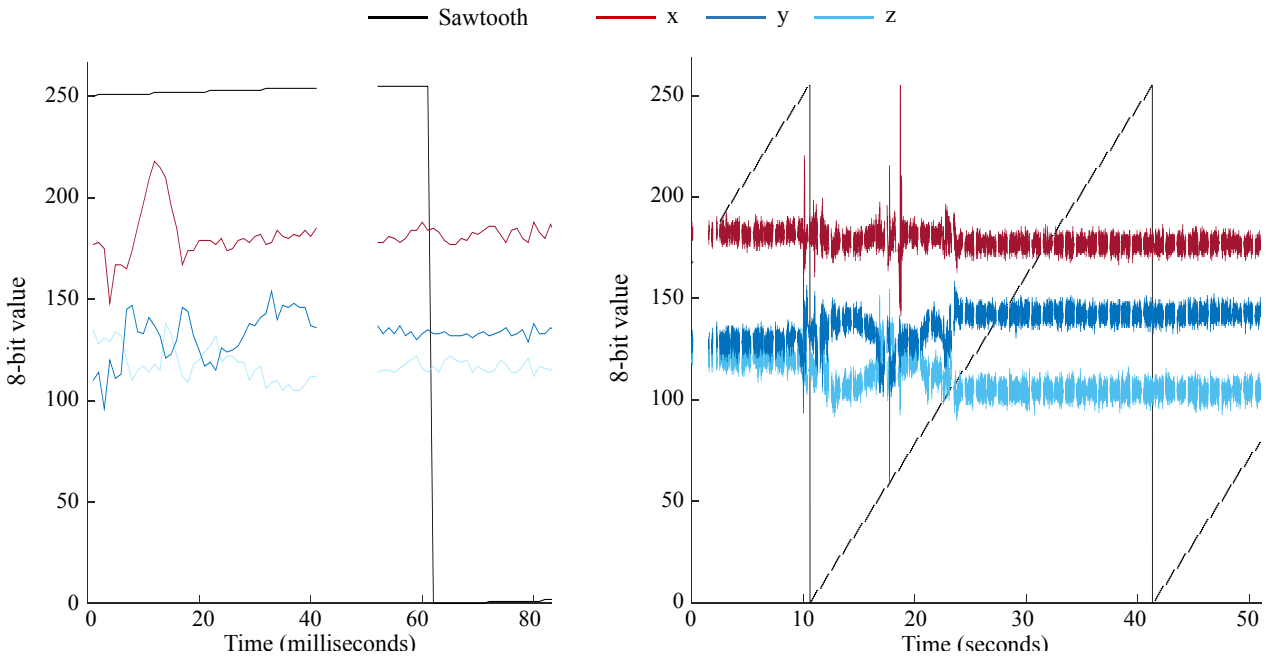
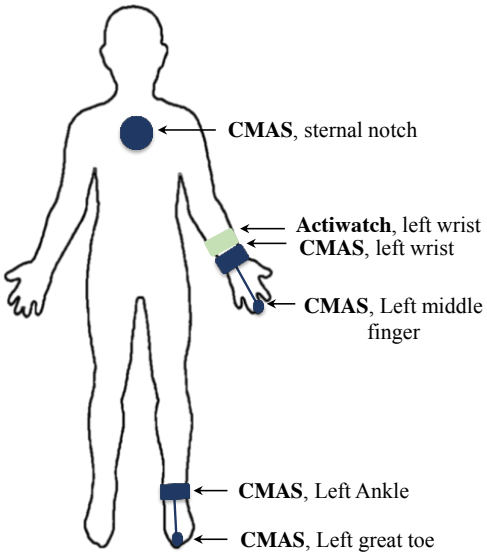


Figure 3.3: Example CMAS data showing the sawtooth and raw tri-axial accelerometry.

The studies were performed at the Mater Children's Hospital Sleep Unit in Brisbane, Australia between approximately 4pm and 8am, as illustrated in the timeline in Fig. 3.6. The manual sleep-scoring was performed by sleep technicians at approximately 11am of the day post-study. The sleep unit at the Mater Children's Hospital followed the recommendations by the Australian Sleep Association (ASA) and Australian Sleep Technologists Association (ASTA) commentary and addendum to the American Academy of Sleep Medicine (AASM)



Accelerometry

Actiwatch Mini, CamNTech
- 2 second epochs;
- Uni-axial accelerometry; and
- Zero crossing (ZC) representation method.

Continuous Multisite Accelerometry System (CMAS)
- Raw 8-bit accelerometry, sampled at 100Hz;
- Tri-axial accelerometry;
- Multi-site capabilities (finger, wrist, upper thorax, ankle and toe); and
- Capabilities for any representation method (including TAT, DI and ZC).

Polysomnography, Somnologica

Brain activity with electroencephalography (EEG)
- Referential F4-A1 (200Hz);
- Referential C4-A1 (200Hz); and
- Referential O2-A1 (200Hz).

Eye movement with electrooculography (EOC)
- Left & Right (200Hz).

Heart rhythm with electrocardiography (ECG) (200Hz)

Muscle activity with electromyography (EMG)
- Left & Right legs (200Hz);
- Submental Chin (200Hz); and
- Diaphragm (200Hz).

Oxygen flow
- Flow subnasal with a pressure sensor (20Hz);
- Flow subnasal & oral with a thermistor sensor (20Hz); and
- Chest & abdominal respiratory inductance plethysmography (RIP) (10Hz).

Blood oxygenation and heart rate
- Sp02 and heart rate (Masimo Radical) oximeter (10Hz); and
- TcpC02 (Radiometer TCM CombiM) (2Hz).

Snoring or sleep talking with an external microphone

Figure 3.4: Sensors used to measure movement and sleep characteristics.

guidelines [38, 43, 44], as outlined in Table 2.1 in Section 2.1.1.

The Actiwatch and CMAS receiver unit and individual modules were typically charged on the morning of the study. The CMAS receiver unit and modules were configured and connected to the polysomnogram computer prior to the patient's arrival at the hospital. The receiver unit was taped to a wall inside the study room and the USB cable was used to connect the receiver unit to the polysomnogram computer located outside of the room. Once the hardware was configured, the CMAS recording software on the CMAS PC was initiated. This also checked that data transmission from the CMAS modules was working. The patient's polysomnography file was then configured using Somnologica Studio (Embla). Additional traces were added to the polysomnography montage to record raw accelerometry data and a timing channel from CMAS.

Upon arrival at the hospital, the child and parent would be presented with information pertaining to the study. If the child and parent agreed to undertake the study they would then sign the consent forms. The child and parent would then often leave the hospital for dinner, etc. When the child was ready for sleep, a full diagnostic polysomnogram was configured on the child by a trained paediatric sleep nurse or by the research scientist. The CMAS modules were connected to the left index fingertip, left wrist, upper thorax, left ankle and left great toe using a combination of tape and hospital bands. The CMAS modules were taped twice to

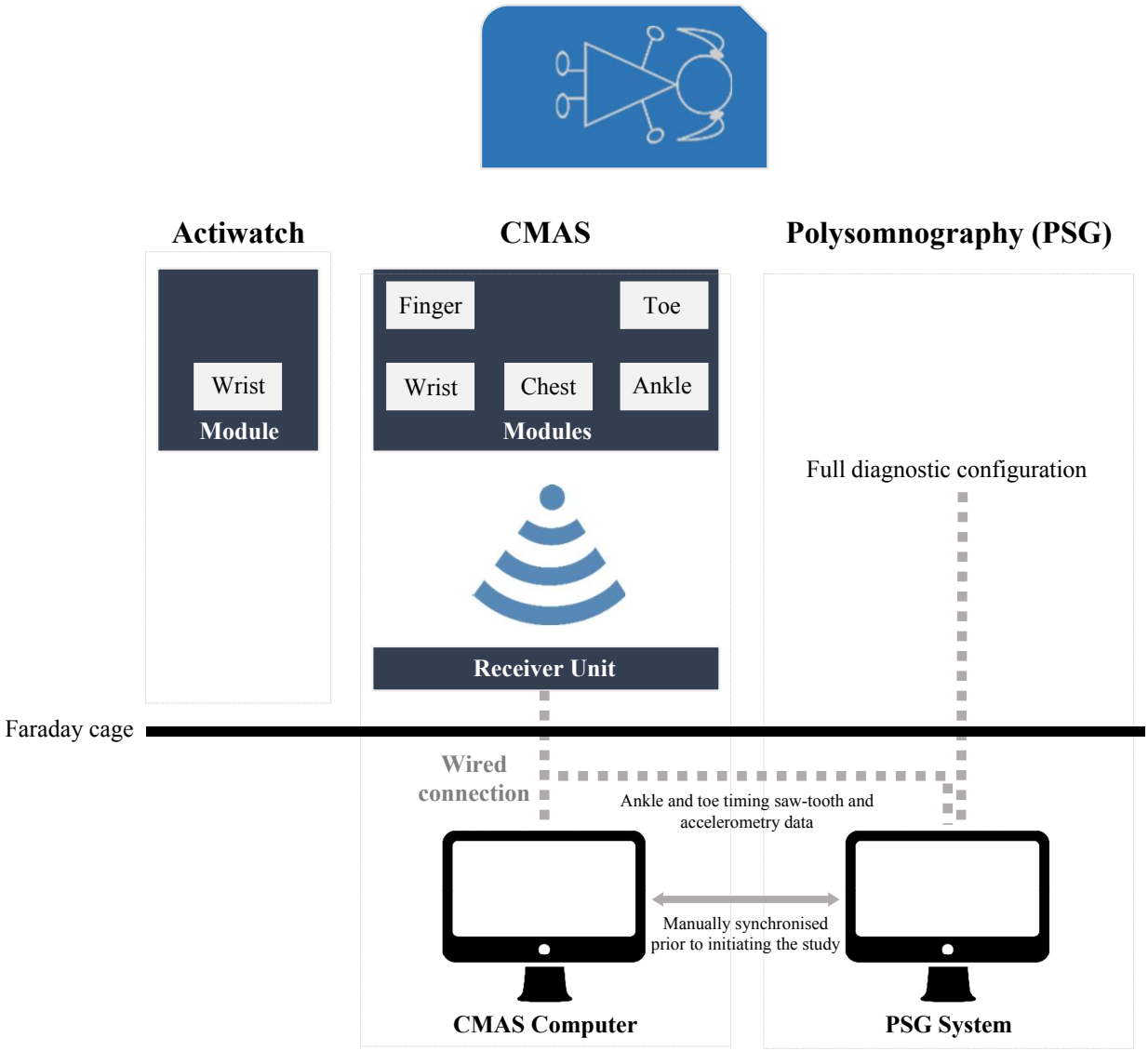


Figure 3.5: Relative study configuration for Actiwatch, CMAS and polysomnography. The polysomnography diagnostic configuration is detailed in Fig. 3.4.

ensure that the housing did not get removed while the patient slept, exposing cell batteries and electronics. The Actiwatch was also configured on the patient's left wrist. The child slept in their own clothing and a parent was normally present during the study. During the study, the CMAS accelerometry data was logged and stored on the CMAS PC with the study ID and date as the folder name. The CMAS ankle and toe modules were also logged in the polysomnogram montage with the other physiological signals.

After the study, the sleep stages (rapid eye movement (REM), non-REM 1:3 and WAKE) and physiological events (sleeping position, arousal, and central or obstructive apnoea or hypopnea) were manually scored by a sleep technician according to the AASM guidelines (detailed below). The physiological recordings in the polysomnogram montage were exported to the European Data Format (.edf), which was then de-identified and stored in the patients study folder

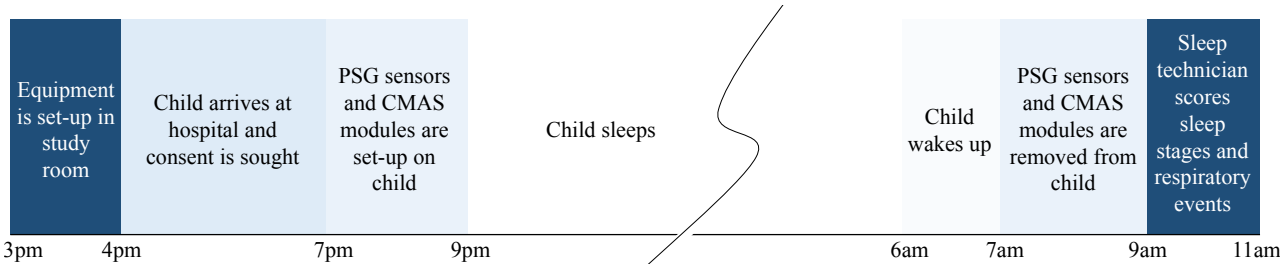


Figure 3.6: Timeline of general sleep study routine.

for post-processing and analysis. The hypnogram and events were exported from Somnologica and stored as text files. The Actiwatch data was downloaded from the Actiwatch and stored in the patient's study folder.

3.2 Post-study processing

The MATLAB (v.R2014b, The MathWorks Inc.^{TM}) environment was used to perform all processing and analysis in this thesis. Data were manually processed post-study to ensure that the three devices (Actiwatch, CMAS and polysomnography) were temporally aligned. Without this, it is difficult to analyse the relationships of movements and their corresponding sleep stages and/or association with physiological and pathological events, as the sampling rate of the different devices differ: CMAS records raw accelerometry at 100Hz (every 0.01s), Actiwatch records activity counts at 0.5Hz (every 2s) and polysomnography produces scored sleep stages at 0.033Hz (every 30s). The methods for synchronising the recordings are detailed in the following sections. CMAS transmits data from the modules wirelessly to the receiver unit and can consequently lose data packets. The techniques used to account for missing data and process the CMAS signals are also detailed below.

3.2.1 Synchronisation of the polysomnogram and accelerometers

The CMAS and polysomnography recordings were often temporally misaligned because they logged data on separate computers with different on-board clocks. Furthermore, the different devices were often started at different times. Both of these offsets were removed to ensure that movements during sleep and wake were accurately analysed.

To account for the different system clocks, the CMAS receiver unit logged the ankle accelerometry data and the saw-tooth timing channel as additional traces in the polysomnography montage. A custom software graphical user interface (GUI) (developed in MATLAB) was used post-study to display and manually align the CMAS traces and the corresponding CMAS raw recordings. As illustrated in Fig. 3.7a, the temporal offset between the CMAS recordings and the recorded accelerometry traces in the polysomnography montage can easily be seen in the

custom GUI. Synchronisation was performed in two steps: at a low temporal resolution, significant regions of movement were aligned using the ankle and toe accelerometry recordings; and, once these were approximately aligned, the software module was zoomed-in to display data at a greater temporal resolution to then align the saw-tooth recordings (shown in Fig. 3.7b). This offset was used to permanently shift the CMAS signals. The synchronised CMAS recordings were saved in the corresponding patient file upon exiting the custom software module.

Next, the start of the polysomnography sleep scores and the CMAS recordings were temporally aligned. A heuristic was applied to each of the recordings to determine the time of the first sleep score: if the polysomnography data occurred first, then the sleep scores were discarded until the first CMAS data sample that aligned with a 30s sleep score, otherwise CMAS samples were discarded until the first 30s sleep score that aligned with the CMAS data.

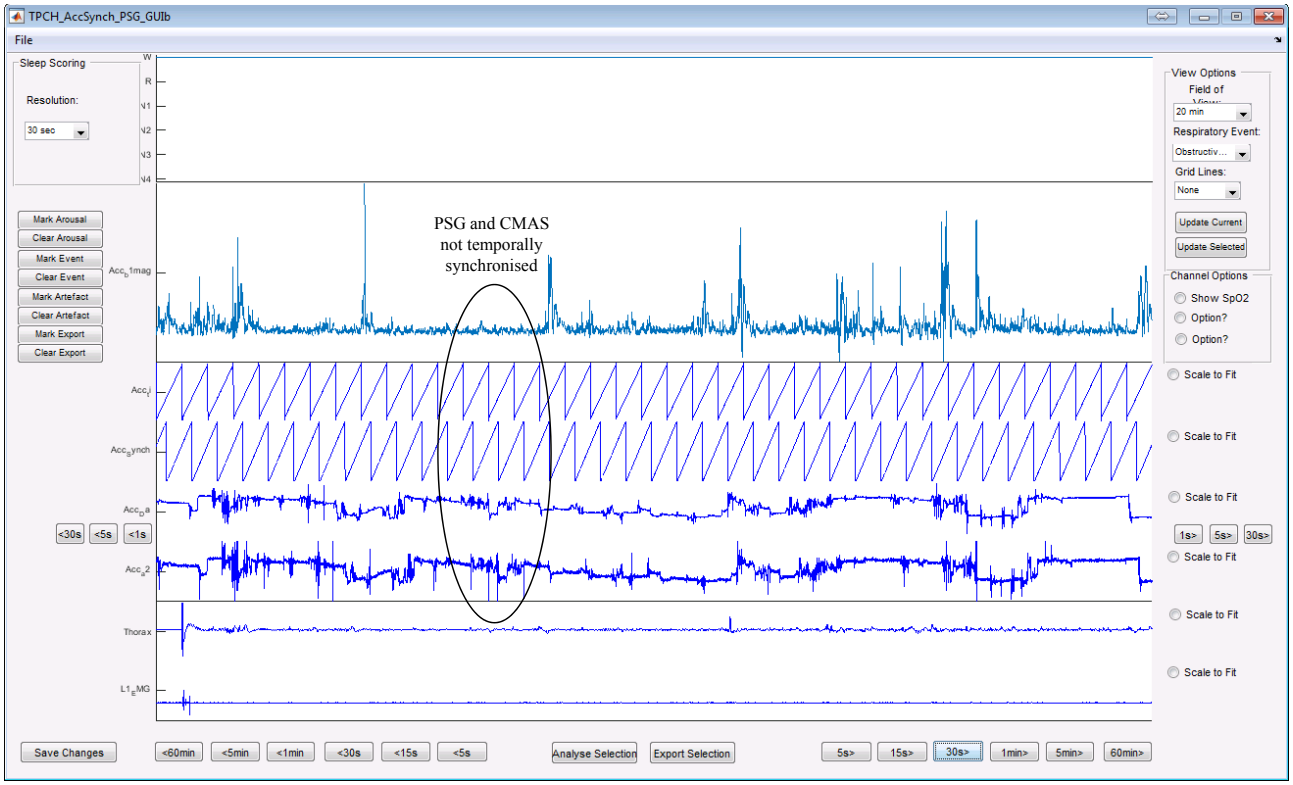
To allow for sample-by-sample analysis, the manual sleep scoring was re-sampled to 100Hz: each manual sleep score was duplicated every 100 times within each second for each 30s sleep score (resulting in 3000 duplications for each sleep score). This ensured that every CMAS sample had a corresponding sleep score, while retaining the 30s resolution of the manual sleep scores.

3.2.2 Representation of sleep scoring from polysomnography

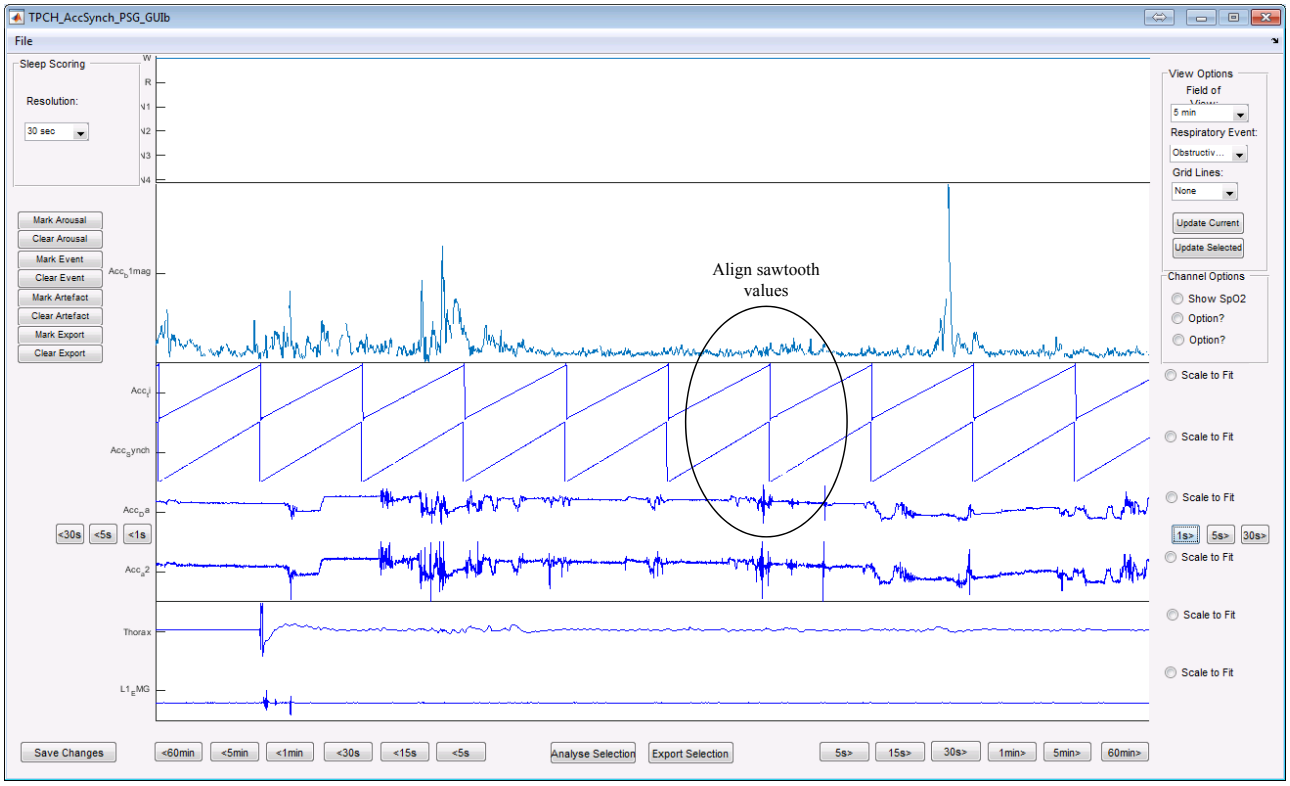
The sleep staging from polysomnography was manually scored into REM, non-REM 1:3 and WAKE stages for each child. The manual sleep scores were exported into a text file using the Somnologica Suite. This text file was then imported into MATLAB where the sleep stages were extracted and stored as a matrix within the corresponding patient study file. To fully encapsulate the appropriate sleep data, the matrix stores the time-stamp, a representation of the sleep stage and whether the sleep stage corresponds to *sleep* or *wake*. The sleep stages are represented as:

Sleep Stage	Representation	
	Stage	Sleep/Wake Flag
WAKE	5	1
REM	4	0
non-REM sleep stage 1 (N1)	3	0
non-REM sleep stage 2 (N2)	2	0
non-REM sleep stage 3 (N3)	1	0

where all sleep stages are flagged as 0 and the wake stage is flagged as 1. This flag was mostly used during the analyses because movements during all sleep (REM and all non-REM stages, flag 0) and wake (flag 1) were primarily analysed. Movement during REM (flag 4), all non-REM stages (flags $1-3$) and wake (flag 5) were used when specifically stratifying events into those occurring during non-rapid eye movement (NREM) and REM sleep.



(a)



(b)

Figure 3.7: Custom MATLAB graphical user interface for synchronising polysomnography with CMAS, (a) prior and (b) post synchronisation alignment.

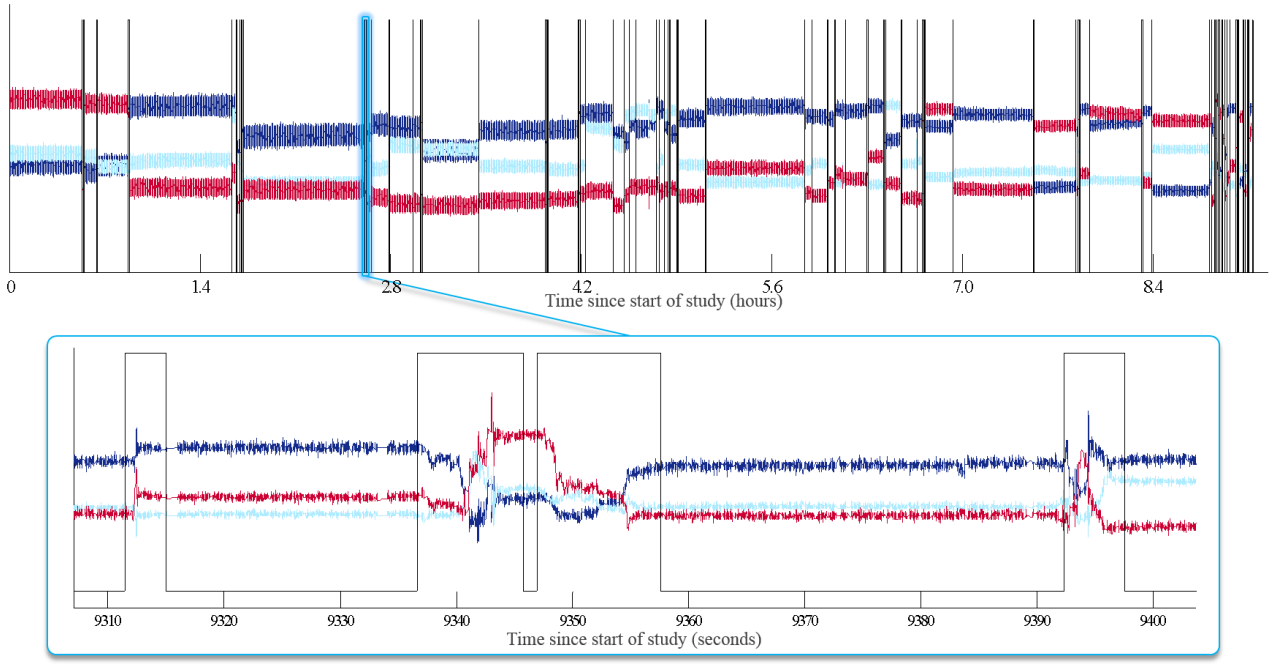


Figure 3.8: Movement detection using the t-test to isolate changes in sample mean. Top shows a full nights study of $(x, y, z)-$axis wrist movement. Bottom shows a segment with four detected movements. The black sections show the regions of detected movement. See Fig. 3.10 for the accuracy.

3.2.3 Movement segmentation algorithm

Movements were segmented to analyse specific characteristics in Chapter 5 and temporally compare movements with physiological and pathological events in Chapter 6. The movements for these experiments were segmented using the procedure outlined below.

The location of movements in the raw accelerometry data was determined using a t-test at a 0.01% significance level on a 1s sliding window of data. The t-test compares the 1s window with the next 0.25s of raw data and returns a 1 if the means differ and a 0 if the means do not differ. This method results in a binary value at each sample location, indicating the occurrence of a movement. A 0.5s dilation was performed on the detected regions to smooth any transitions and to ensure that the movement bounds are well within the detected region. An example of the applied process is shown in Fig. 3.8.

The start temporal location of each movement was aligned to ensure that characteristics can be compared accurately. The start of a movement was defined when the gradient change of the signal to noise ratio (SNR) of the signal $S\dot{N}R$ is greater than the $N = 25$ samples (0.25s) prior to the detected region. This 0.25s prior to the detected region primarily contained noise because the 0.5s dilation expanded the detected region outside of movement, as illustrated in Fig. 3.9. The gradient change of the SNR was determined by,

$$S\dot{N}R[n] = 100 \left| \frac{d}{dt} \left(\sum_{i=n}^{n+N} | \, SNR(x, y, z)[i] \, | \right) \right|. \tag{3.1}$$

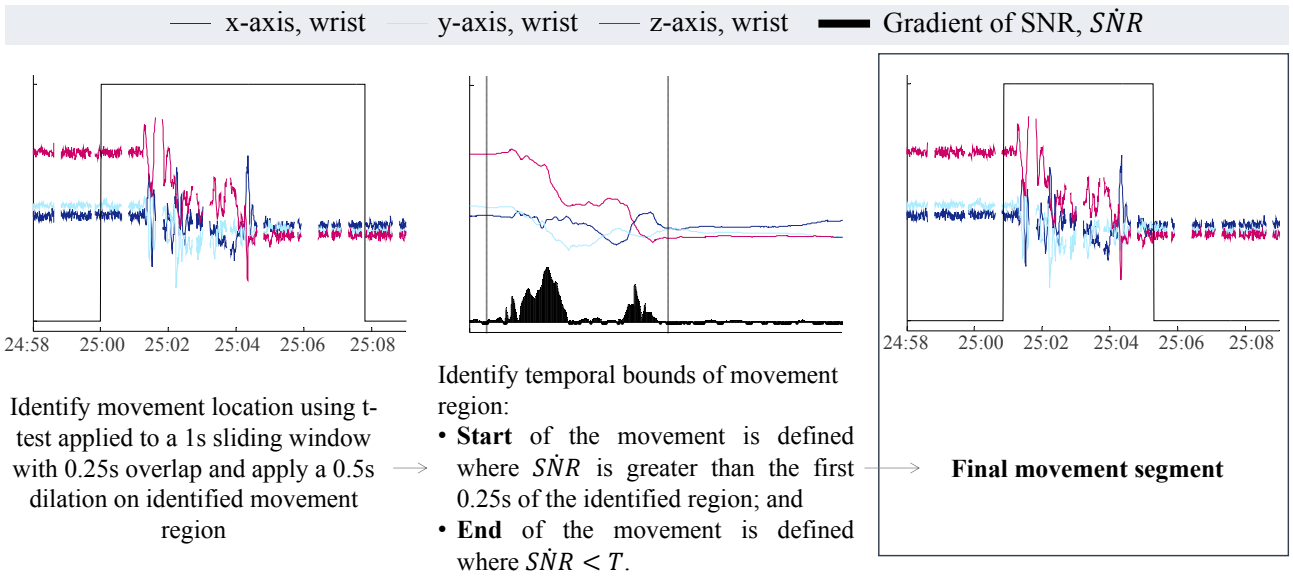


Figure 3.9: Process for segmenting movement regions from raw tri-axial accelerometry.

The end temporal location of each movement is defined as when the SNR of the gradient change drops below a threshold for the final time within the detected region. This process ensured that the movement segments were consistently defined. After segmenting movements, detected movements that occurred within 2s of each other were considered to belong to a single movement. The process is illustrated in Fig. 3.9. After segmenting movements, the segments and their corresponding sleep stage label were extracted.

The accuracy of this algorithm was determined using a custom MATLAB graphical user interface (see Appendix D) with 10 patients ($6-13$ years, median 8.5 years, 5M/5F, AHI range $0-16.9$, median 1.4). Of the 91% of actual movements that were detected by the algorithm, only 6% did not contain movement. The confusion matrix is shown in Fig. 3.10.

3.2.4 Pre-processing procedure and data representation of accelerometry signals

CMAS records 8−bit accelerometry data into a text file for each module. This text file is imported into MATLAB and saved into the patient's study data as a single matrix for each

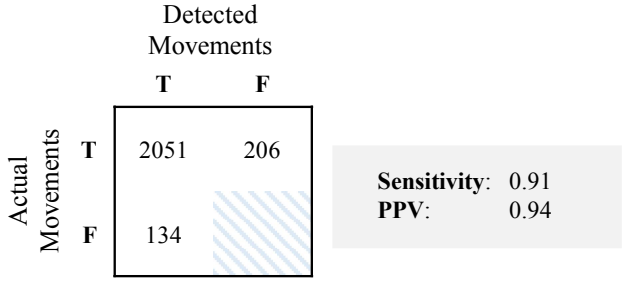


Figure 3.10: Confusion matrix for assessing the accuracy of the automated movement detection algorithm on 10 patients.

module. The accelerometry matrix for each module records the time-stamp, saw-tooth value and the $x-$, $y-$, and $z-$axis samples (an example of a CMAS data matrix is shown in Fig. 3.12).

CMAS accelerometry packets can be lost due to the wireless transmission between the CMAS modules and the receiver unit. For regions of missing data that were less than 2s (200 samples) in duration, the missing samples of each axis for each CMAS module were interpolated using cubic interpolation. The restriction in duration ensured that the interpolation procedure did not create any artificial movement artefacts. Regions of missing data greater than this duration were represented as NaNs (i.e. missing data).

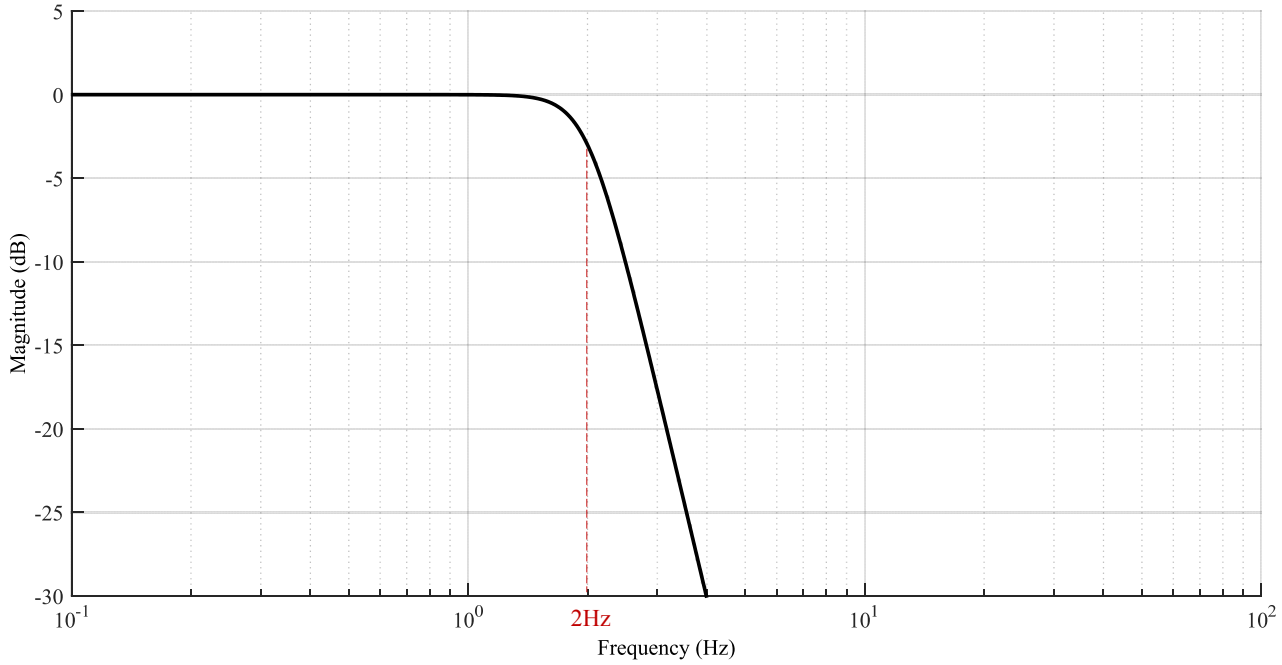
Chapter 4 derives the conventional movement representations (described in Section 2.3.1) using the CMAS accelerometry data. A band-pass filter (BPF) and a low-pass filter (LPF) were used to isolate the high-frequency and low-frequency spectral content necessary for deriving these representations using the CMAS raw accelerometry data. The high-frequency content was used to derive ZC, time above threshold (TAT) and digital integration (DI), and the low-frequency content was used to derive integrated angle of posture change (SUMPST) and maximum magnitude of acceleration (MAXACT) (described in Section 4.1.1). As shown by the frequency responses in Fig. 3.11, the characteristics of the filters were:

- 10^{th} order (5 up, 5 down) band-pass Butterworth filter with cut-off frequencies 2Hz and 12Hz; and
- 5^{th} order low-pass Butterworth filter with a cut-off frequency of 2Hz.

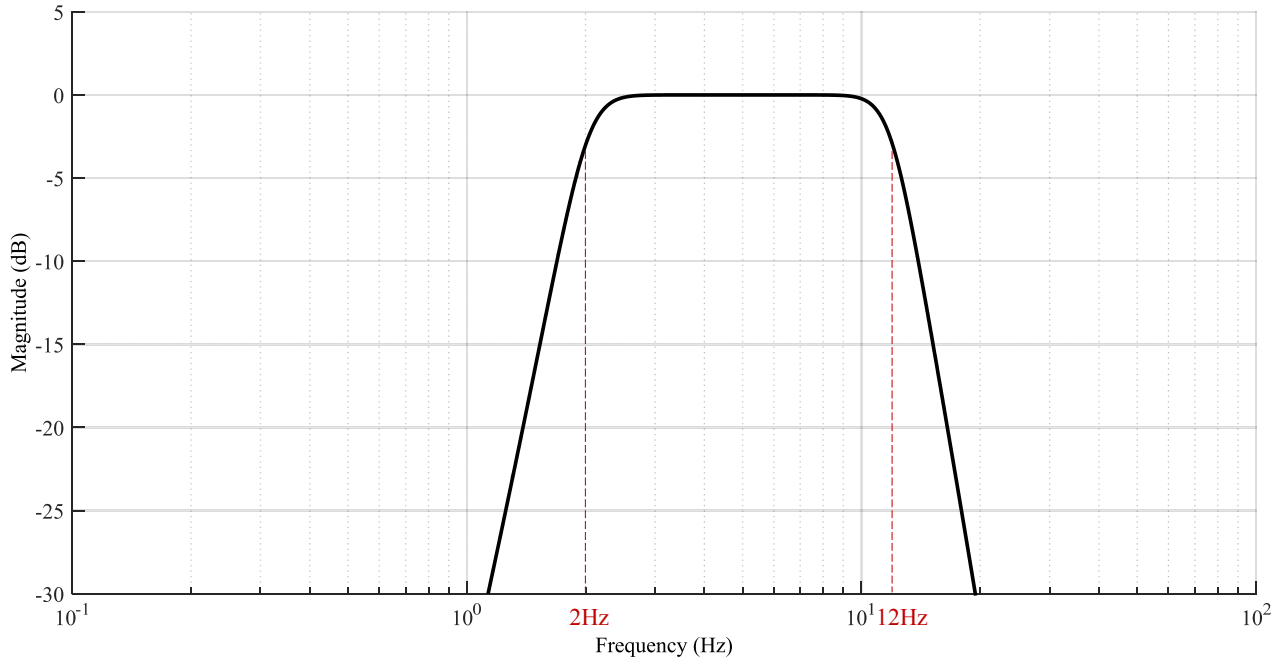
A zero-phase Butterworth filter was chosen because it gives minimal signal distortion in the pass-band and linearly rolls off in the stop-band. Zero-phase is required so that the characteristics of movements (e.g. postural changes) in the accelerometry signals are not affected by the filtering process [129]. The higher-order term was chosen to account for the relatively slow roll-off that is characteristic of Butterworth filters (with a roll-off of 100dB/decade only 10% of the original signal amplitude is present at 2.2Hz for the LPF, and 1.8Hz and 14Hz for the BPF). The cut-off frequencies were chosen on the basis of sample spectral analysis of raw accelerometry during typical movements, which indicated that information content predominantly occurred at frequencies lower than 12Hz.

3.3 Overview of final data set format for analysis

The final data set consists of the CMAS time-stamp, saw-tooth, and $x-$, $y-$, and $z-$axis data, the Actiwatch activity counts and the polysomnography time-stamp, sleep stage flag and sleep/wake flag. These are shown in Fig. 3.12.

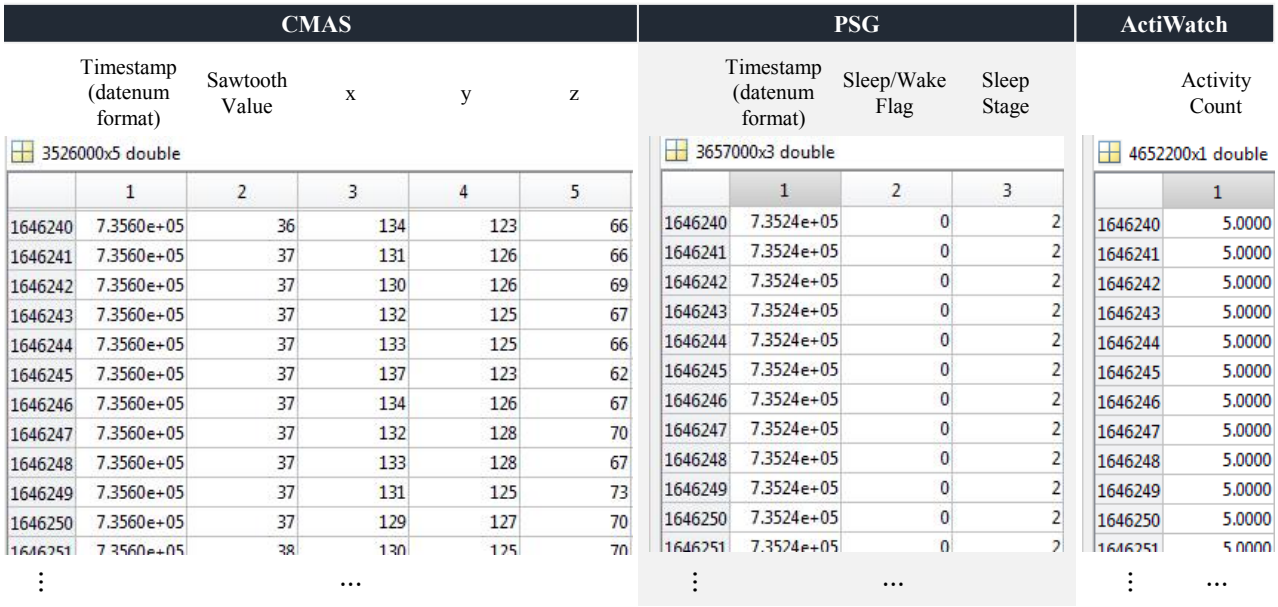


(a)



(b)

Figure 3.11: Filter response for the (a) low-pass filter and (b) band-pass filter.



	1	2	3	4	5
1646240	7.3560e+05	36	134	123	66
1646241	7.3560e+05	37	131	126	66
1646242	7.3560e+05	37	130	126	69
1646243	7.3560e+05	37	132	125	67
1646244	7.3560e+05	37	133	125	66
1646245	7.3560e+05	37	137	123	62
1646246	7.3560e+05	37	134	126	67
1646247	7.3560e+05	37	132	128	70
1646248	7.3560e+05	37	133	128	67
1646249	7.3560e+05	37	131	125	73
1646250	7.3560e+05	37	129	127	70
1646251	7.3560e+05	38	130	125	70

	1	2	3
1646240	7.3524e+05	0	2
1646241	7.3524e+05	0	2
1646242	7.3524e+05	0	2
1646243	7.3524e+05	0	2
1646244	7.3524e+05	0	2
1646245	7.3524e+05	0	2
1646246	7.3524e+05	0	2
1646247	7.3524e+05	0	2
1646248	7.3524e+05	0	2
1646249	7.3524e+05	0	2
1646250	7.3524e+05	0	2
1646251	7.3524e+05	0	2

	1
1646240	5.0000
1646241	5.0000
1646242	5.0000
1646243	5.0000
1646244	5.0000
1646245	5.0000
1646246	5.0000
1646247	5.0000
1646248	5.0000
1646249	5.0000
1646250	5.0000
1646251	5.0000

Figure 3.12: CMAS, polysomnography and Actiwatch data format for analysis in MATLAB.

3.4 Outcome measures

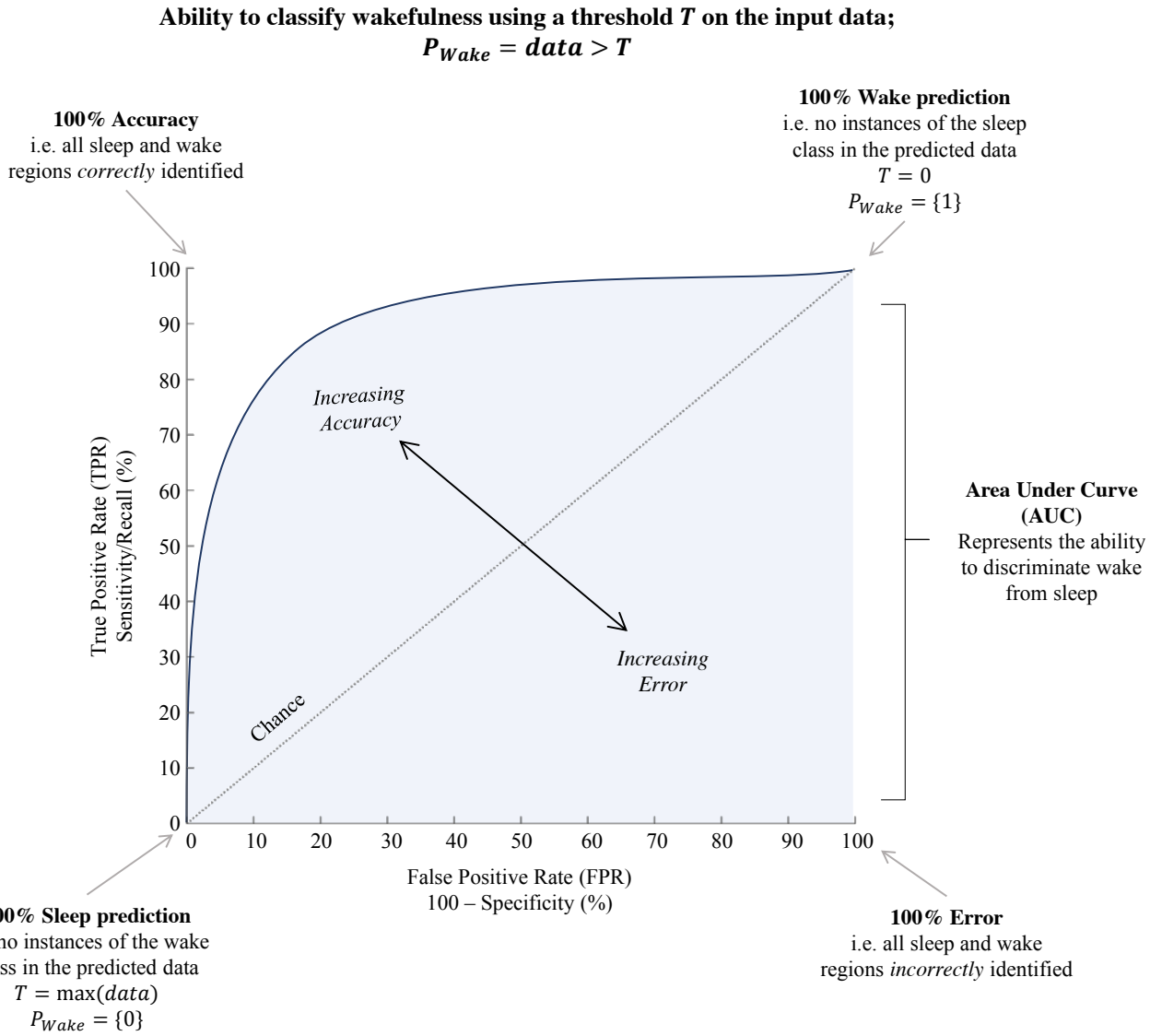
Chapter 4 and Chapter 5 seek to predict binary class labels (i.e. sleep and wake). Throughout these analyses, the classification performance will be compared using receiver operating characteristics (ROC) analysis. The different metrics derived from this technique are described in Fig. 3.13. The ROC curve represents the predictive performance for all possible classification thresholds $0 \leq T \leq \max(data)$. For most classification problems in this thesis we will be analysing the area under the receiver operating characteristics curve (AUC), sensitivity and specificity (defined below). The AUC represents the ability to rank a randomly chosen positive instance higher than a randomly chosen negative instance. Therefore, AUC represents the ability of the techniques to discriminate between the binary classification states (sleep/wake) [130]. An advantage of analysing AUC is that, unlike measures like agreement rates, it is prior probability invariant. This is particularly important for sleep analysis because the prior probability of sleep and wake differs between each patient.

The ability to accurately detect sleep is represented by the sensitivity of a method. Sensitivity is defined as the percentage of actual sleep epochs or samples (as identified with gold standard or manual classification) correctly predicted as 'sleep',

$$Sensitivity = \frac{TP}{TP + FN}. \tag{3.2}$$

Similarly, the ability to accurately detect wake is represented by the specificity. Specificity is defined as the percentage of wake epochs or samples correctly predicted as 'wake',

$$Specificity = \frac{TN}{TN + FP}. \tag{3.3}$$



		Actual: *Wake*	Actual: *Sleep*	
Predicted	*Wake*	True Positive TP	False Positive FP	$Precision = \frac{TP}{TP + FP}$
	Sleep	False Negative FN	True Negative TN	
		$TPR, Sensitivity, Recall = \frac{TP}{TP + FN}$	$FPR = \frac{FP}{FP + TN}$ $Specificity = \frac{TN}{TN + FP}$	

Figure 3.13: ROC analysis used to compare sleep and wake prediction throughout this thesis, and the associated outcome measures.

Considering that specificity is a major limitation of conventional actigraphy (detailed in Section 2.3), it is one of the main outcome measures throughout this thesis. The specificity at a fixed sensitivity and the identified operating point will often be compared.

An ROC curve indicative of the average performance across the population will often be used to illustrate the generalised performance. An example is shown in Fig. 3.14. All sensitivity values corresponding to each specificity value are grouped across the population. These values are then averaged and plotted as an average ROC curve. The solid line represents a fitted line to the median sensitivity value for each specificity value. The shaded region represents the 75^{th} and 25^{th} percentiles of the grouped sensitivity values. This graphical representation takes the average of all ROC curves across the population. If an individual ROC curve finishes before the top-right corner, the rest of the curve is linearly interpolated. As explained in Fig. 3.13, this would occur if the spread of activity is bimodal; i.e all epoch activity values are high-range, but there still exists many epochs of 0 activity. Reducing the threshold beyond the minimal activity value (i.e. moving the threshold towards the top-right corner of the curve) would not affect predictions because the activity values (excluding 0) cannot fall below the threshold. A set of different incomplete ROC curves means that it is possible for the median sensitivity value to be lower than a previous value, and the average curve will not be monotonically increasing, which is expected of a typical ROC curve. This graphical representation is not intended as a quantitative analytical tool, but as an indication of the general performance across the cohort.

The final outcome measure is the agreement with manually scored polysomnography. Both standard agreement and Cohen's Kappa (κ) [131] are used to compare the accuracy of the different techniques. Kappa adjusts the agreement for chance and is consequently considered a more accurate metric than standard agreement. However, standard agreement will be noted for comparison with literature. Kappa agreement is defined as:

$$\kappa = \frac{p_o - p_e}{1 - p_e}. \tag{3.4}$$

where p_o is the relative observed agreement, and p_e is the hypothetical probability of chance agreement derived from the class prior-probabilities [132]. It is important to note that only one technician manually scored the sleep-related events. Despite the scoring guidelines, the identification and temporal boundaries of events can differ between scorers. Considering events where there is some consensus from multiple scorers would provide additional confidence in the 'ground truth' scores; however, this is not always practical in a clinical setting.

3.5 Statistical procedure for combining distributions

Within Chapter 5 and Chapter 6, multiple distributions were combined using meta-analysis techniques to form a resulting *averaged* pooled distribution [133]. Meta-analysis techniques are often used to integrate results of independent studies that measure the same variable [134].

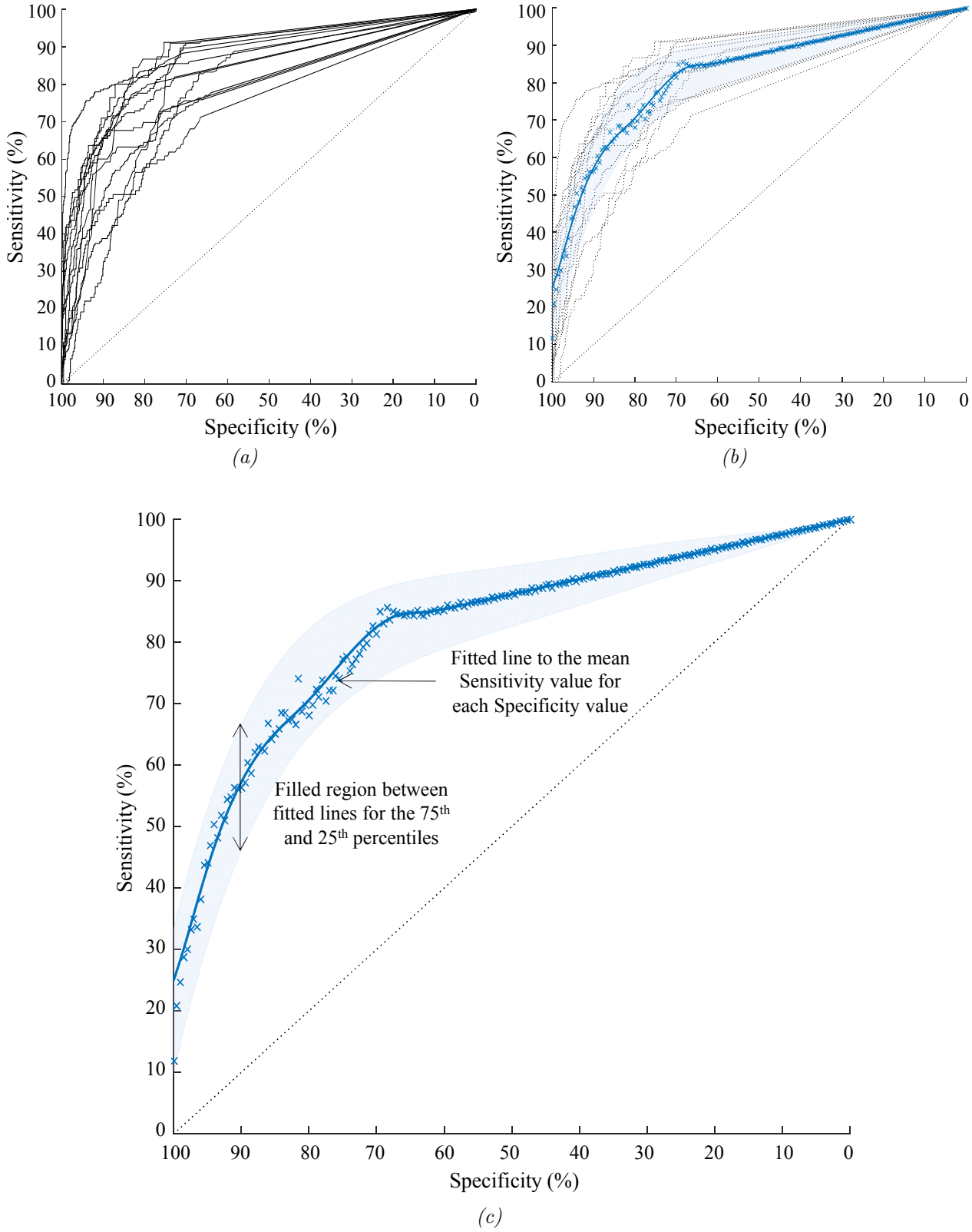


Figure 3.14: Combining ROC curves: (a) individual ROC curves, (b) individual ROC curves overlaid with the average representation, and (c) average ROC curve used to illustrate the performance across the cohort throughout this thesis.

Results for each study are weighted to ensure that large studies have more effect on the resulting distribution than smaller studies. This is done to reduce estimate error that is caused by chance because estimates with a small sample size are more influenced by chance than estimates with a large sample size. The weights w_n represent how much each patient n contributes to the final combined result. The combined mean $\hat{\mu}$ and variance $\hat{\sigma}^2$ is given by [133],

$$\hat{\mu} = \frac{\sum_{n=1}^{N} w_n \mu_n}{\sum_{n=1}^{N} w_n}, \quad w_n = \frac{1}{\sigma_n^2}, \tag{3.5}$$

$$\hat{\sigma}^2 = \sum_{n=1}^{N} \frac{\sigma_n^2}{N}, \tag{3.6}$$

where N is the total number of patients in the meta-analysis, and μ_n and σ_n^2 are the mean and variance of the sample distribution respectively for patient n.

3.6 Outline of research process

The general outline of the study procedure and common pre-processing techniques between each analysis are outlined in Fig. 3.15.

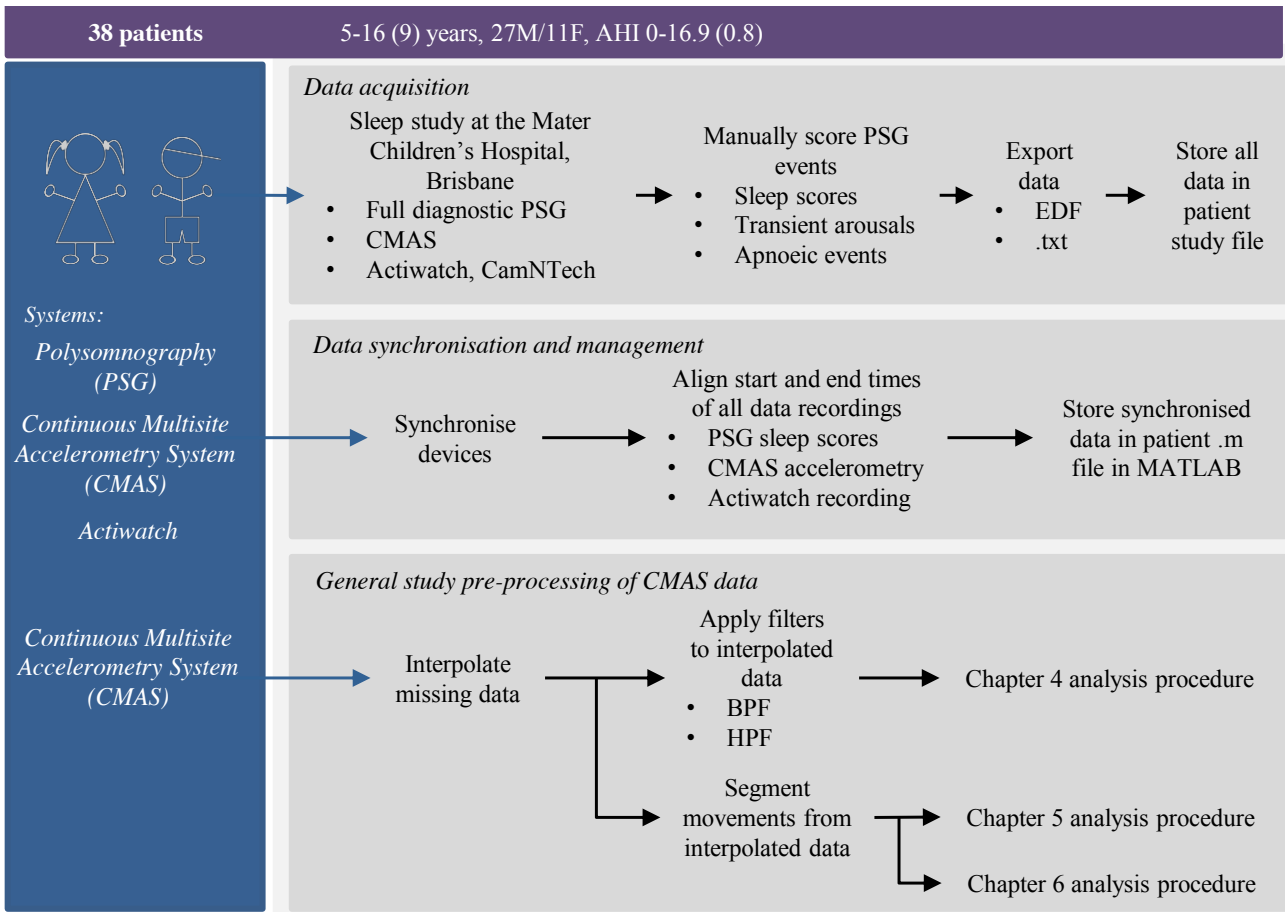
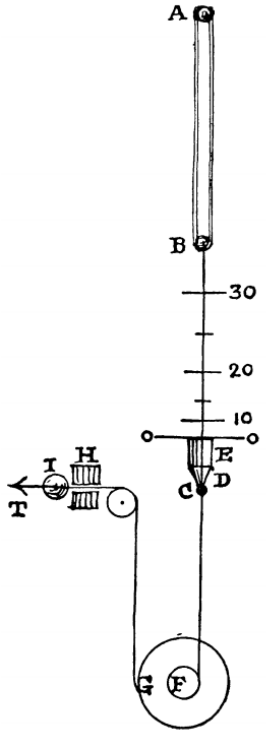


Figure 3.15: Outline of general methodology common to each analysis.

4

Conventional representations of tri-axial multisite accelerometry



In 1891, Sir Francis Galton invented a device for measuring the rate of limb movement using the momentum of an excited ivory bead. The ivory bead (E), threaded on a string (T), sat freely on a stopper (CD). Movement of the limb (attached to the string (T)) caused tension in the string, which then excited the bead. A rubber band (AB) provided the backwards momentum from the excitation. A scale (feet per second) was used to measure the velocity of the ivory bead (and thereby the limb movement) as the position on the scale where the bead reached zero velocity (at the vertex of the trajectory) before falling back to the stopper (CD). This measurement was not mechanically recorded, but noted by an observer.

— Sir Francis Galton, F.R.S, *1822 - 1911*[1]

[1]Description of measurement device detailed in [135].

Conventional actigraphy identifies wake as regions of increased activity and sleep as regions of low activity. As discussed in Chapter 1, and again in Section 2.3, one of the main limitations of this technique is that regions of no activity will always be classified as sleep. This is a limitation because there are often epochs of low activity during wake. These low-activity epochs can be due to inadequate sensor configurations: uni-axial accelerometry measured at the wrist cannot detect movements that are orthogonal to the measurement axis, or that occur elsewhere on the body. Actigraphy will misclassify low-activity wake periods as sleep, resulting in an increased number of false negatives. Therefore, the accuracy of actigraphy-based estimates of wake is dependent on effectively detecting wake periods. In this chapter we seek to reduce false sleep detections by evaluating multisite tri-axial accelerometry.

Commercial systems conventionally use time-series techniques to summarise movement as 'activity counts' (detailed in Section 2.3.1). These methods have been extensively validated in literature across different devices (see Table 2.4). Activity counts are calculated using one of three common methods: zero crossing (ZC), time above threshold (TAT), or digital integration (DI) (see Section 2.3.1) [83]. There are some distinct limitations to these methods that may impact the ability of actigraphy to effectively represent movement: TAT and ZC ignore the amplitude of the acceleration signal because they only determine the time spent above, or the number of times the acceleration signal crosses, a set acceleration value; and DI is unable to differentiate many high amplitude, short-duration motions from a low amplitude, long-duration motion because DI identifies the total 'area' of acceleration that occurs within an epoch. Although these methods are commonly used to derive activity counts for different commercial devices, the method employed is often not documented; only 'activity counts' are reported. This makes it difficult to compare the performance of actigraphy in different studies. Furthermore, direct comparisons of the different techniques on the same clinical data have not been performed in literature.

It is likely that some of the shortcomings of conventional actigraphy are related to the inherent hardware limitations of previously validated systems. Initial actigraphs measured movement with uni-axial accelerometers because, until recently, the cost of additional axes was not worth the performance benefits [121, 136]. However, uni-axial accelerometers are unable to detect movements that are orthogonal to the monitored axis. In addition to the limited movement detection of uni-axial accelerometers, activity counts may not accurately represent movement. This, in turn, limits the number of wake movements that are detected with conventional actigraphy. In addition to this, actigraphy in sleep assessment typically only measures wrist movement. However, movement from other limbs may occur predominantly during wake. Detecting these movements would reduce the number of wake epochs with low activity, consequently reducing false sleep detections.

There have been some studies in sleep literature that have investigated the implications of replacing the conventional wrist accelerometer with shoulder, thorax or hip accelerometers.

These studies found high correlation between waist and non-dominant wrist accelerometers for children [137, 138]; however, the waist accelerometer overestimated total sleep time and sleep efficiency. This suggests that while the waist placement is effective at detecting large movements associated with sleep (such as body positional changes from supine to right or left lateral), it is unable to detect small movements (such as hand twitches) that may occur during short periods of wake. Other literature has compared the shoulder placement to the wrist placement in children [139, 140]. The findings in these studies suggest that the shoulder may be over-sensitive to noise or movements. These studies were performed on children on the autism spectrum because of their inability to tolerate the conventional wrist placement. As such, these conclusions may not be extended to children without these disorders. These studies indicate that other accelerometer placements can be substituted for the conventional wrist placement in sleep assessment. However, each placement has its own limitation and it is unclear in literature if combining data from each accelerometer counters the individual limitations.

Analysis of combining multiple accelerometers has been performed in physical activity assessment and task identification. Gjoreski et al. [141] found that combining accelerometers improved the accuracy of posture-identification in the elderly. This finding is consistent with literature on task detection, which reports that multiple accelerometers aid with discriminating different activities [116]. Despite these results, literature that combines data from multiple accelerometers to predict sleep and wake is lacking.

This chapter specifically addresses the first limitation of actigraphy, discussed in Chapter 1 and illustrated in Fig. 4.1:

False negatives: wake epochs with no observed movement are incorrectly identified as 'sleep'.

This chapter will address the hypothesis:

> *Uni-axial accelerometry measured solely at the wrist limits sleep and wake prediction accuracy because movements orthogonal to the measurement axis, or occurring elsewhere on the body, cannot be detected.*

In this chapter we will identify if incorporating tri-axial multisite accelerometry into the conventional actigraphy framework improves false sleep detections by increasing the number of detectable wake movements. We will first compare movement measured with tri-axial accelerometers to the conventional uni-axial accelerometers, and explore additional time-series movement representations. We will then explore the effect of incorporating additional accelerometers on sleep and wake estimates.

This analysis will be implemented in two sections:

Tri-axial movement representations

Section 4.1 aims to directly compare the conventional movement representations using the same data; and explore the accuracy of sleep and wake predictions when representing movement with tri-axial techniques, relative to the conventional uni-axial representations.

Accelerometer placements

Section 4.2 aims to identify the performance benefits and clinical implications of incorporating multiple accelerometer placements into the actigraphy scoring routine.

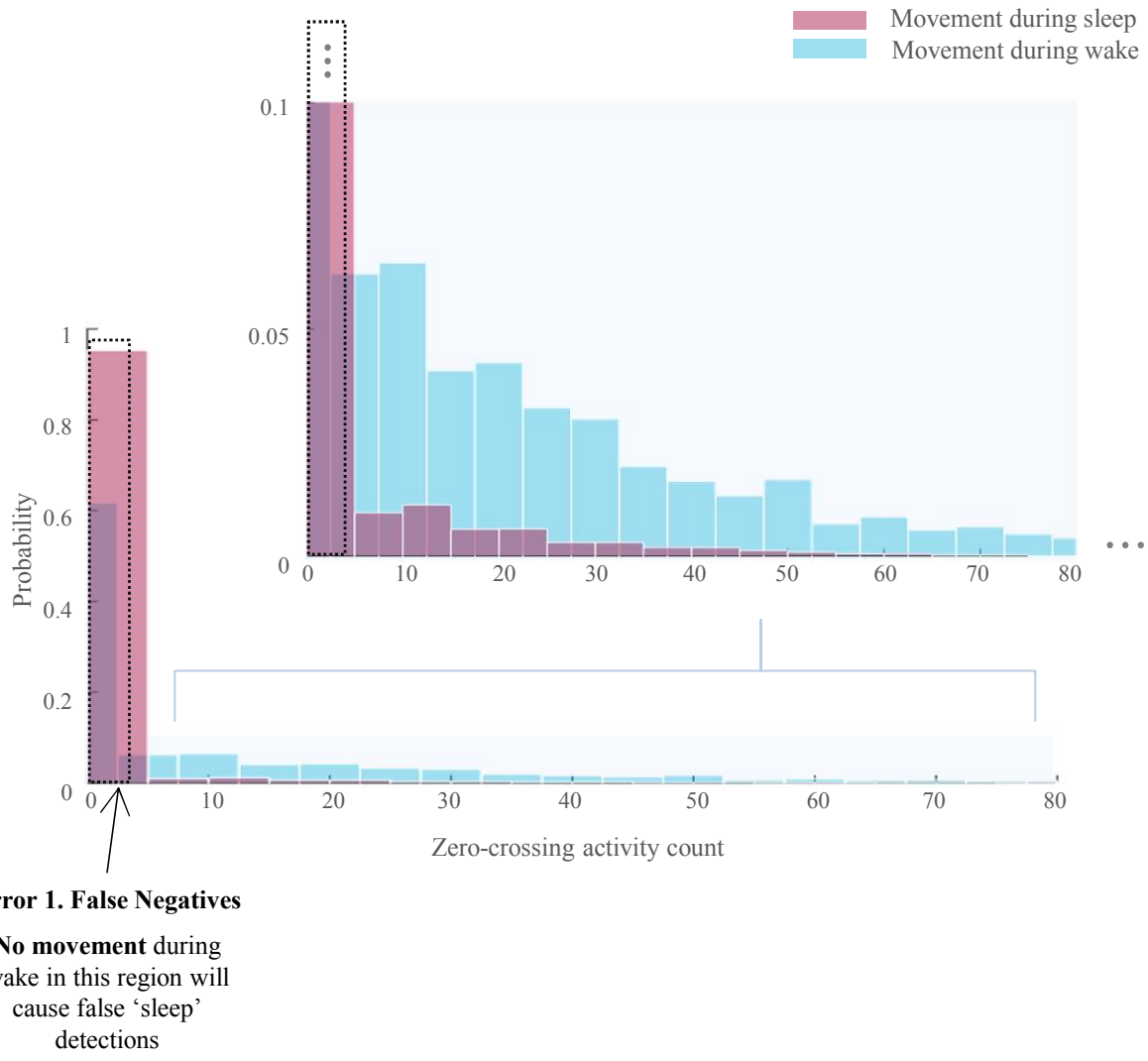


Figure 4.1: Histogram of activity during sleep (red) and wake (blue) for 24 *patients using the conventional activity counts derived using the zero-crossing method. Highlighted region shows the source of the false negatives.*

4.1 Tri-axial movement representations

Activity counts are widely used throughout literature; however, there are re-occurring problems when estimating the sleep quality of patients who have atypical sleeping behaviour (as previously discussed in Section 2.3). Existing actigraphy systems suffer from false sleep detections because of wake epochs occurring with 'no activity'. As discussed in Section 2.3 and specifically highlighted in Fig. 2.8, monitoring movement with tri-axial accelerometry detects a larger range of movement, which increases the amount of representable activity. Incorporating tri-axial accelerometry may consequently reduce the number of wake epochs with no activity. This section will explore improvements to false sleep detections by representing movement with tri-axial information.

In this section we will:

1. Evaluate and compare the ability of the conventional movement representations to differentiate sleep and wake periods using the same data;
2. Compare the performance of the conventional representations derived from uni-axial data with representations derived from tri-axial data; and
3. Evaluate the performance of two novel time-series movement representations that exploit tri-axial accelerometry.

4.1.1 Method

Each patient underwent the study procedure outlined in Section 3.1.4 of Chapter 3. The full methodology for this analysis is summarised in Fig. 4.2. The conventional movement representations were derived for the $x-$axis of the Continuous Multisite Accelerometry System (CMAS) wrist accelerometer. These representations were compared to the uni-axial commercial actigraph, Actiwatch Mini (CamNTech)[2]. The Actiwatch Mini was set to record raw activity in 2s epochs, which were cumulatively combined post-study to form 30s epochs. This was done to allow direct comparison with the 30s manual scoring from polysomnography. This process was performed prior to applying any weighting function; a raw activity count is not reliant on its surrounding epochs. Oakley [123]'s weighted moving average filter was then applied to the uni-axial activity counts (see Section 2.3.1 for more details). Custom weighted moving average filters were developed for the tri-axial activity counts (detailed below).

[2]Actiwatch Mini, CamNTech, `http://www.camntech.co.uk/files/Actiwatch_Mini_Insert.pdf`. While the Actiwatch Mini is now marketed for veterinary use, it was initially developed for paediatric research as it is small and lightweight. For these reasons, and because it was readily available in our sleep laboratory at initial data collection, the Actiwatch Mini was chosen for use.

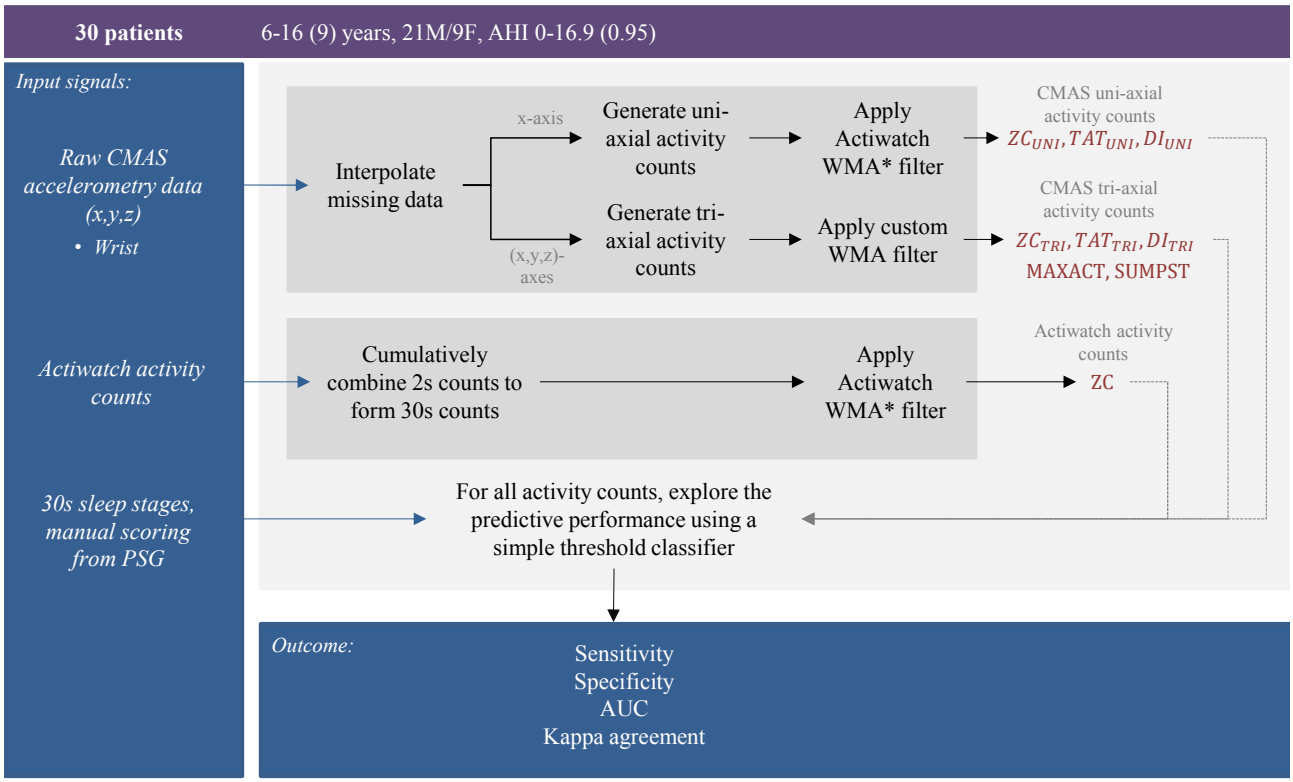


*Figure 4.2: Methodology for analysing and comparing the utility of the conventional activity count derivation techniques for uni-axial and tri-axial data. WMA refers to 'weighted moving average' filters. *The Actiwatch weighted moving average filter is detailed in Section 2.3.1 and in [123].*

Patient characteristics

The full analysis in this section used data from the full 30 patients detailed in Table 3.1.3 (i.e. eight patients in Table 3.1.3 were excluded due to technical issues). The comparison with the Actiwatch used data from 14 patients (patients $2-6$, $11-12$, 15, $19-20$, 22, $24-25$, and 27) aged $6-15$ years (median 8 years, 14 male) with obstructive sleep apnoea syndrome (OSA) severity ranging from healthy to moderate (median apnoea hypopnea index (AHI) 1.15, range $0-5.1$). Seven patients aged $8-16$ years (median 12 years, 4 male) with OSA severity ranging from healthy to severe (median AHI 1.3, range $0-16.9$) (patients $7-10$, 21, 23 and 28) were used for training custom filters. To ensure that there was no bias in the results, these patients were not used in the main analysis.

Derivation of the conventional movement representations

A number of pre-processing steps were required before deriving the movement representations from the CMAS data. Firstly, the raw CMAS data was filtered to remove high-frequency noise and the DC offset caused by the gravitational component of acceleration. To determine the general frequency bounds of movement during sleep and wake, spectral analysis of a sample

window of movement was conducted. This process is described in Section 3.2.4 and illustrated in Fig 3.11 in Chapter 3. ZC, TAT and DI were derived using the band-pass filtered data.

As discussed in Section 2.3 and illustrated by Fig. 2.8, the projection of a movement vector onto the $x-$, $y-$ and $z-$ axes differ. The derived movement representations would consequently differ between these axes. However, the limitation of uni-axial accelerometry (i.e. the inability to detect movements orthogonal to that axis) applies, regardless of the chosen axis. Therefore, the $x-$axis was used to derive the conventional uni-axial representations. The band-pass filter (documented in Section 3.2.4 and illustrated in Fig. 3.11b) was applied to isolate the higher-frequency signal from the $x-$axis. The conventional representations were derived using the process described in Section 2.3.1. Henceforth, the uni-axial representations are termed ZC_{UNI}, TAT_{UNI} and DI_{UNI}.

The tri-axial movement representation derivation process was the same as for the uni-axial data; however, the Euclidean norm operation was applied to the band-pass filtered tri-axial data $\|\mathbf{a}_B\|$ prior to deriving the representations. Similar to the uni-axial representations, the tri-axial representations are termed ZC_{TRI}, TAT_{TRI} and DI_{TRI}.

The tri-axial representations are defined similarly to the uni-axial representations (detailed in Section 2.3):

$$TAT_{TRI} = \sum_{n=1}^{N \cdot f_s} \mathbb{1}\{\|\mathbf{a}_B[n]\| > T\}, \tag{4.1}$$

$$DI_{TRI} = \sum_{n=1}^{N \cdot f_s} \|\mathbf{a}_B[n]\|, \tag{4.2}$$

$$ZC_{TRI} = \sum_{n=1}^{N \cdot f_s} \mathbb{1}\{(\|\mathbf{a}_B[n]\| - T) \cdot (\|\mathbf{a}_B[n-1]\| - T) < 0\}, \tag{4.3}$$

where $N = 30$s represents the window size, $f_s = 100$Hz represents the sampling rate, T represents a threshold and $\mathbb{1}$ represents the indicator function (i.e. a function that has a value of 1 if the condition is satisfied, otherwise 0) .

In order to exploit the nature of tri-axial accelerometry, two novel representations were also extracted from CMAS: maximum magnitude of acceleration (MAXACT) and integrated angle of posture change (SUMPST). Although the maximum magnitude has been used in physical activity literature [142, 143], it is not a conventional method for representing movement during sleep. As such, MAXACT is a somewhat novel representation for sleep analysis. MAXACT represents high intensity movements by measuring the maximum acceleration within an epoch,

$$MAXACT = \max(\forall n \in N \cdot f_s : ||\mathbf{a}_B[n]||). \tag{4.4}$$

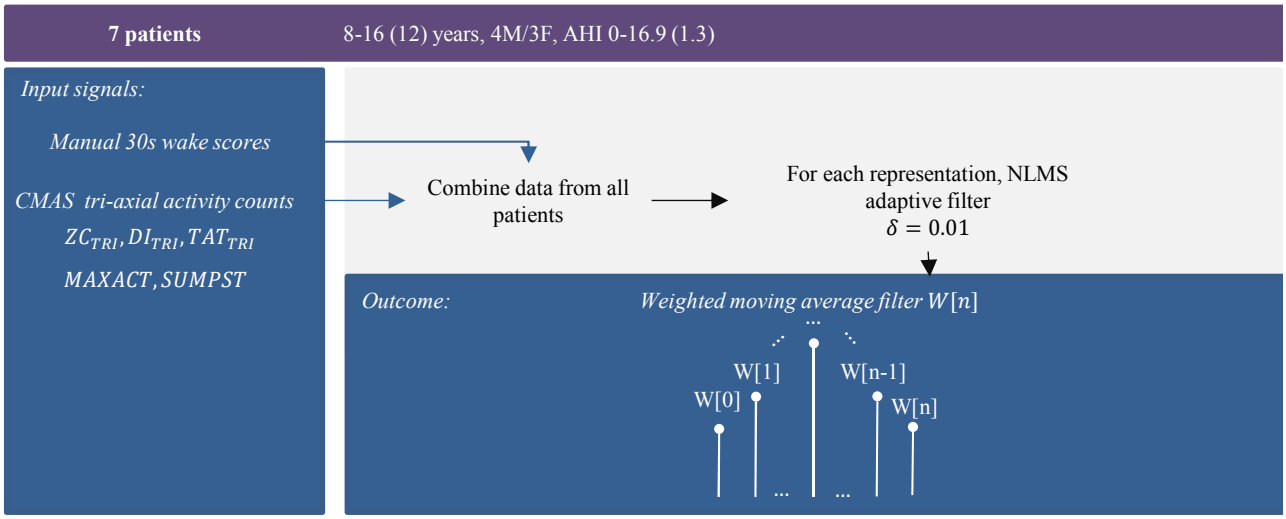


Figure 4.3: Procedure for generating the weighted average filters for the tri-axial activity counts using the adaptive normalised least mean squares (NLMS) filter approach.

The second novel method, SUMPST, is an approximation of the total postural change within an epoch. SUMPST represents limb postural changes and slow motion by approximating the angular displacement of the limb between consecutive samples, relative to gravity. SUMPST is derived using the low-pass filtered data $\mathbf{a}_L$,

$$SUMPST = \sum_{n=1}^{N \cdot f_s} \left| \arccos\left(\frac{\mathbf{a}_L[n]\mathbf{a}_L[n+1]}{||\mathbf{a}_L[n]|| \cdot ||\mathbf{a}_L[n+1]||} \right) \right|. \tag{4.5}$$

Weighted moving average filter development for the tri-axial representations

The Actiwatch uses a weighted moving average filter to combine activity counts between consecutive epochs [123] (see Section 2.3.1 for a detailed description of the algorithms). For consistency with the literature, this weighted moving average filter was also applied to TAT_{UNI}, ZC_{UNI} and DI_{UNI}. To ensure consistent comparisons of all representations, weighted moving average filters were developed for the tri-axial activity counts. These filters were developed using a NLMS adaptive filter. In this application, the NLMS filter creates a weighted function that attempts to maximise the agreement between the activity counts and the manually scored sleep and wake stages from polysomnography for a training patient set. A separate set of patients (detailed in 'Patient characteristics' above) that had missing Actiwatch data were used to determine the filter size and corresponding coefficients for the tri-axial activity counts. The process is shown in Fig 4.3. A filter was generated for varying filter lengths, ranging from $3-25$ epochs. The area under the receiver operating characteristics curve (AUC) was then generated for each representation and filter size for all patients. The filter length that gave the greatest AUC was selected as the appropriate size (illustrated by the larger data points in Fig 4.4). The resulting filters are shown in Fig 4.5.

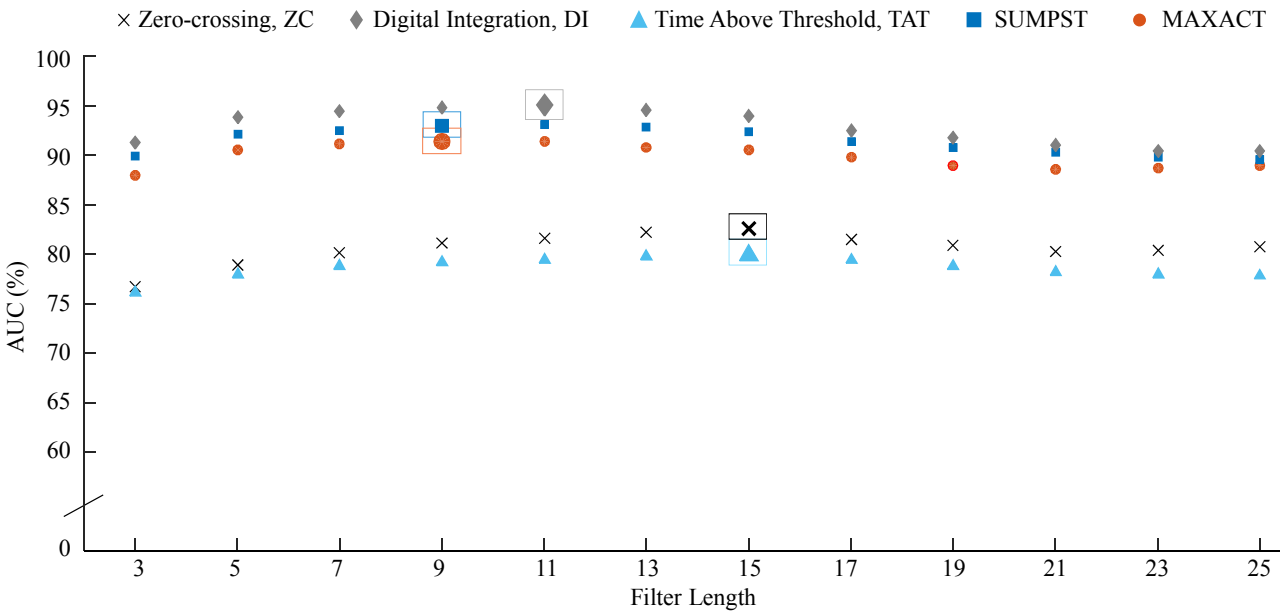


Figure 4.4: AUC for different filter sizes for the tri-axial activity counts.

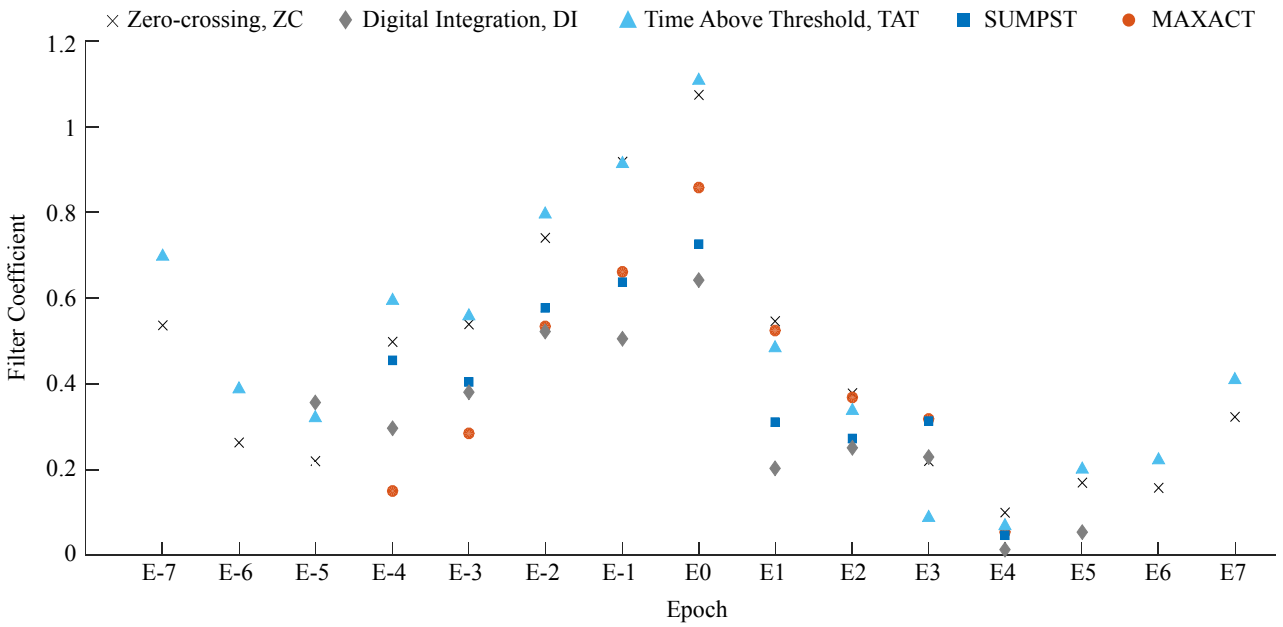


Figure 4.5: Resulting weighted average filters for tri-axial activity counts.

Validation procedure

A threshold classifier was used to analyse the sleep and wake predictive performance of the uni-axial and tri-axial activity counts in a leave-one-out cross-validation design on patients. The threshold was defined as the operating point in the receiver operating characteristics (ROC) analysis that gave the maximum Kappa agreement for the training set in the cross-validation. The predictive performance of the test set at each fold of the cross-validation was quantified using the outcome measures outlined in Section 3.4 in Chapter 3 and documented in Fig 4.2.

To ensure that a meaningful performance comparison between the Actiwatch and CMAS can be made, the standard thresholds for the commercial Actiwatch were not used in this analysis. The threshold for the Actiwatch was defined using the ROC analysis described above. Two-way analysis of variance (ANOVA) was used to assess the statistical significance where metrics exhibited equal variance (as defined by the Brown-Forsythe test). Post-hoc Tukey's honest significant difference test (HSD) test was then used to compare the means of any significant differences as found by the ANOVA. The non-parametric Welch's t-test was used where unequal variance was observed.

4.1.2 Results

Table 4.1 summarises the performance metrics for the uni-axial and tri-axial activity counts. The sleep and wake predictive performance improved significantly when moving from uni-axial to tri-axial activity counts (κ of 0.402 vs. 0.268 for tri-axial vs. uni-axial accelerometry respectively, $p < 0.05$). The discrimination ability (i.e. AUC) was also improved from 81.5% to 86.2%, as illustrated by the greater area under curve for the tri-axial activity counts in Fig. 4.6a. Within the tri-axial representations, the novel representations tended to give better discrimination and predictive performance than ZC, DI or TAT. Table 4.2 shows that there were no significant difference in the performance metrics between CMAS and the Actiwatch.

Table 4.1: Class discrimination ability and predictive performance of the conventional activity counts for uni-axial and tri-axial wrist movement

Activity Count Method	**Sp [%] (at 85% Se)**	**AUC [%]**	**Agreement [%]**	κ
Zero-Crossing, ZC				
Uni-axial	68.5 (27.5)	81.5 (14.6)	80.8 (11.8)	0.268 (0.210)
Tri-axial	73.3 (23.8)	86.2 (9.6)	85.1 (10.3)	0.402 (0.141)[a]
Time Above Threshold, TAT				
Uni-axial	72.5 (26.4)	84.1 (11.6)	83.9 (8.3)	0.405 (0.239)
Tri-axial	72.3 (22.1)	85.7 (10.5)	84.1 (9.8)	0.393 (0.158)
Digital Integration, DI				
Uni-axial	71.3 (27.0)	81.4 (11.8)	83.2 (8.5)	0.359 (0.169)
Tri-axial	74.8 (16.9)	82.4 (10.7)	85.6 (11.9)	0.361 (0.332)
Other tri-axial methods				
Maximum Acceleration, MAXACT	72.3 (28.9)	85.2 (11.4)	82.2 (6.8)	0.422 (0.181)[a]
Total Posture Change, SUMPST	74.8 (20.8)	83.5 (15.3)	85.6 (9.5)	0.439 (0.290)[a]

Values are shown as median (IQR)
[a] $p < 0.05$, greater than ZC uni-axial

Table 4.2: Class discrimination ability and predictive performance of the conventional activity counts for uni-axial and tri-axial wrist movement, comparison with the Actiwatch

Activity Count Method	**Sp [%] (at 85% Se)**	**AUC [%]**	**Agreement [%]**	κ
Actiwatch, ZC	69.1 (19.6)	85.8 (13.5)	85.0 (19.8)	0.445 (0.276)
Uni-axial				
Zero-Crossing, ZC	67.8 (19.6)	81.5 (13.5)	82.7 (19.8)	0.272 (0.276)
Time Above Threshold, TAT	70.0 (24.0)	82.5 (12.3)	85.0 (11.0)	0.421 (0.275)
Digital Integration, DI	72.2 (24.1)	80.2 (8.3)	84.4 (14.1)	0.410 (0.177)
Tri-axial				
Zero-Crossing, ZC	69.9 (21.1)	86.4 (9.9)	86.3 (10.0)	0.403 (0.269)
Time Above Threshold, TAT	69.2 (21.5)	85.7 (11.6)	85.6 (10.0)	0.376 (0.269)
Digital Integration, DI	71.5 (15.6)	79.6 (8.8)	85.7 (20.8)	0.328 (0.230)
Maximum Acceleration, MAXACT	71.8 (28.9)	83.4 (11.4)	84.3 (8.6)	0.369 (0.196)
Total Posture Change, SUMPST	72.9 (17.1)	80.6 (10.9)	88.1 (14.4)	0.421 (0.238)

Analysis performed on the 14 patients that had Actiwatch data (detailed in Table 3.1.3)
Values are shown as median (IQR)

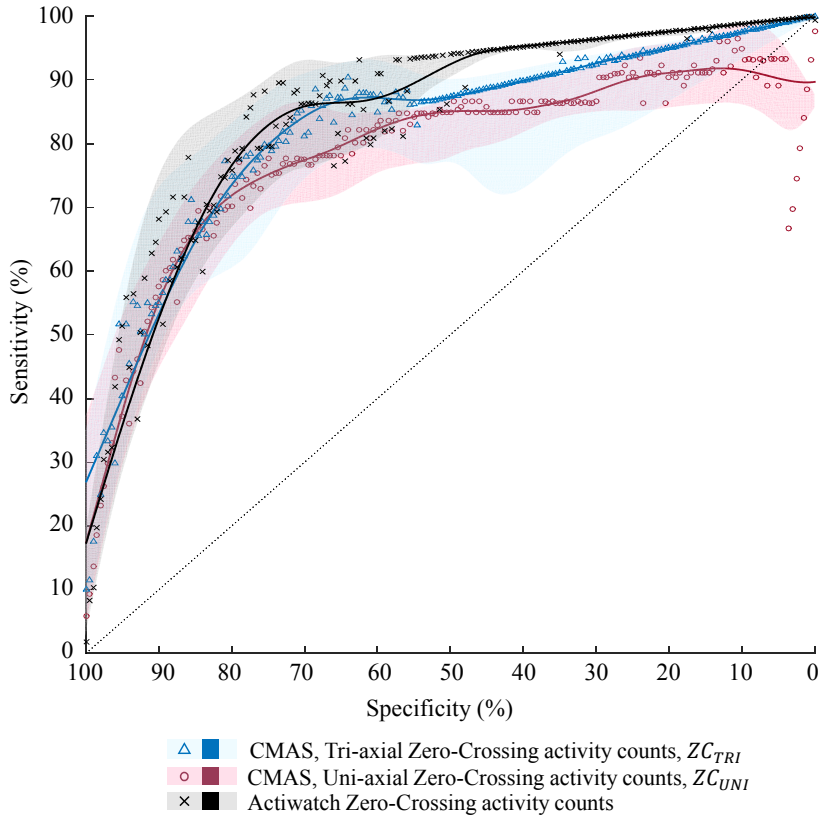
4.1.3 Discussion

The objective of this section was to explore the efficacy of the conventional activity counts for estimating sleep and wake, particularly for children with atypical sleeping behaviour. Although the tri-axial data appeared to improve the predictive performance, it was only statistically significant for one of the representations (i.e. ZC). It is likely that any advantages of tri-axial accelerometry are lost when generating the activity count.

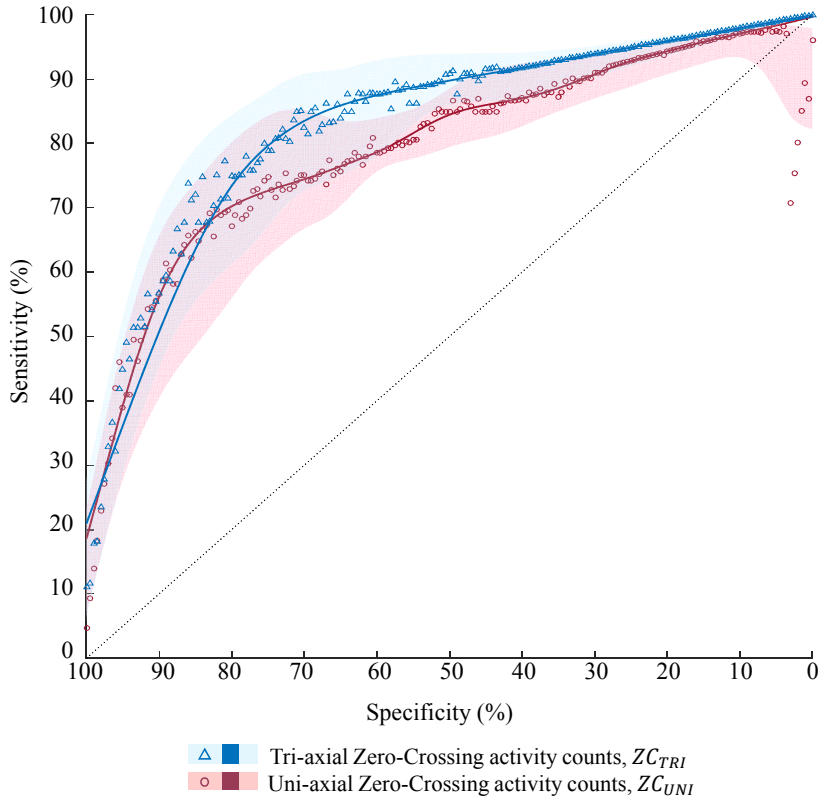
Effectiveness of the conventional activity counts

As shown in Table 4.1, the performance did not significantly differ between the activity count derivation techniques. Any discriminatory movement information is lost when summarising within large epochs. The conventional activity counts do not have adequate resolution for differentiating movement types.

The novel representations, MAXACT and SUMPST, summarise movements that are large in magnitude or result in a positional change. From Table 4.1, these representations had significantly greater predictive performance than ZC, TAT and DI. It is likely then that the maximum acceleration and positional change differ between movements that occur during sleep

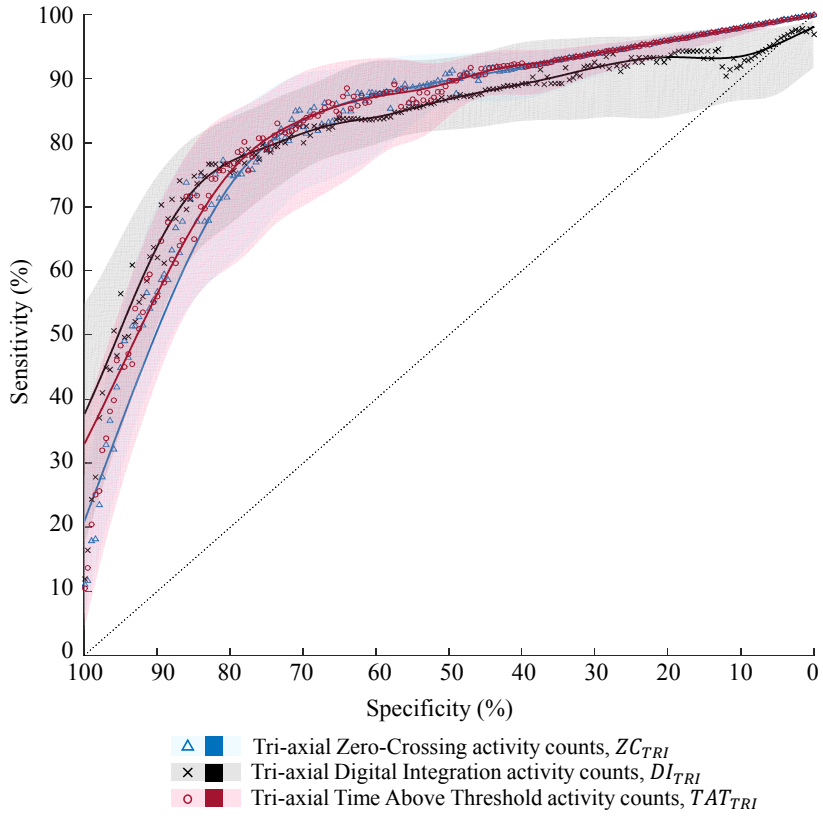


(a)

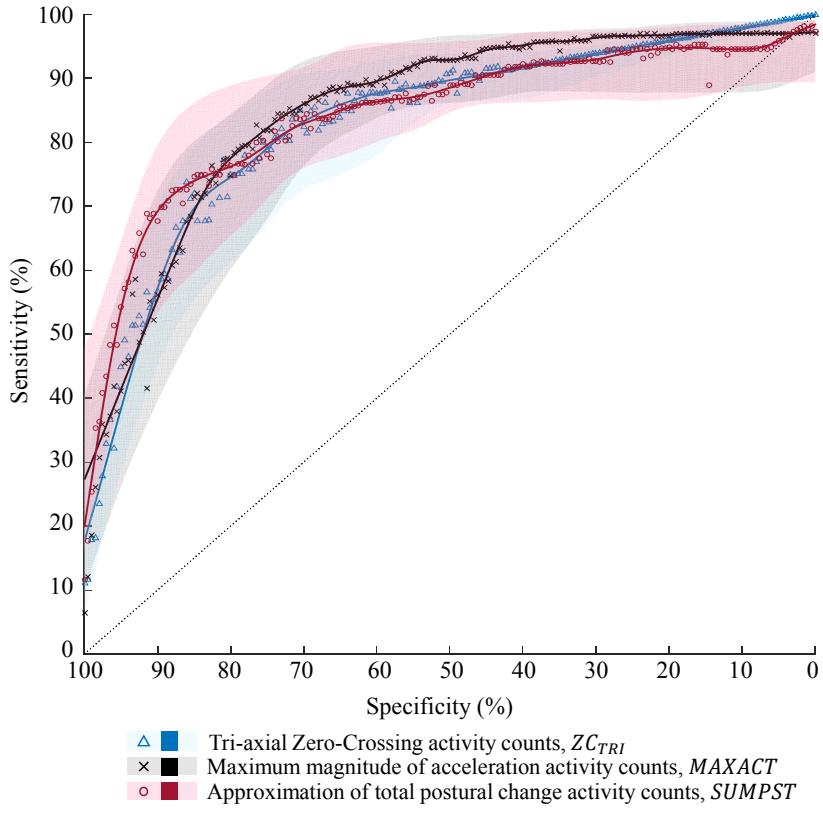


(b)

Figure 4.6: Median ROC curves for the full population used in this analysis for (a) the Actiwatch and comparative CMAS activity counts, and the (b) uni-axial activity counts. Note that the analysis for (a) was performed on a subset of 14 *patients.*



(a)



(b)

Figure 4.7: Median ROC curves for the full population used in this analysis for the (a) conventional and (b) CMAS tri-axial activity counts.

and those that occur during wake. However, as already discussed, any significance is lost when summarising the motion over the 30s epoch.

For consistency, the published Actiwatch weighted moving average filter was applied to the uni-axial activity counts. This filter has been optimised for the Actiwatch and not for the derived uni-axial activity counts in this analysis. This explains the performance disparity between the Actiwatch and CMAS ZC_{uni}.

Information gain from tri-axial accelerometry

Tri-axial data is expected to improve the performance of actigraphy by capturing movement in all three axes, rather than a single axis. In contrast to uni-axial accelerometry that cannot detect movements orthogonal to the measurement axis, capturing movement in all three axes ensures that no movement can go undetected. As expected, the performance of the tri-axial versions of the conventional movement representations were generally superior to the corresponding uni-axial representations. There is evidence then that movement components do occur in the other axes of movement and contribute to detecting sleep and wake. Therefore, all three axes of measurement need to be considered when characterising movement during sleep and wake.

4.2 Accelerometer placements[3]

We have explored uni-axial accelerometry as a limitation of conventional actigraphy for sleep assessment. Another limitation is the use of a single accelerometer. Commercial sleep actigraphy systems typically only measure movement at the wrist. As a consequence of this, wrist movement may be misidentified as whole body movements and movements that occur elsewhere on the body may be missed entirely. This limitation likely contributes to the poor specificity inherent in conventional actigraphy, as some movements during wake may go undetected, resulting in false sleep detections. Therefore, this section will explore the sleep and wake predictive performance when combining data from multiple limbs.

4.2.1 Method

Each patient underwent the study procedure outlined in Section 3.1.4 of Chapter 3. The methodology is summarised in Fig. 4.8 and described in detail below. The data was pre-processed using the procedure outlined in Section 3.2.4 and movement was quantified using each time-series method described in Section 4.1.1 for tri-axial data. Two feature selection techniques were used to select a restricted combination of movement representations and accelerometer placements.

[3]This work has been published in Physiological Measurement: "Multisite accelerometry for sleep and wake classification in children" [144]

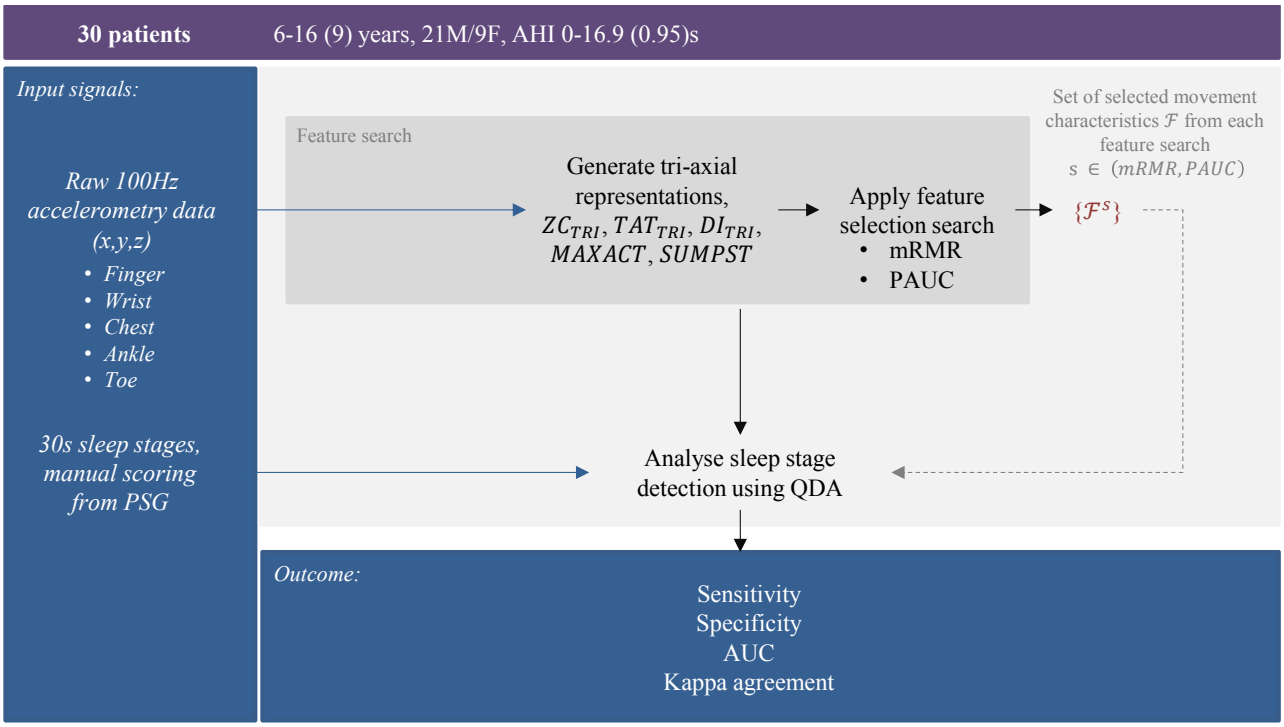


Figure 4.8: Methodology for generating and comparing selected movement representations for two different selection techniques.

Patient characteristics

The full 30 participants detailed in Table 3.1.3 were used in this analysis.

Identifying placement importance

Two well established feature selection techniques were used to judge the reliability and generalisability of the selected movement representations: a sequential forward selection search optimising for partial receiver operating characteristic area under curve (PAUC) [145]; and minimum redundancy maximum relevancy (mRMR) [146]. Forward selection search was used because it is a simple starting point in the feature search. This technique aims to maximise the area under the ROC curve above 60% specificity. Restricting the search in this way ensures that only representations that perform well at a high specificity will be selected. However, it is possible that redundant information can be introduced, which may limit the performance when restricting the number of representations [147]. The second method addresses this by minimising the redundancy in the selected representations, while maximising the relevancy to sleep and wake. The methods are described in more detail below. Each selection method uses leave-one-out cross-validation on patients to select the final set of movement representations.

Sequential forward search with partial AUC

Sequential forward selection search performs a greedy search for representations that improve a performance metric [148]. The search finishes when the metric is no longer improved by

additional representations. The performance metric used in this analysis was the PAUC, which attempts to maximise the sleep/wake discrimination ability while ensuring that the specificity cannot fall below a set value. In this analysis, PAUC is defined as the AUC above a specificity of 60%. AUC represents the ability of a set of representations to rank a randomly chosen positive instance (wake) higher than a randomly chosen negative instance (sleep) [130]. Therefore, maximising AUC above a specificity of 60% ensures that only a set of representations that have good discriminatory power at a relatively high specificity will be chosen. Similar selection techniques have been used in literature. Thiemjarus et al. [149] perform feature and sensor reduction using a Bayesian framework for detecting activities with accelerometers placed on many areas on the body. They rank features based on the cumulative AUC when combining multiple accelerometers. Their search is stopped when the AUC is no longer increasing. It is unclear if this method takes into account redundancy. The selected features and/or placements may also differ with a different starting point. The sequential forward search specifically attempts to improve the low specificity of sleep actigraphy systems. However, it is possible that redundant representations can be selected. Reducing redundancy in the selected representations can improve how well they generalise [150, 147]. This is addressed by the next method.

Minimum redundancy maximum relevancy

The second selection approach uses mutual information to determine the relevancy and redundancy of each representation with respect to sleep and wake [151]. Minimum redundancy maximum relevancy is a supervised selection technique that seeks to minimise the redundancy between representations, while maximising the relevancy to the class label [146]. This method avoids the case where features are highly relevant to the class label, but redundant, resulting in a larger subset than necessary [152]. mRMR has been validated in a number of studies ranging from gene expressions [153] to drug interactions [154]. mRMR is particularly suited to analysing movements with accelerometry, since it is likely that some accelerometer locations or movement representations will result in redundancy. A similar method to Peng et al. [151] is used here to define and optimise the redundancy and relevancy of each representation. The redundancy Ru of a set of representations is defined as the mean mutual information I between each representation f_n within that set s,

$$Ru = \frac{1}{|s|^2} \sum_{f_1, f_2 \in s} I(f_1; f_2), \tag{4.6}$$

where mutual information is defined by Hutter [155],

$$I(s; y) = \sum_{y \in Y} \sum_{f \in s} \log \left(\frac{p(f, y)}{p(f) \cdot p(y)} \right), \tag{4.7}$$

where $p(f, y)$ is the joint probability distribution function of the class y and the feature f,

and $p(f)$ and $p(y)$ are the marginal probability distribution functions of the feature and the class respectively. Mutual information quantifies both the contribution of each representation to wake prediction and the similarity between each representation. Similar to the redundancy, the relevancy Re is defined as the mean mutual information I between each representation f within a set s and the class label c,

$$Re = \frac{1}{|s|} \sum_{f \in s} I(f; c). \tag{4.8}$$

A simple approach to optimise for maximum relevancy and minimum redundancy is to maximise the difference between the two metrics,

$$mRMR = \max(Re - Ru). \tag{4.9}$$

Validation procedure

The sleep and wake predictive performance of the selected representations was determined using quadratic discriminant analysis (QDA), with a binomial distribution in a leave-one-out cross-validation design on patients. QDA was used because the variability of the accelerometry data for sleep and wake differs, which violates the assumptions for a linear discriminant [156]. The probability of the predicted sleep stages given by the discriminant analysis was compared to the actual sleep stages using ROC curve analysis. The ability of the predictor to discriminate between sleep and wake was summarised by the AUC. The predictive performance at each fold of the cross-validation was quantified using standard and Kappa agreement κ with polysomnography [131], where wake was defined as a predictive probability greater than the threshold that gave the maximum Kappa agreement in the ROC analysis. Specificity at 85% sensitivity was also reported. The Wilcoxon rank sum test was used to assess the significance of the class prediction from the QDA because the performance distributions were not normally distributed (as defined by the Kolmogorov-Smirnov test).

4.2.2 Results

The ROC analysis and agreement with polysomnography are summarised for the best N representations of each of the methods in Table 4.3 and for the commercial actigraph and CMAS in Table 4.4. Similar to Section 4.1, tri-axial accelerometry improved the performance when compared to uni-axial accelerometry (specificity at 85% sensitivity: 72.0(22.9)% vs. 66.1(32.8)%, $p < 0.05$). These values differ slightly from Section 4.1 as the full 30 patients were used here, whereas Section 4.1 only used 23. Combining data from the accelerometers significantly improved the ranking performance and agreement when compared to the single wrist accelerometer (AUC for five characteristics: 92.2(9.5)% vs. 84.6(13.6)%, $p < 0.05$; Kappa agreement:

Table 4.3: Class discrimination ability and predictive performance for the N best representations of each selection method.

N	Selected Characteristics	Sp [%] (at 85% Se)	AUC [%]	Agreement [%]	κ
	Uni-axial, Wrist_{ZC}	66.1 (32.8)	79.9 (11.3)	84.4 (9.8)	0.440 (0.258)
1	Tri-axial, Wrist_{ZC}	72.0 (22.9)^{a}	84.6 (13.6)	85.0 (10.3)	0.488 (0.257)
Multivariate Locations, Forward Selection with PAUC					
2	Wrist_{ZC}, Toe_{ZC}	78.1 (29.6)^{b}	88.2 (11.3)^{b}	87.2 (12.5)	0.548 (0.230)
3	Wrist_{ZC}, Toe_{ZC}, Toe_{TAT}	78.8 (30.2)^{b}	87.9 (11.4)^{b}	86.6 (11.7)	0.551 (0.244)
$\leq$5	Wrist_{ZC}, Toe_{ZC}, Toe_{TAT}, Ankle_{DI}, $\text{Toe}_{\text{MAXACT}}$	83.5 (17.4)^{bc}	92.2 (9.5)^{bc}	87.2 (12.2)	0.565 (0.231)^{a}
Multivariate Locations, mRMR					
2	Wrist_{ZC}, Toe_{ZC}	78.1 (29.6)^{b}	88.2 (11.3)^{b}	87.2 (12.5)	0.548 (0.230)
3	Wrist_{ZC}, Toe_{ZC}, $\text{Finger}_{\text{DI}}$	76.3 (26.5)^{a}	87.2 (12.8)^{b}	88.2 (12.8)	0.600 (0.209)^{a}
$\leq$5	Wrist_{ZC}, Toe_{ZC}, $\text{Finger}_{\text{DI}}$, Toe_{DI}, $\text{Wrist}_{\text{TAT}}$	81.9 (22.2)^{b}	90.0 (10.1)^{b}	86.4 (13.1)	0.545 (0.218)^{a}

Values are shown as median (IQR)
^{a} $p < 0.05$, ^{b} $p < 0.01$, greater than uni-axial wrist placement
^{c} $p < 0.05$ greater than single tri-axial wrist placement

Table 4.4: Class discrimination ability and predictive performance of the Actiwatch and the CMAS uni-axial and tri-axial wrist accelerometers.

Selected Characteristics	Sp [%] (at 85% Se)	AUC [%]	Agreement [%]	κ
Actiwatch	78.2 (14.0)	86.4 (9.2)	88.0 (7.2)	0.464 (0.429)
Uni-axial, Wrist_{ZC}	52.0 (30.1)	80.0 (11.2)	84.6 (11.7)	0.471 (0.268)
Tri-axial, Wrist_{ZC}	73.7 (19.2)	85.6 (7.6)	87.3 (12.6)	0.461 (0.271)

Commercial actigraphy system: Actiwatch Mini, CamNTech, using the zero-crossing mode.
Analysis performed on the 14 patients that had Actiwatch data (detailed in Table 3.1.3)
Values are shown as median (IQR)

0.565(0.231) vs. 0.488(0.257), $p < 0.05$). This trend in improvement was consistent across all metrics. The selection algorithms (mRMR and forward selection search) selected similar characteristics and yielded consistent performance metrics. We can therefore be confident that these results are not an artefact of a specific selection algorithm.

4.2.3 Discussion

The objective of this section was to determine if combining data from multiple accelerometers improved the ability to detect sleep and wake in children. Combining accelerometry data improved the sleep and wake discrimination performance and agreement with polysomnography. In this analysis, the predictive performance of the single wrist accelerometer differed slightly to the performance seen in Section 4.1. This is likely due to the chosen operating point from the different representations in the different analyses: Section 4.1 identified the operating point from the raw activity counts; and Section 4.2 identified the operating point from the probability output of QDA with the raw activity counts, which will have additional constant, scalar and/or quadratic terms.

Impact of additional accelerometers

Combining data from multiple accelerometers showed a trending performance improvement, which was significant for the Kappa agreement and sleep and wake discrimination ability (an example of comparative ROC curves is illustrated in Fig. 4.9). This improvement suggests that

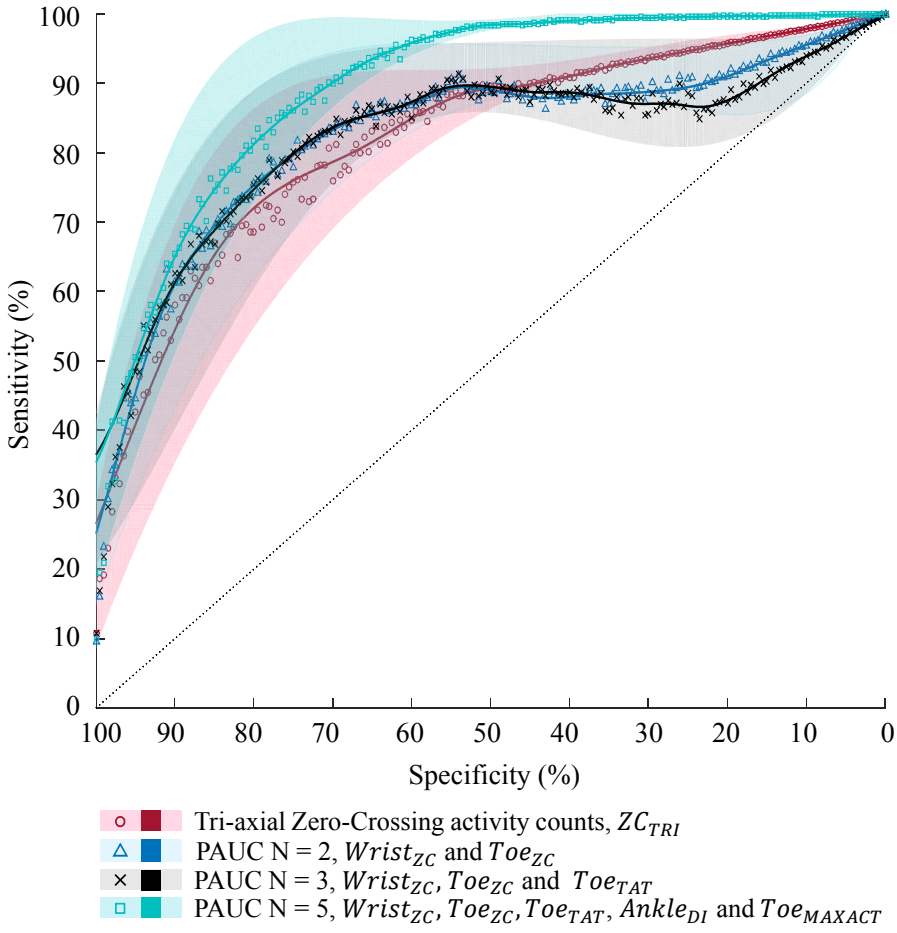


Figure 4.9: Median ROC curves for the full population used in this analysis for multi-site activity counts.

combining movement information from multiple accelerometers improves the ability to detect sleep movements. As evidenced by the improvement in performance when incorporating additional accelerometers, the wrist accelerometer alone is not sufficient for accurately measuring movement. This improvement in performance also suggests that there are some physiologically different movements between sleep and wake that are captured using multiple accelerometers. Although there is a slight performance benefit to measuring movement with multiple accelerometers, additional accelerometers negatively impacts the usability of the system in a non-clinical environment: system complexity, cost and patient discomfort are all impacted by incorporating additional accelerometers into the actigraphy hardware framework.

Significance of selected representations and accelerometer placements

The wrist placement was consistently selected first, suggesting that the wrist accelerometer captures more wake movements than the other accelerometers. Although we can observe some improvement when combining data from additional accelerometers, it was only significant with five features. Supporting the findings in Section 4.1, this suggests that the conventional activity counts do not adequately summarise movement, and that they can be improved by including activity from other representations (for example, including ZC and TAT activity counts for toe movement in Table 4.3). Analysing the relationships between the representations showed that, in general, summaries of total movement within an epoch (TAT, DI, SUMPST) were complementary to those that summarise abrupt movements (ZC, MAXACT) and were often selected together. Conversely, representations that summarise similar movements at the same location on the body were seldom selected together.

Despite being highly relevant to wake, there was considerable redundancy found between wrist and finger movement, suggesting that the finger and wrist accelerometers may be interchangeable. Although the finger placement is commonly used for devices such as pulse oximeters, it is considered unfavourable for children because of comfort dependencies (e.g. thumb-sucking or self-soothing) and it may distract from sleeping [157]. Information from the ankle and toe accelerometers were also found to be redundant. Interestingly, the upper thorax was seldom selected in the search, indicating that the chest offers little additional useful information. Indeed, this agrees with previous literature that has shown that 90% of variation in chest movement is summarised by the ankle (or toe) movement [158], which was often selected.

4.3 Summary

In this chapter we explored techniques to reduce the number of false sleep detections. We hypothesised that these false sleep detections are caused by inaccurate movement representations; some wake movements cannot be detected with the conventional actigraphy framework. Therefore, in this chapter we explored tri-axial multisite accelerometry. In Section 4.1 we explored

sleep and wake classification performance using movement information from tri-axial summary techniques. We then incorporated tri-axial representations of movement measured from multiple accelerometers on the body in Section 4.2 and assessed the sleep and wake predictive performance.

The analysis performed in Section 4.1 shows that tri-axial accelerometry does improve predictive performance when compared to uni-axial accelerometry. We identified that this is likely due to improved movement detection from monitoring additional axes. The analyses in Section 4.1 and Section 4.2 highlighted a significant limitation of the conventional activity counts; activity counts are unable to provide adequate resolution to detect specific characteristics of movements, which we saw is required for improving sleep estimates. It is likely then that identifying these characteristics, and detecting sleep and wake on a movement-by-movement basis (instead of the conventional epoch-by-epoch summaries), may improve the sleep estimates.

Incorporating data from accelerometers at different locations on the body in Section 4.2 improved the discrimination ability and predictive performance when compared to the conventional wrist placement. However, including additional accelerometers in the actigraphy routine increases the cost, system complexity and patient discomfort, all of which we are trying to minimise; the limitations of additional accelerometers impact the usability of actigraphy in a non-clinical environment, and may impact the measured sleep behaviour. For these reasons, the performance benefits of multisite accelerometry will need to be considered against the practical limitations.

In this chapter we conclude that:

- Tri-axial accelerometry significantly improves sleep and wake detection when compared to uni-axial accelerometry;
- Incorporating additional accelerometers at different locations on the body improves discrimination ability and predictive performance for detecting wake; and
- The conventional activity counts are unable to accurately quantify differing characteristics between movements associated with sleep and those associated with wake, as epoch-by-epoch quantification does not provide adequate resolution.

In the next chapter, we will address the second limitation of actigraphy: false wake detections caused by movements during sleep. We will isolate and remove the effect of some of these sleep movements, and identify differentiable characteristics between sleep and wake movements. We will move away from representing movement as a summarised activity count by and explore sleep and wake detection on a movement-by-movement basis. In Chapter 6 we will identify the association between sleep movements and pathological and physiological events related to sleep disorders.

5

Differentiating sleep and wake movements

"... sleep ... is a seizure of the primary sense-organ, rendering it unable to actualize its powers; arising of necessity (for it is impossible for an animal to exist if the conditions which render it an animal be not fulfilled), i.e. for the sake of its conservation; since remission of movement tends to the conservation of animals."

— Aristotle, *Philosopher, 384 - 322 BC*

Since the early 1900s, movement has been considered a method of detecting sleep [159]; sleep onset was measured at the point of complete muscle control loss as visually observed. Although naïve, this theory has persisted and evolved to include specific movements during the different sleep stages. These movements include the sharp limb twitches that are often observed during rapid eye movement (REM) sleep [160], movements caused by parasomnias during non-rapid eye movement (NREM) sleep [19], or general sleep positional changes. Since the conventional actigraphy classification framework identifies 'wake' as regions of increased activity, these movements can confound wake predictions if they correspond with an activity reading that is above the classification threshold: as illustrated in Fig. 5.1, high-activity epochs during sleep will result in false wake detections. Differentiating sleep movements from wake movements would consequently improve this limitation. However, a distinct limitation of representing movement as activity counts that we observed in Chapter 4 was that the temporal resolution of activity counts can cause characteristically different movements to have the same

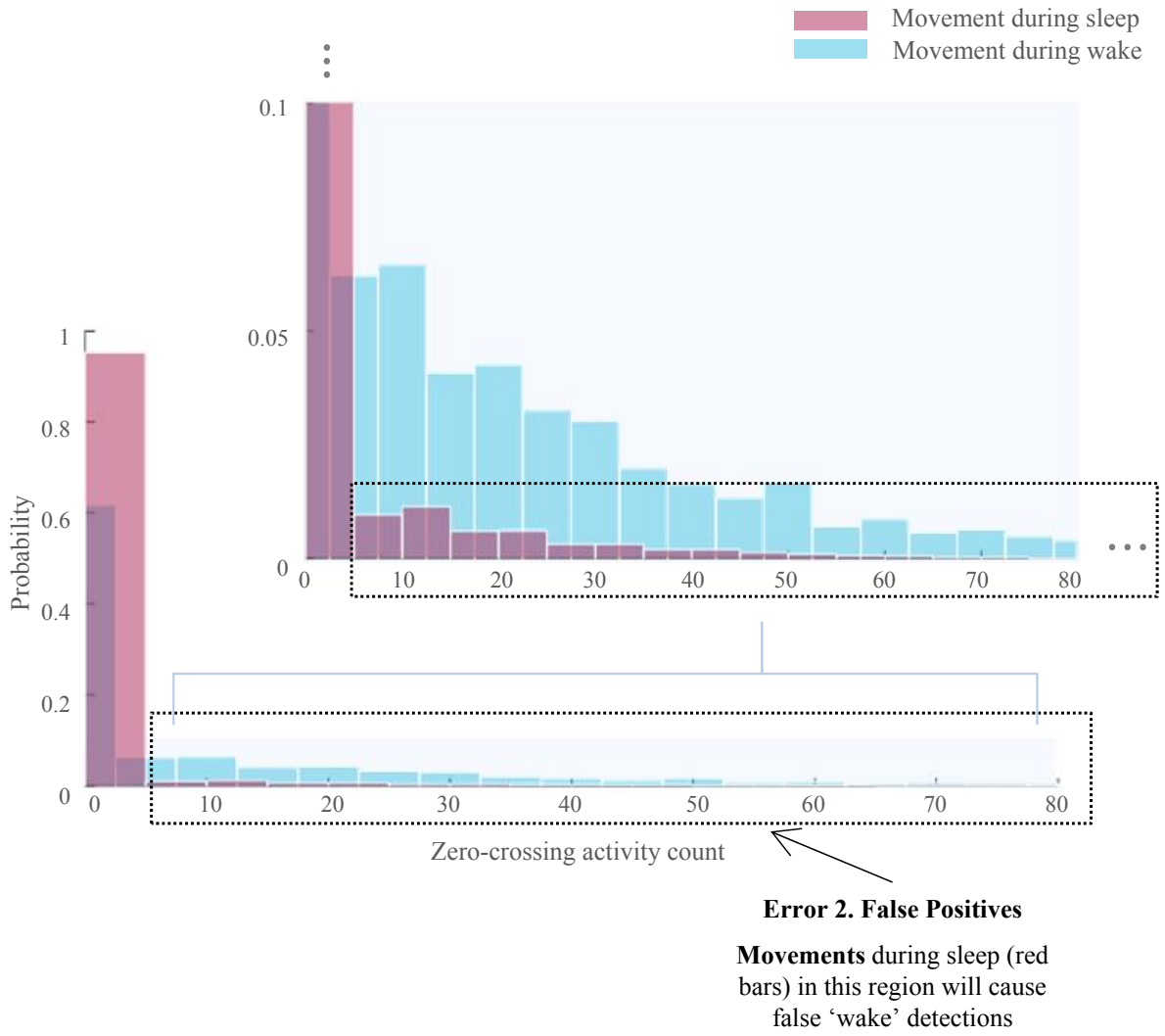


Figure 5.1: Histogram of activity during sleep (red) and wake (blue) for 24 patients using the conventional activity counts derived using the zero-crossing method. Highlighted region shows the source of the false positives.

numerical activity count value; i.e. multiple short-duration movements will appear identical to a single lengthy movement within an epoch. Therefore, the accuracy of actigraphy-based wake estimates is dependent on accurately differentiating these active sleep epochs from active wake epochs. In this chapter we seek to reduce false wake detections by identifying if movement characteristics during wake are characteristically different to those during sleep. We will first explore heuristic identification of movements specific to sleep as a pre-processing step on the raw accelerometry data. We will then explore specific characteristics that differ between sleep and wake movements, segmented from high-resolution raw accelerometry data. Finally, we will compare sleep and wake predictions on a movement-by-movement basis to the conventional fixed epochs.

The brain deliberately inhibits motor activity during sleep, reducing muscle tone and impacting the nature of sleep movements. As a consequence of this atonia, we can speculate

that the physiological characteristics of sleep movements differ to the deliberate movements of wake. The body of literature for sleep movement exploration has focused on identifying movements during sleep that are associated with specific disorders (such as bruxism [161] and epilepsy [162]); there have been few attempts to specifically differentiate restless sleep (i.e. regions of activity during sleep) from wake. Although some studies have observed that the general prevalence [163, 164, 165] and coherence [18, 17] of movement during sleep differs to wake, specific movement characteristics have not explicitly been explored. As detailed in Section 2.3.1, conventional actigraphy uses time-series methods to represent movement. These representations have distinct limitations which were specifically explored in Chapter 4. We saw that the conventional activity counts have inadequate temporal resolution to identify single movements, and time-series methods cannot capture all movement information. Time-frequency analysis is a useful tool for analysing non-stationary signals like sleep and wake movements because it summarises both spectral and temporal information [166, 167]. Given that the spectral content of non-stationary movement typically varies over time, the combined spectral and temporal information is particularly important for characterising non-periodic movement measured with accelerometry. For this reason, we will focus on identifying localised spectral characteristics that differ between sleep and wake movements. There are two common approaches for analysing this in literature: template matching and wavelet transform co-efficient analysis.

The wavelet transform provides an overview of the similarity of the signal with a wavelet or basis function [115]. In template matching, prior knowledge of movement patterns are exploited to identify movements by matching the pattern with a basis function. This is a common approach for identifying movements that conform to a certain pattern or behaviour, such as seizures and tremors [112, 168, 169]. Nijsen et al. [168] developed a model to detect motor activity from epileptic patients at night. Using this model, they then identified temporal and spectral characteristics of epileptic movements for accurate event detection [112]. Geman et al. [169] also used template matching to identify tremors for early detection of Parkinson's disease. They compared the coefficients of different basis functions to optimise tremor detection.

These approaches are effective at identifying movements where the characteristics conform to some pattern. Although some movements during sleep may match a template (for example, limb twitches during REM sleep) we can hypothesise that the majority of movements are likely to be stochastic in nature because of the inhibited muscle control [170, 8] and identified incoherent nature [18, 17]. For this reason, pattern matching may not be effective for identifying characteristic differences between movements during sleep and wake. The wavelet transform can detect localised regions of high spectral power for different frequencies. This application is often adopted for identifying periodic activities [21, 116, 115, 171]. Although movements during sleep are not likely to be periodic, there will be localised regions of high spectral power at varying frequencies that the wavelet transform can isolate. Excluding tremors and myoclonic twitches, the literature has not explored spectral characteristics of movements during sleep.

This chapter specifically addresses the second limitation of actigraphy, discussed in Chapter 1 and illustrated in Fig. 5.1:

False positives: sleep epochs with observed movement are incorrectly identified as 'wake'.

This chapter will address the hypothesis:

> *Movement characteristics can differentiate sleep from wake because the physiological nature of these movements differ.*

In this chapter we will manually identify movements specific to sleep and explore the effect of removing these movements as a pre-processing step on the raw accelerometry data. We will then explore localised spectral characteristics of movements that are specific to sleep or wake. Using these characteristics, we will compare the accuracy of predicting wake on a movement-by-movement basis to the conventional epoch-by-epoch summaries.

This exploration will be performed in three sections:

Heuristic removal of restless sleep
Section 5.1 aims to reduce false wake detections by identifying and removing the effect of movements specific to sleep on the conventional activity counts.

Differentiable spectral movement characteristics
Section 5.2 aims to identify differentiable spectral characteristics between sleep and wake movements.

Sleep/wake prediction on a movement-by-movement basis
Section 5.3 aims to explore the efficacy of predicting sleep and wake regions on a movement-by-movement basis.

5.1 Heuristic removal of restless sleep[1]

A significant limitation of the conventional actigraphy framework is the inability to accurately differentiate regions of activity during sleep from wake. As we observed in Section 4.1, this limitation may be due to the poor temporal resolution of activity counts. Since activity counts are restricted to large epochs (commonly 30s), activity counts are unable to determine specific characteristics or types of individual movements; activity counts can only indicate the occurrence of movement, or provide a measure of the intensity of movement within the epoch. For this reason, activity counts are unable to isolate a single movement, as movements are likely to last for a fraction of 30s (particularly for sleep where myoclonic and incoherent movements are common [22, 17]). Literature has focussed on detecting periods of 'activity' and 'inactivity'. However, manually analysing finger, wrist, chest, ankle and toe accelerometry data concurrently with 6-hour video segments of 10 sleeping patients (5 male, aged $6-15$, patients 15, $19-22$, 26 and $30-33$ in Table 3.1.3) using custom software (see Appendix C) showed that there are characteristics of movements that differ between sleep and wake.

The observations of particular note were:

(a) children are likely to completely change their hand positions during sleep;
(b) children are likely to slightly shift their body during sleep;
(c) 'bursty' movements during sleep are likely to be short in duration (< 2s);
(d) movements such as face scratching and full sleeping position changes (e.g. supine to left or right lateral) are likely to occur during both sleep and wake; and
(e) movements associated with wake are likely to have a longer duration than movements during sleep.

Identifying the observed sleep characteristics (a)-(c) may reduce false wake detections because these characteristics are specific to sleep. Therefore, this section heuristically identifies and removes the influence of these characteristics as a pre-processing step on the raw accelerometry data. The standard activity counts (calculated as described in Chapter 4) are then generated and the sleep and wake prediction accuracy is then comparatively assessed, pre- and post-heuristic.

5.1.1 Method

Each patient underwent the study procedure outlined in Section 3.1.4 of Chapter 3. The methodology is summarised in Fig. 5.3 and described in detail below. The heuristic was applied

[1]This work has been published in Physiological Measurement: "Multisite accelerometry for sleep and wake classification in children" [144]

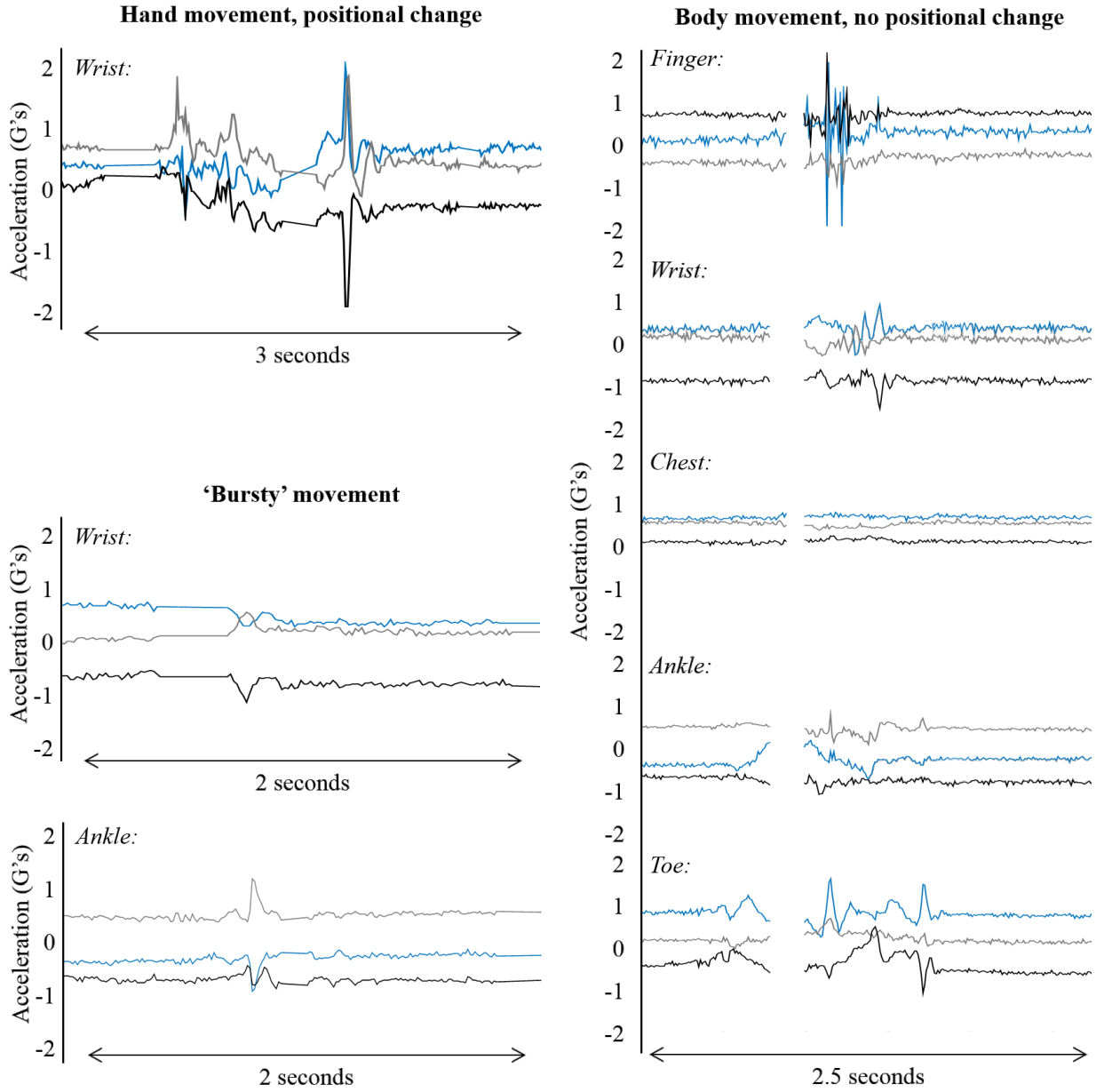


Figure 5.2: Example of the observed movements used to derive the heuristic: hand movement with a positional change, 'bursty' movements and body movements with no positional changes.

to the raw accelerometry data prior to the standard pre-processing procedure outlined in Section 3.2.4. Movement was quantified using each time-series method described in Section 4.1.1 for tri-axial data. The sleep and wake predictive performance of the multisite combination that gave the greatest performance in Section 4.2 was compared prior to, and after, applying the heuristic using QDA. Similar to Section 4.2, QDA was used because the variability of the accelerometry data for sleep and wake differs.

Patient characteristics

The full 30 participants detailed in Table 3.1.3 were used in this analysis.

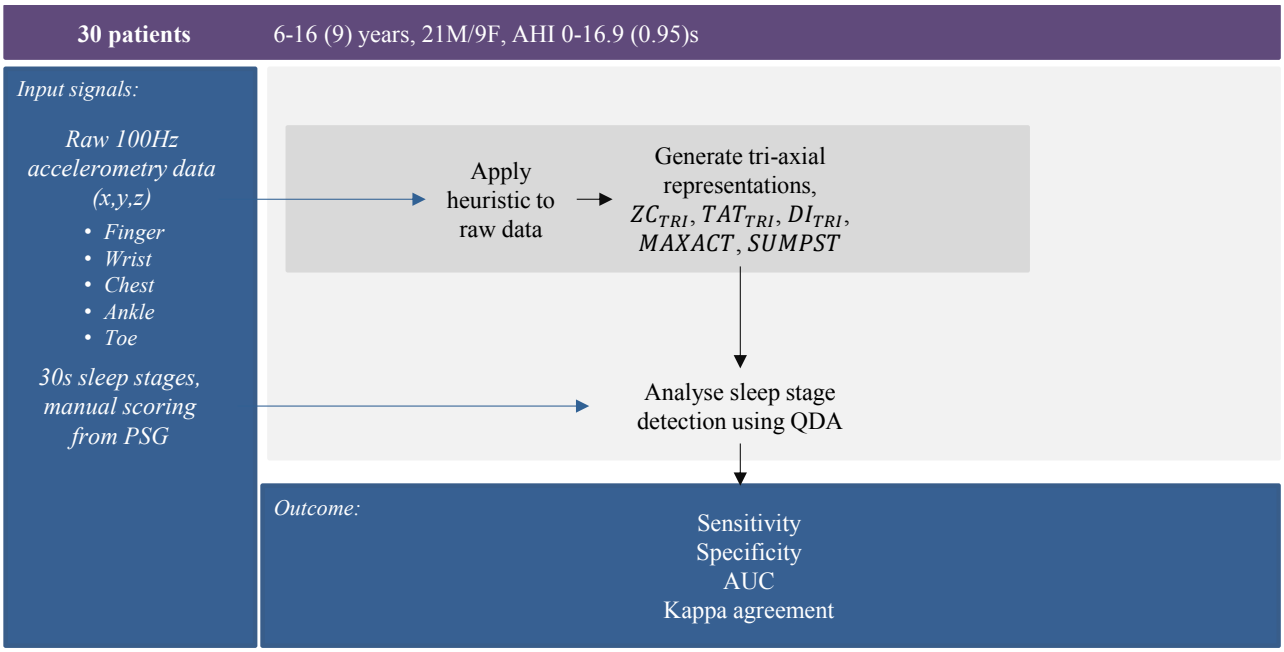


Figure 5.3: Methodology for identifying the effect of heuristically removing movements specific to sleep.

Heuristic procedure

A heuristic was developed to identify and quantify the observed differentiable characteristics (described as (a)-(c) above). This heuristic attempts to negate the influence of the sleep characteristics on wake detection by setting the accelerometry signal to zero when these characteristics are identified. This ensures that the predictor will identify that region as 'no movement' and consequently correctly classify that region as 'sleep'. The characteristics of movements that were observed during sleep are defined technically as:

(a) hand position change: wrist movement with a final DC offset;

(b) body shift: wrist, ankle and chest movement with no chest DC offset; and

(c) 'bursty' movement: any movement of less than 2s in duration.

The procedure for negating the effects of these movements is described in Fig. 5.4 and the algorithm is detailed in Appendix A. Movements are detected from the raw accelerometry data

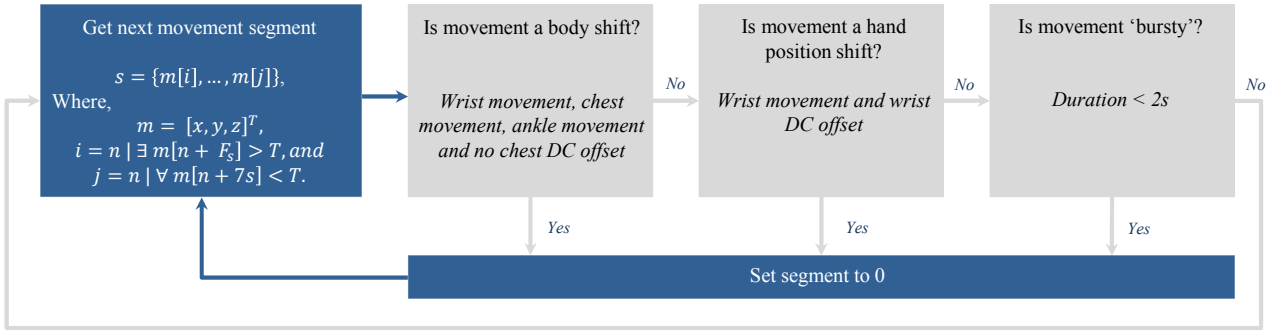


Figure 5.4: Procedure for removing movements associated with sleep from the raw accelerometry data.

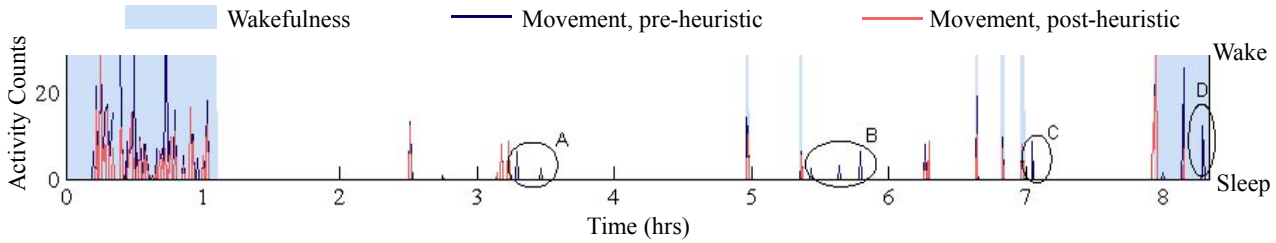


Figure 5.5: Example of the effect of applying the heuristic to the raw accelerometry data, prior to generating the features. Sleep (defined as REM and all non-REM stages) and wake (shaded) periods are shown. Movement is represented by wrist ZC_{TRI}.

as any data point above the noise floor. Since the magnitude of the noise is consistent across the raw signal, the noise floor is defined as the median of the absolute magnitude of the high-pass filtered accelerometry data across each full study. The duration of movement is defined as the start of movement until there is no movement for a minimum of 7s. This is to ensure that movements with transitional properties are detected as a single movement, which is particularly important for movements such as sleeping position changes (e.g. supine to left or right lateral).

An example of the effect of removing differentiable sleep movements is illustrated in Fig. 5.5. The heuristic removed movement associated with restless sleep (shown in regions A, B and C). The heuristic also removed some movements associated with wake (shown in region D) because there is a likelihood that some wake movements may appear similar to movements characteristic of restless sleep (as defined by the heuristic). We can see from Fig. 5.5 that removing sleep characteristics (a)-(c) reduces the number of large activity counts during sleep, without significantly affecting the activity counts during wake. After applying the heuristic to the raw accelerometry data, the conventional time-series features were derived for each accelerometer. Since the heuristic required the assessment of movements from each accelerometer, the best multisite feature set selected from Section 4.2 was then used to classify sleep and wake using QDA. The performance was then compared to the feature set with data that was not pre-processed with the heuristic.

Leave-one-out cross-validation on patients was used to assess the performance of the conventional wrist placement and the feature set prior- and post-heuristic removal of restless sleep movements. The predictive performance at each fold of the cross-validation was quantified using standard and Kappa agreement κ with polysomnography [131], where wake was defined as a predictive probability greater than the threshold that gave the maximum Kappa agreement in the ROC analysis. The Wilcoxon rank sum test was used to assess the significance of the class prediction from the QDA because the performance distributions were not normally distributed (as defined by the Kolmogorov-Smirnov test). AUC and the specificity at 85% sensitivity was also reported.

Table 5.1: Class discrimination ability and predictive performance for tri-axial multisite accelerometry, pre- and post-removal of movement characteristics specific to sleep.

Selected Characteristics	**Sp [%] (at 85% Se)**	**AUC [%]**	**Agreement [%]**	κ
Single accelerometer, $\mathbf{Wrist_{ZC}}$				
Pre-heuristic	72.0 (22.9)	84.6 (13.6)	85.0 (10.3)	0.488 (0.257)
Post-heuristic	70.2 (28.1)	85.1 (14.0)	87.1 (12.8)	0.557 (0.261)
Multiple accelerometers, $\mathbf{Wrist_{ZC}}$, $\mathbf{Toe_{ZC}}$, $\mathbf{Toe_{TAT}}$, $\mathbf{Ankle_{DI}}$, $\mathbf{Toe_{MAXACT}}$				
Pre-heuristic	83.5 (17.4)[a]	92.2 (9.5)[a]	87.2 (12.2)	0.565 (0.231)
Post-heuristic	85.6 (16.3)[b]	93.2 (6.6)[b]	90.6 (10.8)	0.630 (0.292)[b]

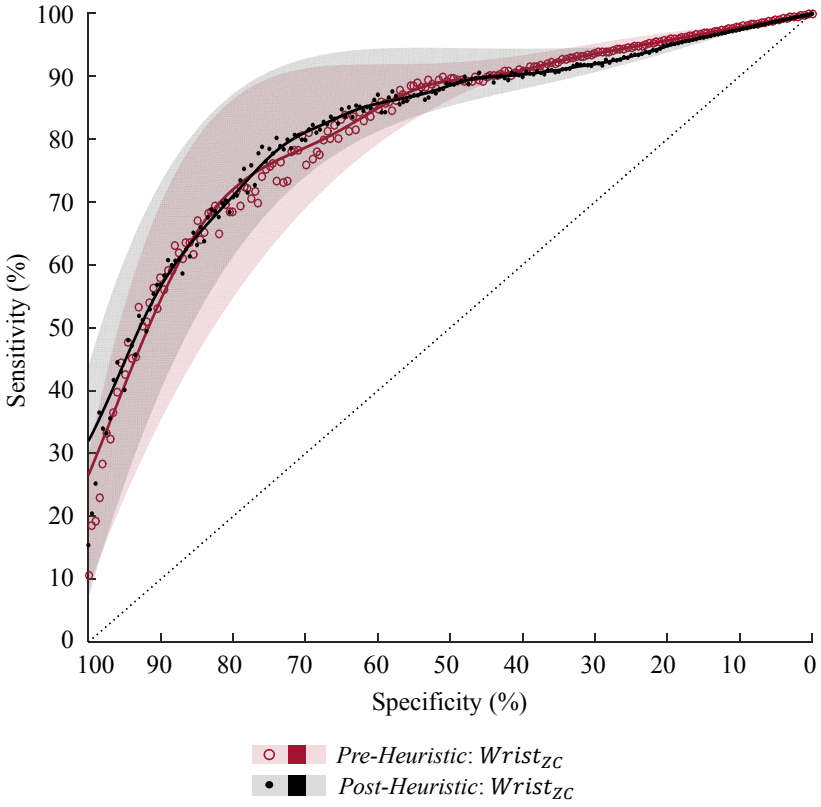
Values are shown as median (IQR)
[a] $p < 0.05$, [b] $p < 0.01$, greater than standard tri-axial wrist placement (no heuristic pre-processing)

5.1.2 Results

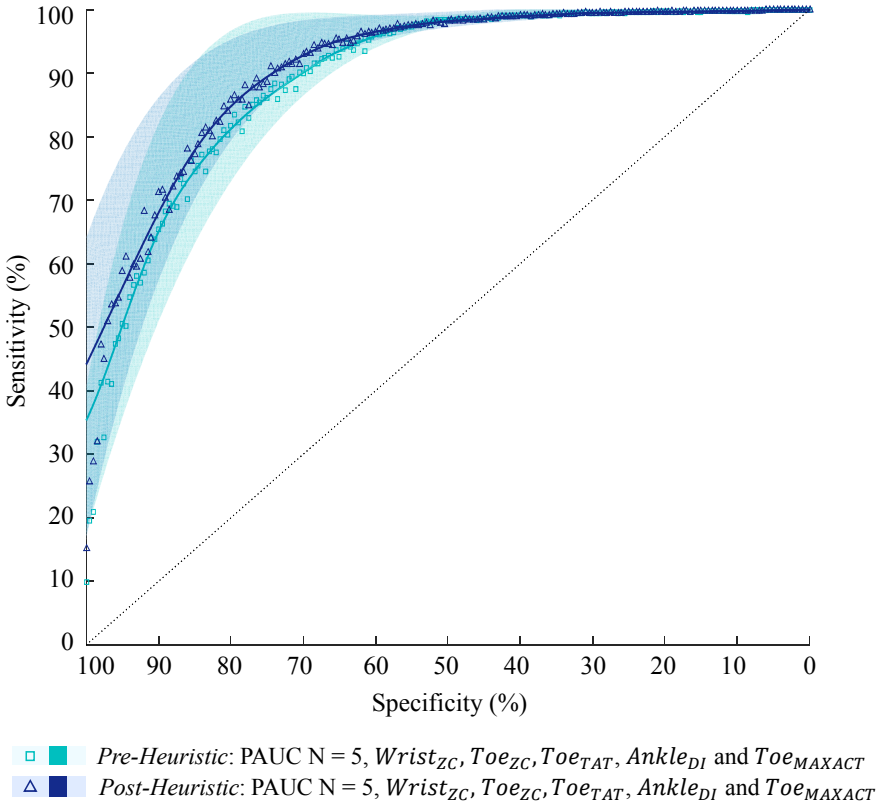
The ROC analysis and agreement with polysomnography are summarised for the conventional tri-axial wrist accelerometer placement and for multisite accelerometry pre- and post-removal of movements specific to sleep in Table 5.1. Applying the heuristic improved predictive performance with multisite accelerometry when compared to the conventional single wrist accelerometer (specificity at 85% sensitivity: 85.6(16.3)% vs. 72.0(22.9)%, $p < 0.01$; Kappa agreement: 0.630(0.292) vs. 0.488(0.257), $p < 0.01$). The heuristic appeared to generally improve the predictive performance for the conventional wrist placement, however it was not statistically significant.

5.1.3 Discussion

The objective of this section was to determine if removing the influence of movements associated with restless sleep significantly improved the predictive performance by reducing false wake detections. Predicting sleep and wake regions with activity counts generated from the heuristically pre-processed raw accelerometry data generally predicted sleep and wake better than standard activity counts (illustrated by comparative ROC curves in Fig. 5.6); however, the improvement was not statistically significant. Using the heuristic to preliminarily remove the influence of movements associated with sleep significantly improved the predictive performance when compared to the conventional single wrist accelerometer placement. This improvement demonstrates that some movements associated with sleep can be successfully isolated and differentiated from movements associated with wake. Similar to Chapter 4, the heuristic indicates that activity counts are unable to distinguish activity during sleep from wake because they



(a)



(b)

Figure 5.6: Median ROC curves for the full population used in this analysis for (a) the conventional wrist accelerometer placement and (b) multisite accelerometry, pre- and post-heuristic.

summarise movement within a fixed epoch. Therefore, activity counts cannot determine the type of activity or specific characteristics of the activity, which heuristically pre-processing the high-resolution raw accelerometry data has demonstrated is likely to aid in improving the poor specificity.

5.2 Differentiating movements on the basis of spectral characteristics[2]

Section 5.1 demonstrated that there are characteristic movements that appear only in sleep, and, together with Chapter 4, verified that activity counts do not have the resolution to specifically identify movements or movement characteristics. Therefore, in this section we will identify localised spectral characteristics that differ between segmented high-resolution wrist movements during sleep and wake. To minimise system complexity, we will restrain analysis to a single accelerometer placement. Considering that the wrist placement was identified as the single most effective sensor in Section 4.2, and that wrist movement appeared in each of the heuristic definitions in Section 5.1, we will only analyse wrist movements.

In this section we will:

1. Characterise the spectral properties of movements segmented from high-resolution, raw tri-axial accelerometry; and
2. Identify spectral characteristics that significantly differentiate sleep movements from wake movements.

5.2.1 Method

The spectral characteristics of wrist movement during sleep and wake were analysed using the discrete wavelet transform. Movements were segmented using a custom algorithm (described in Section 3.2.3), and separated into bins based on duration ($<$ 2s, $2-5$s, $5-10$s, $10-15$s and $>$ 15s). A spectrogram was derived for each axis of movement using the over complete discrete wavelet transform (OCDWT) [173]. The OCDWT was used to ensure shift-invariance (i.e. the response remains the same, regardless of when it occurs temporally), while retaining computational efficiency [173, 174, 175]. Eight scales were used for the decomposition, representing $0.78-1.56$Hz, $1.56-3.13$Hz, $3.13-6.25$Hz, $6.25-12.5$Hz, $12.5-25$Hz, $25-50$Hz, and $50-100$Hz. The Mallat algorithm (i.e. sub-sampling by 2) was applied to the first two scales of the decomposition ($25-50$Hz and $50-100$Hz), and the *à trous* algorithm (i.e. no sub-sampling) was applied to the remaining six scales [175]. We will ignore scales $1 = 25-50$Hz and $2 = 50-100$Hz in the analysis because of the shift-variant nature of these scales, and as

[2]This work has been published in the 36th Annual International Conference of the IEEE Engineering in Medicince and Biology Society: "Characterization of Movements during Restless Sleep in Children: A Pilot Study" [172]

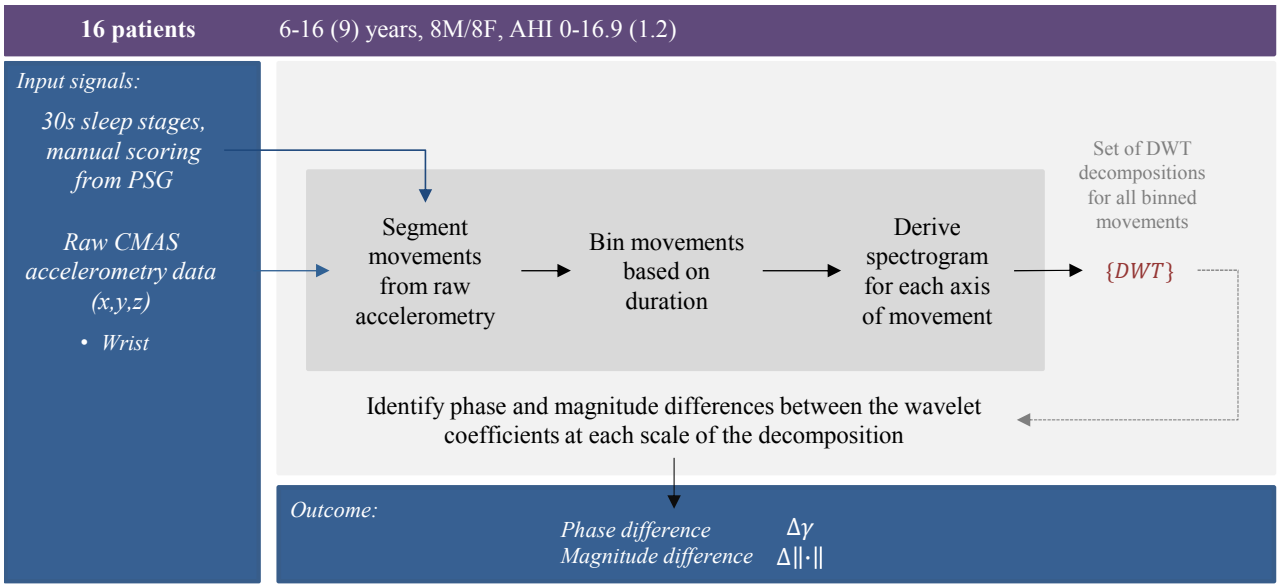


Figure 5.7: Methodology for exploring the time-frequency characteristics of movement during sleep and wake for 16 *patients.*

discussed in Section 3.2.4 in Chapter 3, movements in sleep assessment are not likely to occur in these frequency ranges. The phase and magnitude difference between wavelet coefficients at each scale was then determined at a scale-dependent temporal offset (described below). Differing spectral characteristics were identified by comparing an average sleep and wake spectrogram within each temporal bin. Sixteen patients were used to derive these spectrograms (detailed below). Considering that the significant regions in the difference OCDWT spectrograms were manually extracted, it is not feasible to describe these regions using a cross-validation design. Section 5.3 requires the remaining patient set to test the predictive performance of a model that uses these characteristics. The process for deriving the differentiable characteristics is illustrated in Fig. 5.7 and described in detail below.

Patient characteristics

A subset of 16 patients ($3-12$, 15, 21, and $28-31$ in Table 3.1.3) were used to identify spectral characteristics that differed between sleep and wake movements. The patients were manually selected based on a generalised spread of characteristics: i.e. aged $6-16$ years (median 9 years, 8M/8F) with an AHI range of mild to severe (range $0-16.9$, median 1.2). This set of patients was used to identify significantly different movement characteristics that are representative of the cohort.

Representing movement in the 3-dimensional plane

Tri-axial accelerometers have been used to effectively estimate body positions using the mathematical framework of inclinometry [176]. Inclinometers measure position changes relative to

the line of gravity. Unlike inclinometers, accelerometers can only measure position changes relative to the orientation of the acceleration vector. As such, these systems rely on *reference positions* to determine the angle and azimuth of postural changes. Here we are interested in measuring the *relative* movement differences and so reference positions are not necessary. The magnitude and phase difference between consecutive samples of the tri-axial accelerometry data were used to summarise movement in three dimensions, similar to [176]. The magnitude and phase difference were calculated from the wavelet decomposition of each $x-$, $y-$ and $z-$axis of each movement (described below). Fig. 5.8 illustrates two vectors, $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$, in three dimensions and their projections onto the $X-Y$ and $X-Z$ planes ($\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ represent consecutive samples of the tri-axial accelerometry data).

The phase difference represents an approximation of the phase change between the vectors in the $X-Y$ (α) and $X-Z$ (β) planes, illustrated in Fig. 5.8b and Fig. 5.8c, and is found using the dot product rule,

$$\alpha_{AB} = \arccos\left(\frac{\tilde{\mathbf{A}}_{\mathbf{XY}} \cdot \tilde{\mathbf{B}}_{\mathbf{XY}}}{\|\tilde{\mathbf{A}}_{\mathbf{XY}}\|\|\tilde{\mathbf{B}}_{\mathbf{XY}}\|}\right),$$

$$\beta_{AB} = \arccos\left(\frac{\tilde{\mathbf{A}}_{\mathbf{XZ}} \cdot \tilde{\mathbf{B}}_{\mathbf{XZ}}}{\|\tilde{\mathbf{A}}_{\mathbf{XZ}}\|\|\tilde{\mathbf{B}}_{\mathbf{XZ}}\|}\right). \tag{5.1}$$

The angle given by the dot product in (5.1) represents the angle between the vectors from 0 rad to π rad. As this is an absolute value, it is unable to differentiate a coherent phase change from vector oscillations around a point in the unit circle. To detect these vector oscillations, a heuristic identified when the accelerometry vector returned past the previous vector. However, as the data is sampled every 0.01s, any phase changes will be minute. If every sample is compared, it is also difficult to detect large phase changes that occur over a period of time. For this reason, the phase difference between time-delayed samples was analysed. The time-delay for each scale was defined as 25% of the minimum period present in that scale. Since this satisfies the Nyquist criterion, the 25% scaling avoids any aliasing effects. The dynamic time-delay also ensures that high frequency signals have a shorter time-delay than lower frequency signals. For example, the second scale ($25-50$Hz) has a period range of $T = 0.02-0.04$s. The time delay in this band is 25% of the minimum period, $\delta = 0.005$s. The time delay δ for each scale ψ is given by,

$$\delta = \frac{2^{\psi-1}}{4F_s}.$$

The direction of the phase change was identified and used to differentiate coherent phase changes from vector oscillations. That is, if a vector moved to the opposing quadrant of the previous vector, the phase change α, β was represented as a negative phase; otherwise it was represented as a positive phase. Since the phase change is additive, a negative phase change

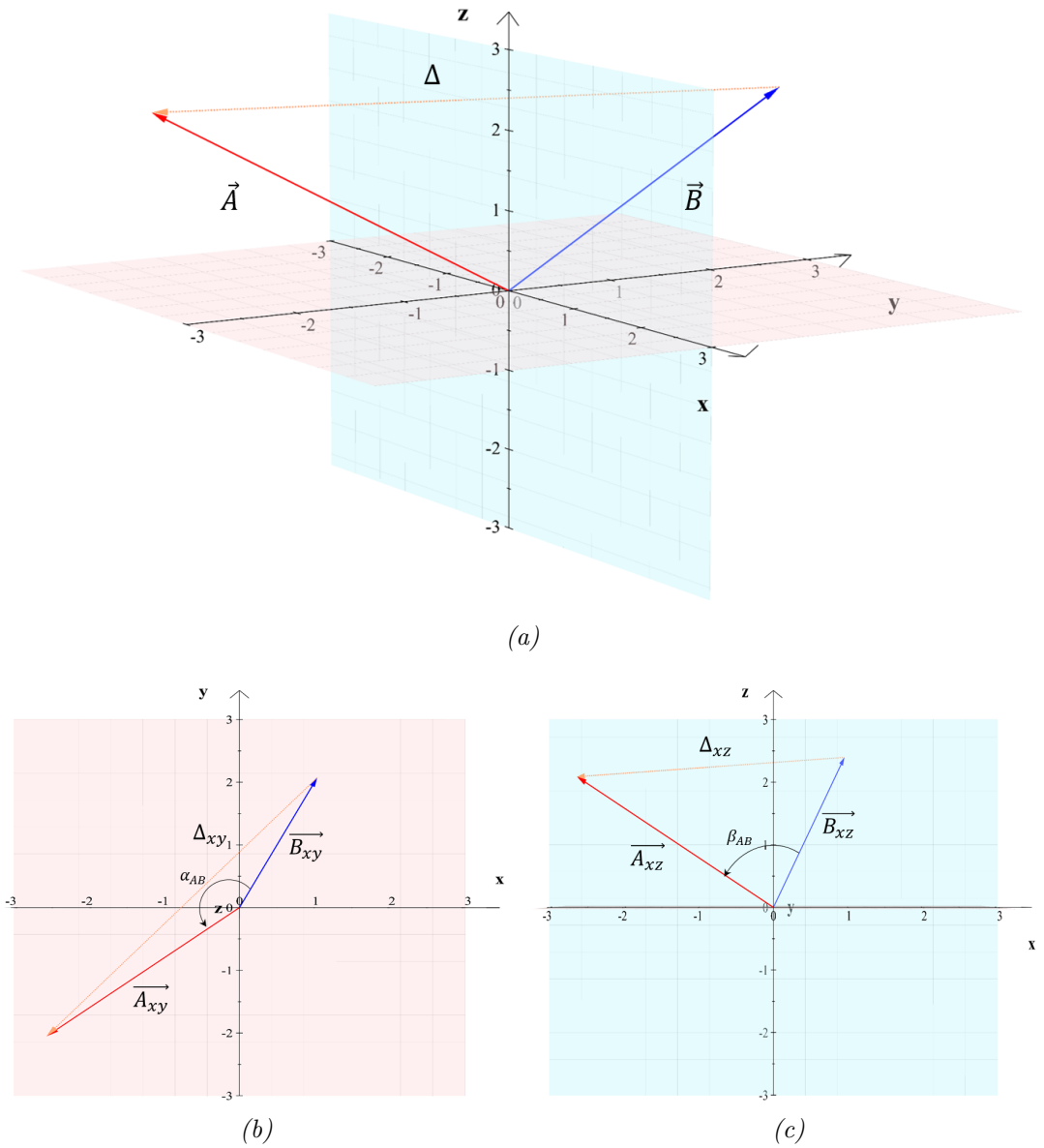


Figure 5.8: Representation of two example movement vectors in three dimensions. Vectors $\tilde{\mathbf{A}}$, $\tilde{\mathbf{B}}$ *and the difference between them,* $\boldsymbol{\Delta}$ *is shown in (a). The projection onto the (b)* $X-Y$*, and the (c)* $X-Z$ *planes are also shown.*

(representing an oscillation) is cancelled. The second previous vector is referred to here as the 'reference vector'. The current movement vector was identified as an oscillation if the difference between the angle of the reference vector $\theta_{n-2\delta}$ and the current vector θ_n had a different sign than the difference between the angle of the reference vector and the previous vector $\theta_{n-\delta}$;

$$\alpha_{AB}, \beta_{AB} = \begin{cases} -\Delta_{AB_n} & \text{if } \text{sign}(\theta_{n-2\delta} - \theta_n) \neq \text{sign}(\theta_{n-2\delta} - \theta_{n-\delta}) \\ & \quad \text{and } \theta_{n-\delta} \neq \theta_{n-2\delta}, \\ \Delta_{AB_n} & \text{otherwise,} \end{cases} \tag{5.2}$$

where Δ_n and Δ_{n-1} represent the phase between the reference vector and the current and

previous vectors respectively, given by (5.1). $\theta_n, \theta_{n-\delta}$ and $\theta_{n-2\delta}$ represent the angles of the current, previous and reference vectors respectively, given by the *atan2* function. This algorithm presents problems if all vectors lay in the $\pi/2 \rightarrow \pi$ and $-\pi/2 \rightarrow -\pi$ quadrants. To work around this, the angle θ of any vectors in the $-\pi/2 \rightarrow -\pi$ quadrant is transformed to the range $0 \rightarrow 2\pi$;

$$\theta = \begin{cases} 2\pi - |\theta| & \text{if } -\pi \leq \theta \leq -\pi/2, \\ \theta & \text{otherwise.} \end{cases}$$

We are interested in the phase with the greatest coherence within either the $X-Y$ plane or the $X-Z$ plane. As such, the phase γ between the vectors was identified as the maximum angle difference from either the $X-Y$ plane or the $X-Z$ plane,

$$\Delta\gamma = \max(\alpha_{AB}, \beta_{AB}), \tag{5.3}$$

where α and β were found using (5.2). Taking the maximum phase ensures that any vector oscillation in an axis of movement is ignored when presented with a coherent phase shift in the other axis. To ignore small oscillations caused by noise, the phase difference was only calculated if the magnitude of the vector was within the top 90% of magnitude values within the transform coefficients. The last piece of information to capture was the change in magnitude between the accelerometry vectors,

$$\Delta\lambda = |\|\tilde{\mathbf{A}}\| - \|\tilde{\mathbf{B}}\||. \tag{5.4}$$

The final feature vector represents the change in angular displacement (5.3) and magnitude (5.4). This is referred to here as the 'differential feature vector' Δf_θ,

$$\Delta f_\theta = \begin{bmatrix} \Delta\gamma \\ \Delta\lambda \end{bmatrix}. \tag{5.5}$$

Although the effects of gravity influence the magnitude of acceleration, studies have shown that the small benefits of removing this influence from the appropriate axis are not worth the complexity of estimating the gravitational component for un-calibrated accelerometers [177]. For this reason, the gravity component was not removed from the acceleration vector.

Movement representation development

The over-complete discrete wavelet transform (OCDWT) with two critically sampled levels was used to represent the varying spectral characteristics for the x, y and $z-$axes separately [173]. Performing the decomposition on the individual axes of movement prior to generating the summary representations ensures that all movements can be represented in the spectral domain; the decomposition is performed on the raw movement, not on a summary metric such as magnitude or phase. Performing the decomposition on a raw axis of movement occurs in

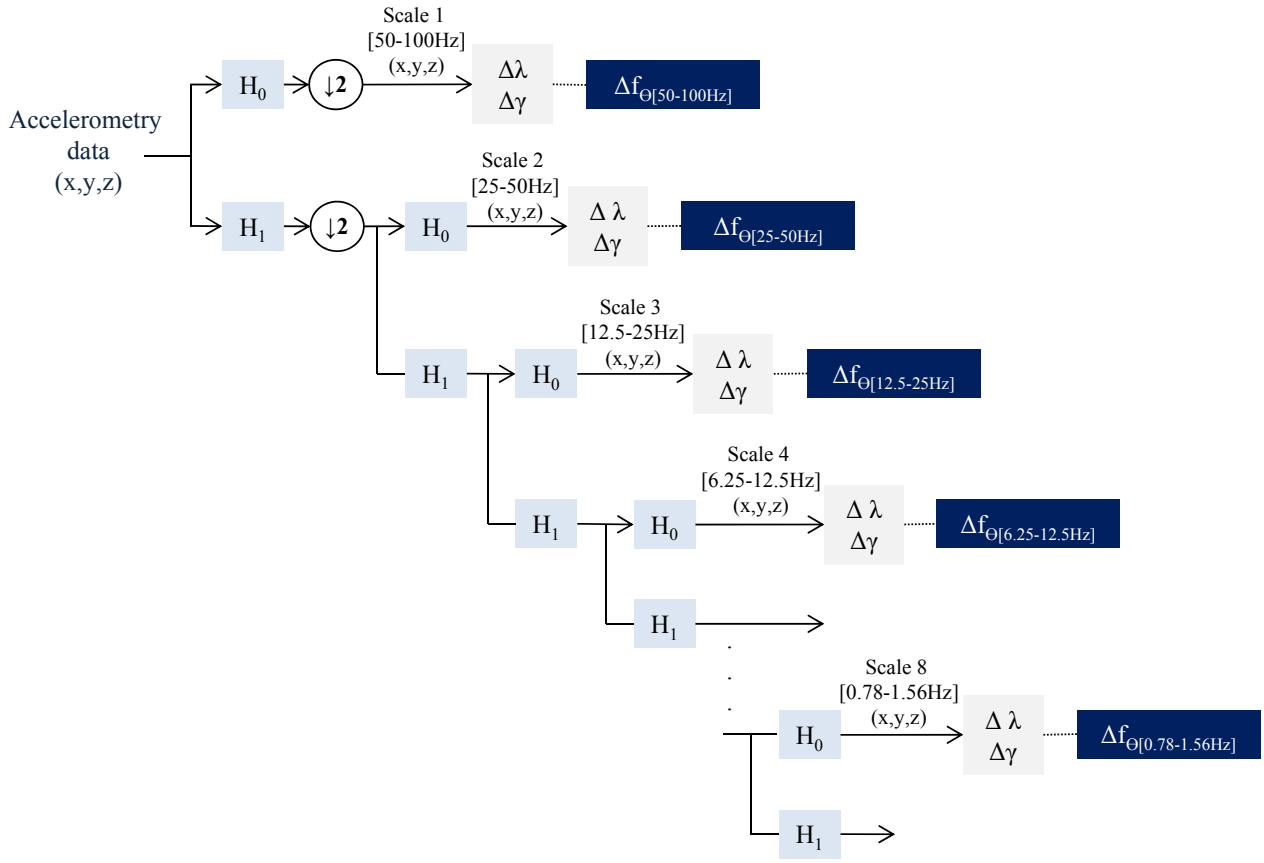


Figure 5.9: Process for the over-complete discrete wavelet decomposition (OCDWT) for representing movement during sleep and wake [173]. The OCDWT is performed separately for each axis of movement (x, y, z) and the magnitude difference $\Delta\lambda$ and coherence of the phase difference $\Delta\gamma$ at each level of the decomposition are calculated from the (x, y, z) wavelet decompositions. $\downarrow 2$ indicates critical sub-sampling.

physical activity assessment [117]. Since movement analysis in Section 3.2.4 found that we are not interested in frequencies above 25Hz, the two critically sampled levels correspond to the $25 - 50$Hz and $50 - 100$Hz frequency bands. These sub-bands were subsequently ignored. The magnitude and phase difference were then calculated at each scale of the wavelet decomposition. The average coefficients for each movement were analysed for sleep and wake. This procedure is illustrated in Fig. 5.9, and the Gaussian averaging technique is described in Section 3.5. The median wavelet coefficient at each temporal position for each scale was then found for the magnitude and phase difference of each movement category (illustrated in Fig 5.10). The final average spectrogram of sleep and wake movements was analysed, and the predominant spectral characteristics that differed between sleep and wake were identified.

Statistical analysis

The distribution of wavelet coefficients within each temporal bin for sleep and wake movements was compared using Welch's t-test because the number of coefficients between the sleep and

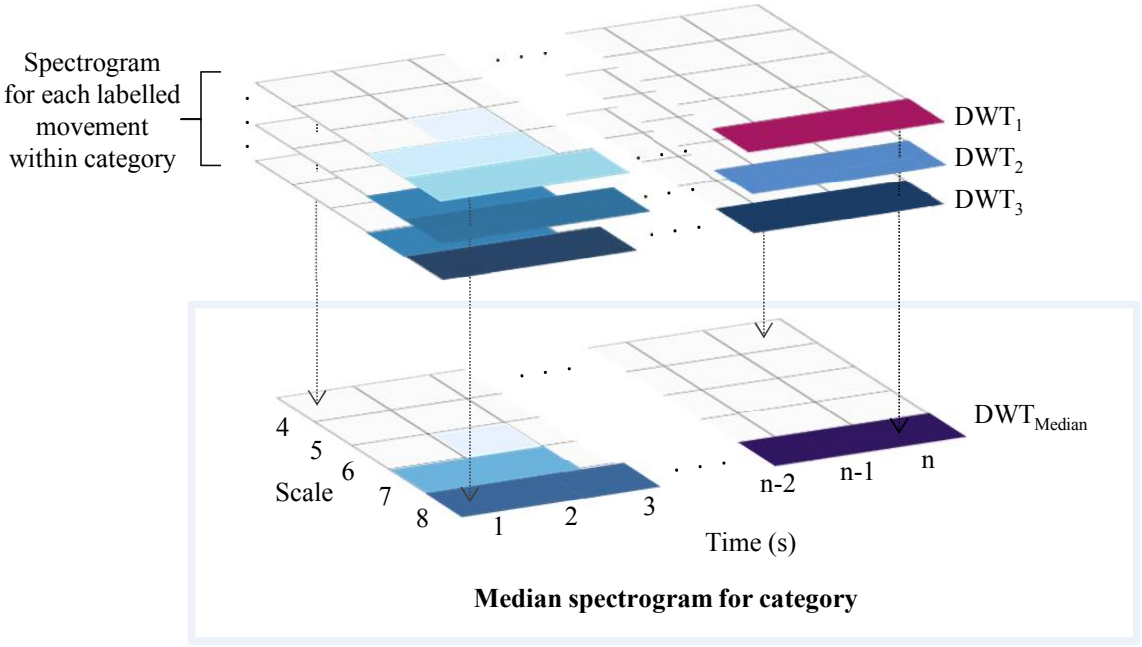


Figure 5.10: Procedure for combining discrete wavelet decompositions for each labelled movement DWT_n within each category. The wavelet coefficients for the median spectrogram DWT_{median} are found by calculating the median coefficient at each time position within each scale from each decomposition within that category.

wake spectrograms differed. A spectrogram representative of the statistical significance given by the $p-$value from Welch's t-test was generated for each comparison.

5.2.2 Results

There were regions of significant differences for movements $2-5$s and $5-10$s in duration, detailed below and summarised in Table 5.2. There were no statistically significant differences in spectral characteristics for wrist movements less than 2s in duration, or greater than 10s in duration. The white regions in the $p-$value difference spectrogram (bottom of Fig. 5.11) for movements $2-5$s in duration illustrate the spectral characteristics that significantly differed between sleep and wake. The energy of the magnitude difference was significantly greater ($p < 0.01$) for wake movements of $0.781-3.13$Hz between $0.2-1$s and $1.2-2$s, and $0.781-1.56$Hz between $3.5-4.5$s. The phase difference was significantly greater ($p < 0.01$) for wake movements of $0.781-1.56$Hz between $0.5-1$s and $2.3-2.6$s, and $1.56-3.13$Hz between $1-1.2$s and $1.7-1.9$s. There were other regions of significant differences; however, these were too short in duration and too scattered throughout the movement to be of practical significance. Unlike short-duration movements (shown in Fig. 5.11), the spectral characteristics of wrist movements $5-10$s in duration did not generally differ. As illustrated in Fig. 5.12, the energy was only significantly greater for the magnitude difference of $0.781-1.56$Hz movements during sleep at $4-4.1$s, $p < 0.01$. Longer-duration sleep and wake movements had similar low-frequency spectral content.

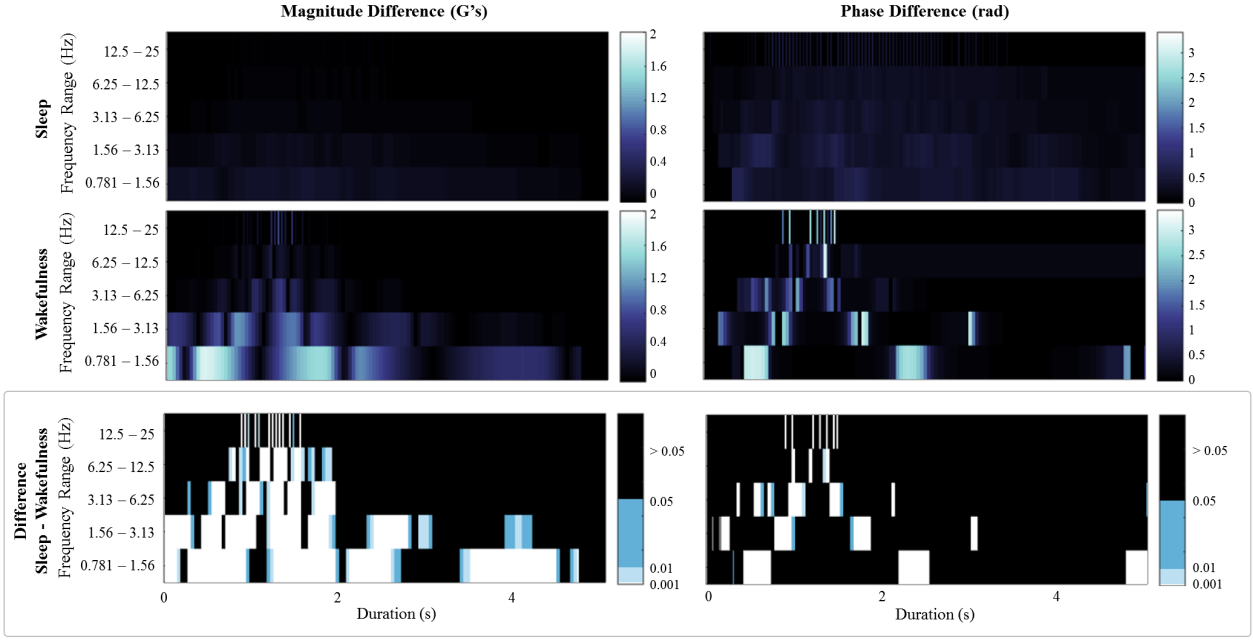


Figure 5.11: Spectrogram of the magnitude and phase difference of wrist movement during sleep and wake for a duration of $2-5s$*. The difference spectrograms display the statistical significance of each time-frequency location (*$p < 0.05$ *blue,* $p < 0.01$ *light blue, and* $p < 0.001$ *white).*

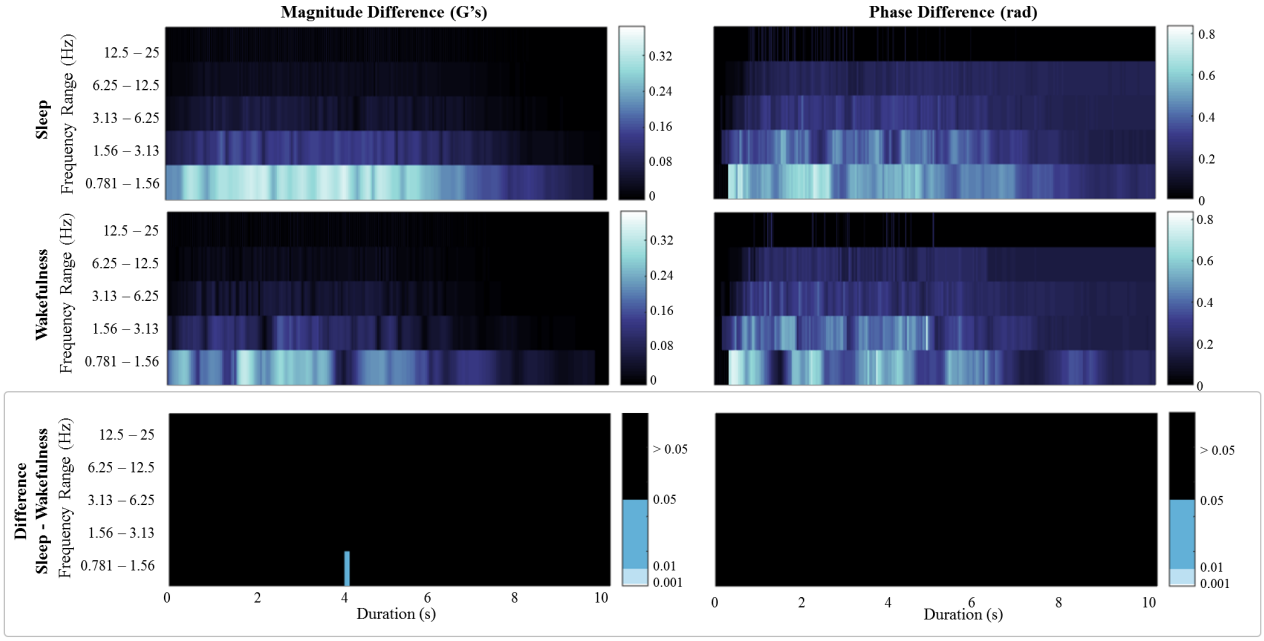


Figure 5.12: Spectrogram of the magnitude and phase difference of wrist movement during sleep and wake for a duration of $5-10s$*. The difference spectrograms display the statistical significance of each time-frequency location (*$p < 0.05$ *blue,* $p < 0.01$ *light blue, and* $p < 0.001$ *white).*

5.2.3 Discussion

The objective of this section was to identify spectral characteristics that differ between sleep and wake movements. We identified some temporal segments of short-duration $(2-5s)$ movements

where there was significantly greater spectral energy for wake movements. The reduced spectral energy of sleep movements is likely due to inhibited muscle tone during sleep [8]. As illustrated by the greater average wavelet coefficients in Fig. 5.11, short-duration wake movements are generally more vigorous than short-duration sleep movements.

As previously discussed, movements during wake are a conscious process and are consequently coherent in nature. This is reflected in the grouped average wavelet coefficients in Fig. 5.11: short-duration movements during wake were fairly consistent in nature, whereas the average sleep coefficients were widely spread. Supporting this finding, Domingues et al. [18] analysed the nature of wake movements and report a high degree of autocorrelation. Considering that the characteristics were fairly consistent across the movements, characteristics that are specific to wake are likely to be effective for a generalised classification model. However, sleep characteristics are irregular and vary across the movements. This is consistent with the behaviour of activity counts seen in Chapter 4: wake movements commonly produce large activity values, whereas sleep movements result in a spread of activity values. Some sleep movements will consequently appear as 'wake' movements and will always impact the accuracy of wake predictions.

Movements $5-10$s in duration had predominantly low-frequency spectral content during both sleep and wake. As expected, the low-frequency spectral content indicates that these lengthier movements are positional changes, rather than limb or muscle twitches. Characteristics of positional changes are unlikely to significantly differ between sleep and wake. Indeed, we saw this in the difference spectrogram in Fig. 5.12, where the low-frequency scales had similar

Table 5.2: Time-frequency characteristic features and their corresponding descriptions

Feature[a]	Representation	Frequency Range	Temporal Location	Description
λ_{W_1}	$\Delta\lambda$	$0.781-3.13$Hz	$0.2-1$s	Wake energy > sleep energy
λ_{W_2}	$\Delta\lambda$	$0.781-3.13$Hz	$1.2-2$s	Wake energy > sleep energy
λ_{W_3}	$\Delta\lambda$	$0.781-1.56$Hz	$3.5-4.5$s	Wake energy > sleep energy
γ_{W_4}	$\Delta\gamma$	$0.781-1.56$Hz	$0.5-1$s	Wake energy > sleep energy
γ_{W_5}	$\Delta\gamma$	$0.781-1.56$Hz	$2.3-2.6$s	Wake energy > sleep energy
γ_{W_6}	$\Delta\gamma$	$1.56-3.13$Hz	$1-1.2$s	Wake energy > sleep energy
γ_{W_7}	$\Delta\gamma$	$1.56-3.13$Hz	$1.7-1.9$s	Wake energy > sleep energy
λ_{S_1}	$\Delta\lambda$	$0.781-1.56$Hz	$4-4.1$s	Sleep energy > wake energy

[a] Features are descriptions of the large significant regions from the difference spectrogram illustrated in Fig. 5.11.

coefficients for both sleep and wake, and the difference spectrogram was not significant. As evidenced by the greater wavelet coefficients, short-duration wake movements generally appeared more vigorous than similar sleep movements. Therefore, characteristics indicative of vigorous movement are likely to have high discrimination ability between sleep and wake movements. Indeed, we saw in Chapter 4 that the zero-crossing method (representative of vigorous movement by counting the number of times the acceleration vector crosses a threshold) generally out-performed all other time-series methods. This is unsurprising as sleep and wake detection with actigraphy works on the principle that wake corresponds to regions of increased activity. The greater energy in the lower frequency scales is indicative of positional and/or postural changes, which are represented by angle approximations (such as SUMPST in Chapter 4).

Section 5.1 identified that 'bursty' movements (i.e. movements < 2s in duration) and hand movements with positional changes were more common during sleep. Although the relative prevalence of these movements were not assessed in this section, we saw that the spectral characteristics of these movements differed between sleep and wake. Differing characteristics further validate the sleep and wake differentiation ability of the heuristic. The analysis in Chapter 4 found that the temporal resolution of the standard 30s epochs effectively averaged movement information, reducing characteristically different movements to the same activity count. This analysis has shown that there is movement information that may significantly differentiate sleep and wake on a movement-by-movement basis. The vigorous movements and postural changes prevalent in wake movements suggest that features representative of these characteristics should be explored within a sleep and wake classification model. This will be explored in the next section.

5.3 Sleep/wake prediction on a movement-by-movement basis

In the previous section we explored the spectral content of movements to identify characteristics that differentiate sleep from wake. In this analysis, we concluded that there were indeed characteristics that could be exploited to classify segmented movements as occurring during 'sleep' or 'wake'. However, as we saw in Chapter 4, conventional actigraphy is unable to differentiate sleep movements from wake movements; sleep and wake movements can produce the same activity count over a fixed epoch of time, and fixed epochs cannot segment specific movements. Detecting sleep and wake periods on a movement-by-movement basis may improve the predictive performance because it analyses characteristics of each movement, rather than fixed epochs of data that may intersect or combine multiple movements. However, there have been no known attempts in the literature to predict sleep and wake with segmented movements.

We have seen in Section 5.2 that characteristics representative of vigorous activity or postural changes are much more common in short-duration wake movements. It is likely then that

these characteristics can differentiate wake movements from sleep movements. Since these metrics are movement-dependent, this analysis cannot be performed with fixed epochs. The aim of this section is to identify if predicting wake on a movement-by-movement basis has improved accuracy over the conventional fixed epoch time-series summary method.

In this section we will:

1. Develop a regression model that predicts wake on a movement-by-movement basis using the spectral characteristics identified in Section 5.2; and
2. Compare the predictive performance of the movement-by-movement regression model to the conventional zero-crossing threshold method.

5.3.1 Method

All 30 patients were used in this analysis: 15 patients that were characteristically representative of the full cohort ($2-10$, 15, 21, 27, and $30-32$ in Table 3.1.3, 10 male, median 9 years, range $6-16$ years) formed the training dataset for the regression model, and the other 15 patients ($11-12$, $19-20$, 22, $23-25$, $28-29$, $33-35$ and $37-38$ in Table 3.1.3, 12 male, median 9 years, range $6-15$ years), were used to test the model using leave-one-out cross-validation on patients.

Movements during sleep and wake were segmented from the raw tri-axial wrist accelerometry data for each patient in the training dataset using the process detailed in Section 3.2.3. The spectral characteristics identified in Section 5.2 (detailed below) were extracted for each movement. These spectral characteristics, together with temporal characteristics that describe vigorous movement and postural changes, formed the final feature set for training the linear regression model. This regression model was applied to the patients in the test dataset. The predictive performance of this model was then identified and compared to the conventional zero-crossing threshold method using ROC analysis. ROC analysis was used to identify the threshold for classification; i.e. the threshold that gave the greatest Kappa agreement in the cross-validation training set within each fold. ROC analysis was performed on the raw zero-crossing data and on the scaled wake likelihood given by the regression model.

The regression model was assessed on a movement-by-movement basis and also translated into a time-series representation. The movement-by-movement regression model only considers the sleep and wake regions that correspond with movement. Translating the prediction scores into a time-series signal also considers both sleep and wake regions that do not coincide with movement. Essentially, the regression model provides a 'summary' value (i.e. a likelihood of coinciding with wake) for each movement. Time-indices with no detected movements are given a value of 0. This process is described in more detail below and illustrated in Fig. 5.13. The methodology is illustrated in Fig. 5.14.

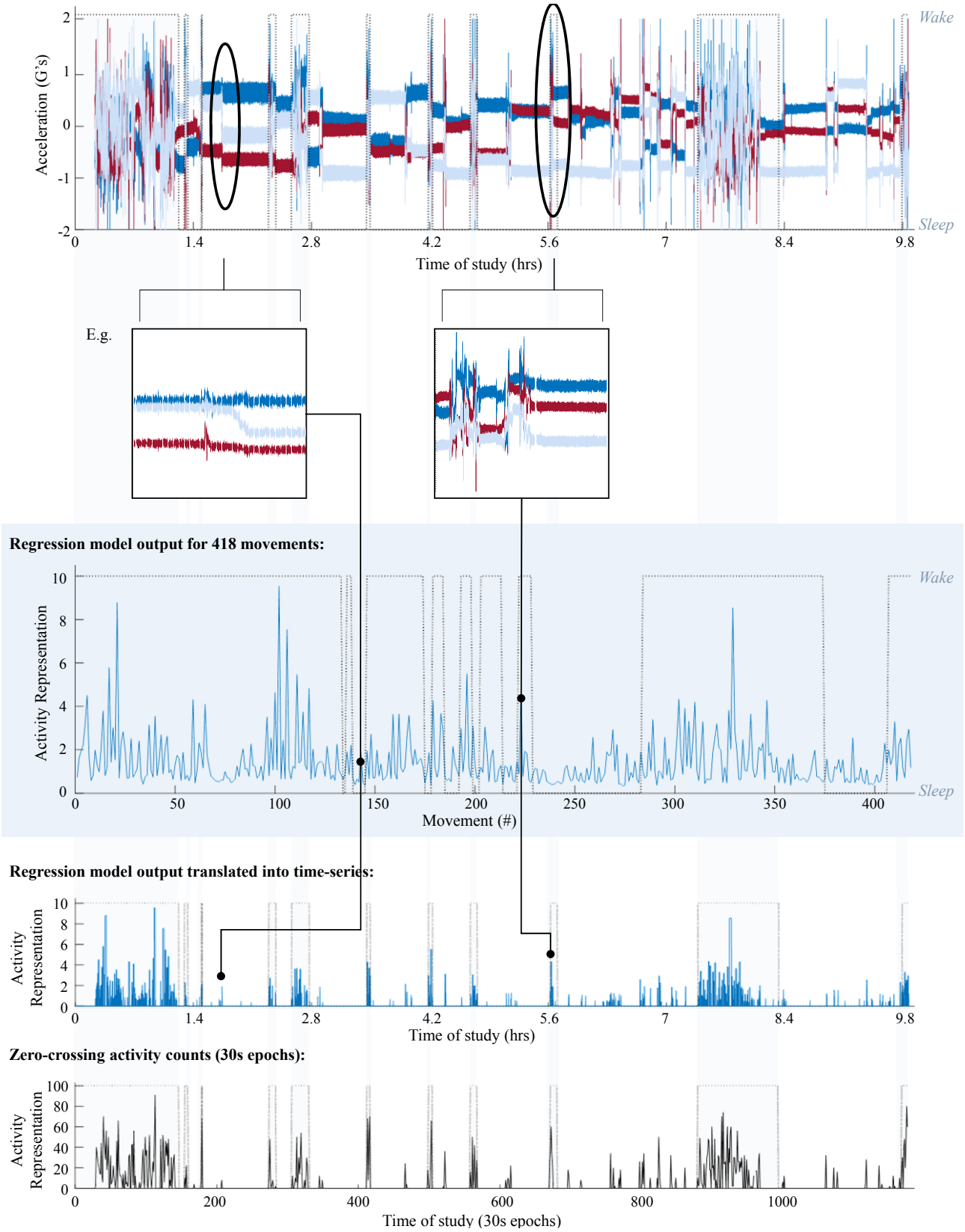


Figure 5.13: Procedure for representing movement with the regression model (movement-by-movement basis and translated into an equivalent time-series representation), as compared to the standard 30s zero-crossing activity counts. Example from one full-night patient study.

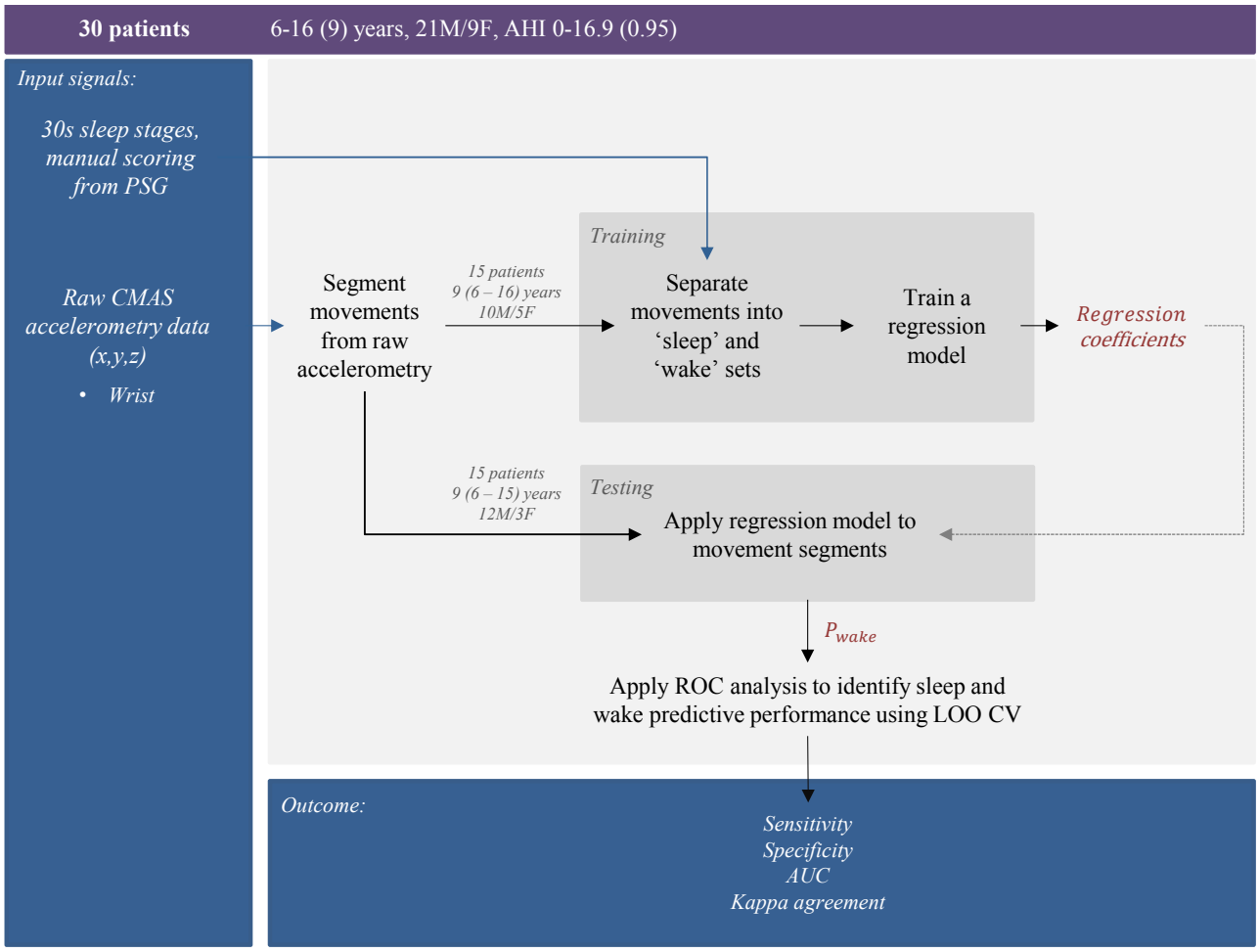


Figure 5.14: Method for identifying the predictive performance of detecting sleep and wake on a movement-by-movement basis.

Regression model

The feature vector for sleep and wake movements includes temporal characteristics that summarise vigorous movement and postural changes, and the specific spectral characteristics identified in Section 5.2.

The spectral features for wake $\mathcal{F}$ were defined as the N wake characteristics summarised in Table 5.2,

$$\mathcal{F} = \{\lambda_{W_n}(f,t), \gamma_{W_n}(f,t)\}, \quad \forall n, \quad 1 \leq n \leq N. \tag{5.6}$$

The differentiable segments of high spectral energy from Fig. 5.11 (summarised in Table 5.2) consist of a *set* of wavelet coefficients. Therefore, different statistical summary operations were performed on each of these segments to summarise the spectral content for consideration in the regression model: maximum, mean, summation and variance.

The final set of temporal features and spectral regions from Section 5.2 used in the regression model were:

Label	Description
	Temporal features
δ	Duration of movement
R	Auto-correlation with a 0.1s offset
$\mathcal{S}_M$	Total magnitude of the movement
$\mathcal{M}_M$	Maximum magnitude of the movement
$\mathcal{S}_P$	Total degree of positional changes within the movement
$\mathcal{M}_P$	Maximum degree of positional change within the movement
	Spectral regions (from Section 5.2)
λ_{W_1}	$\Delta\lambda$, $0.781 - 3.13$Hz, $0.2 - 1$s
λ_{W_2}	$\Delta\lambda$, $0.781 - 3.13$Hz, $1.2 - 2$s
λ_{W_3}	$\Delta\lambda$, $0.781 - 1.56$Hz, $3.5 - 4.5$s
γ_{W_4}	$\Delta\gamma$, $0.781 - 1.56$Hz, $0.5 - 1$s
γ_{W_5}	$\Delta\gamma$, $0.781 - 1.56$Hz, $2.3 - 2.6$s
γ_{W_6}	$\Delta\gamma$, $1.56 - 3.13$Hz, $1 - 1.2$s
γ_{W_7}	$\Delta\gamma$, $1.56 - 3.13$Hz, $1.7 - 1.9$s

Linear regression identified the fit of the model on the training patient data using the movement features described above: the temporal features and the summary operations (maximum Max, mean $\mathbb{E}$, summation Σ and variance Var) applied to the spectral regions. The normalised features were used to identify the significant features for the model (i.e. features with a p-value < 0.05). The final model coefficients were identified using the unnormalised features.The final model found in the regression analysis was,

$$P_{wake} = 0.354R + 0.037\mathcal{S}_P - 0.029\text{Max}(\gamma_{W_7}) + 0.008\mathbb{E}(\lambda_{W_2}) - 0.007\text{Max}(\lambda_{W_2}) + 0.003\Sigma\gamma_{W_5}. \quad (5.7)$$

Validation procedure

The predictive performance of the regression model was compared to the conventional zero-crossing threshold technique using ROC analysis in leave-one-out cross validation on the test patients. The threshold for predicting wake was identified as the threshold closest to $(0, 1)$

Table 5.3: Class discrimination ability and predictive performance of the regression model and the conventional zero-crossing thresholding.

	AUC (%)	Kappa	Sensitivity (%)	Specificity (%)
Regression Model	70.2 (8.5)	0.285 (0.157)	70.0 (18.0)[a]	59.6 (20.4)
Regression Model (Time-series)	63.9 (6.7)	0.331 (0.133)	30.3 (13.4)	96.9 (1.3)
Zero-crossing Thresholding	69.7 (7.9)	0.424 (0.157)[a]	42.8 (14.7)	95.3 (2.4)[a]

Values are shown as mean (±SD)
[a] $p < 0.05$, Regression model vs. Zero-Crossing thresholding

in the ROC space of the training set in the cross validation. In addition to movement-by-movement detection, the regression model was analysed when translated back into the 100Hz time-series. As illustrated in Fig. 5.13, the detected movements were represented by their respective regression value, and regions with no movement were set to 0. The AUC, Kappa agreement, sensitivity and specificity of the chosen threshold (from ROC analysis) were assessed.

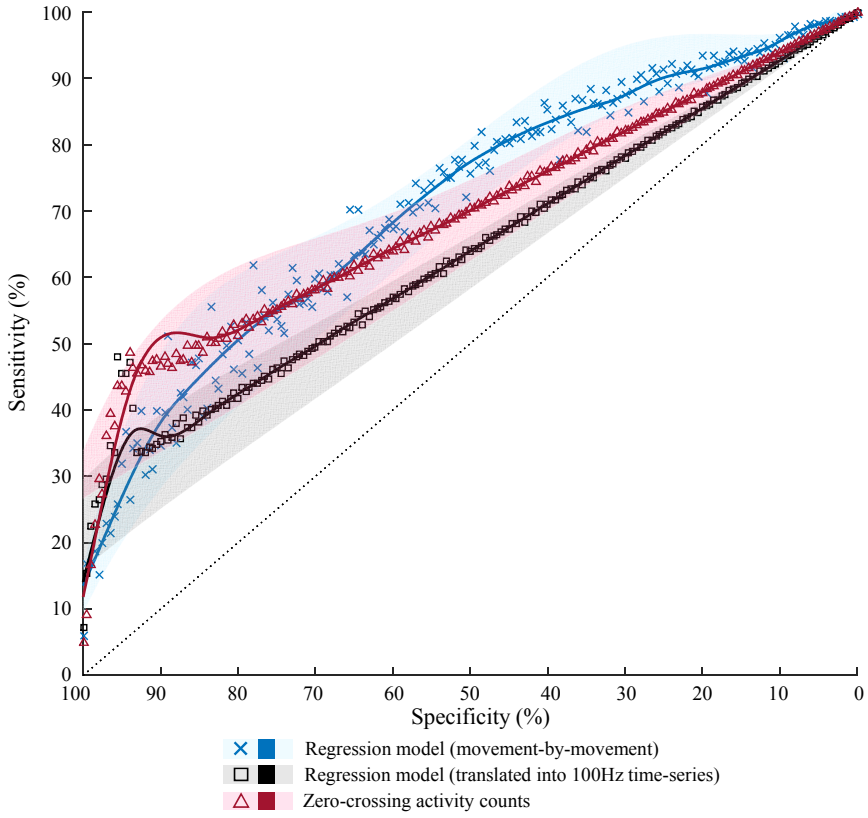
5.3.2 Results

The ROC analysis and Kappa agreement for the regression model and the conventional zero-crossing threshold method are shown in Table 5.3. The sleep and wake discrimination ability was similar between the regression model and the conventional method (AUC: 70.2(8.5)% vs. 69.7(7.9)% respectively). The Kappa agreement was significantly greater for the conventional method (0.424(0.157) vs. 0.285(0.157), $p < 0.05$).

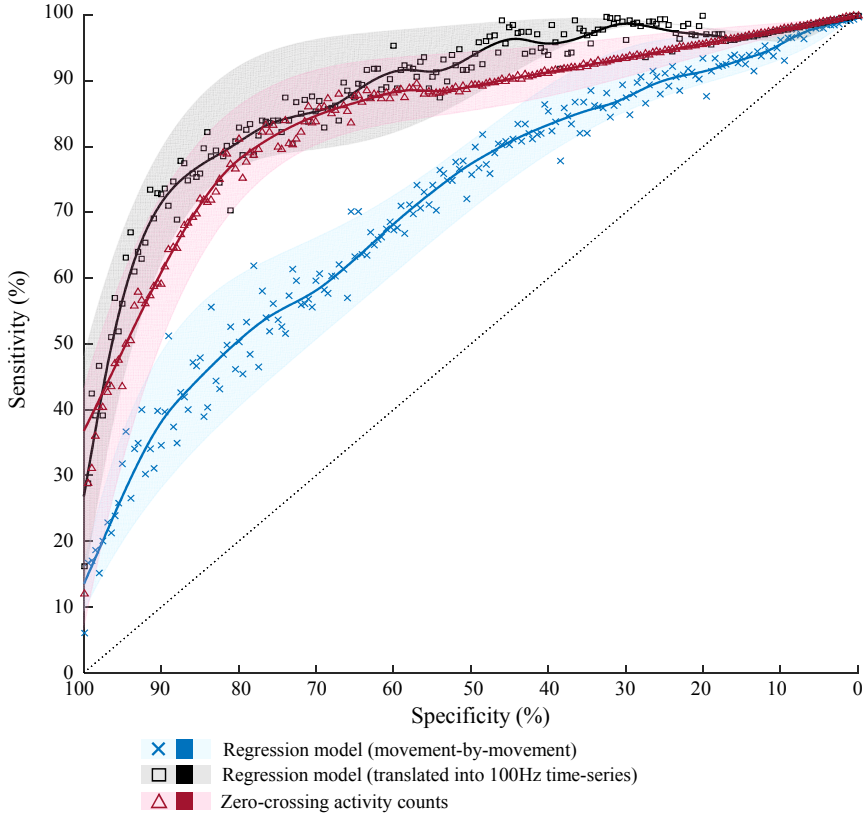
5.3.3 Discussion

The movement-by-movement regression model poorly predicts if movement occurs during sleep or wake. The regression model only classifies segmented movements; it is not affected by quiet rest artefacts or restful sleep. Therefore, the regression model assessed the ability to discriminate sleep regions that are associated with movement from wake regions that are associated with movement. We can speculate from the poor predictive performance that sleep and wake movements cannot accurately be discriminated using a simple threshold or regression model. Considering that the performance of the regression model was similar to the conventional epoch summaries, we can also speculate that the specific method for deriving a movement representation does not significantly affect the predictive performance.

The autocorrelation of the movement was found to be the most significant factor in the regression model. This verifies literature [18] and general data observations in Chapter 4 and Chapter 5 that the coherence of movements during sleep and wake differ. Wake movements are generally more consistent in nature than sleep movements, and are consequently characterised



(a)



(b)

Figure 5.15: Average ROC of the regression model and the zero-crossing thresholding (a) initially and (b) after smoothing with a 5.5 min moving average filter. The filter was not applied to the movement-by-movement regression model (shown in blue).

by a greater degree of autocorrelation. This verifies our movement observations in Section 5.1 that sleep movements appear more 'random'. However, the poor predictive performance suggests that the degree of this characteristic difference is not substantial for discriminating sleep movements from wake movements.

The model offered slightly better performance when a moving average filter was applied (as illustrated by the comparative ROC curves in Fig. 5.15b). This is consistent with the filter effects seen in Chapter 4; smoothing the time-series representations significantly improved the sleep and wake discrimination ability for both methods. The greatest impact on sleep and wake prediction accuracy has consistently been the smoothing filter (seen here and in Chapter 4). The significant increase in discrimination ability when applying the smoothing filter (highlighted in Fig. 5.15) is explained by the addition of activity values in wake periods that, prior to the smoothing, did not contain activity. Applying a threshold to these activity values will correctly identify these regions of quiet rest as 'wake' where, prior to smoothing, these regions were incorrectly identified as 'sleep'. The greatest performance improvement in sleep and wake estimation with actigraphy is seen when addressing the quiet rest limitation.

5.4 Summary

In this chapter we sought to differentiate movements that occur during sleep from movements that occur during wake, with the aim of improving false wake detections. We hypothesised that the physiological characteristics of sleep and wake movements differ, and can consequently directly differentiate sleep regions from wake regions for improved classification performance. First, we explored the effect of specifically removing the influence of movements associated with sleep on wake predictions in Section 5.1. We then identified movement characteristics that differed between sleep and wake in Section 5.2. Finally, we used these characteristics to predict sleep and wake on a movement-by-movement basis in Section 5.3.

Heuristically removing movements that occur during sleep as a pre-processing step in Section 5.1 significantly improved wake predictions. This shows that we can successfully detect some movements that are associated with restless sleep, which indicates that there are some characteristics of movements that occur during sleep that differ from wake. This was verified in Section 5.2, where localised spectral characteristics of short-duration movements differed between sleep and wake. We saw regions of increased spectral energy and characteristics indicative of postural changes during wake movements. Movements during wake were also consistent in nature, whereas the spectral characteristics tended to differ across the sleep movements. Using these characteristics to classify sleep and wake on a movement-by-movement basis in Section 5.3 did not accurately predict the correct class membership of the detected movements. This shows that although some movement characteristics did significantly differ, they were not substantially discriminatory for classification purposes. It is likely that the general inconsistent nature

of sleep movements contributes to this. That is, some sleep movements do differ to wake movements; however, there are also sleep movements that have similar characteristics to wake movements. Therefore, because of the inability to differentiate movements in a generalised context, and the moderate wake predictive performance seen in Chapter 4, we can conclude that movement information alone is not able to accurately classify sleep and wake; there are no *generalisable* movement characteristics that consistently and accurately differentiate sleep from wake.

In this chapter we saw that:

- Some movements associated with restless sleep can be isolated from movements that occur during wake, and removing these sleep movements consequently improves sleep and wake predictive performance;
- Wake movements have greater energy than sleep movements for some localised spectral regions;
- Short-duration wake movements are generally more vigorous than sleep movements and are characteristic of positional changes;
- Movement characteristics in a regression model are unable to accurately classify movements as occurring during sleep or wake; and
- Temporal and spectral characteristics of movements are not able to accurately estimate sleep and wake regions for a generalised cohort; it is unlikely that movement information alone can be effectively manipulated to accurately estimate sleep and wake.

We saw in this chapter that, unlike wake, the nature of sleep movements varied considerably; however, the specific origin of these events has yet to be considered. It is possible that some of these sleep movements are associated with a physiological or pathological event that is characteristic of sleep disturbance associated with a sleep disorder. In the next chapter, we will identify the association between sleep movements and physiological and pathological events (transient arousals, apnoeas and hypopneas) that are associated with sleep disorders.

6

Physiological associations with sleep movements

"Mr. C., aged fifty-six years, ... gave no evidence in the day of respiratory incompetence, ... When in deep sleep he began to breathe less and less deeply, and at last, for a few seconds, not to breathe at all. At this moment he moved, twitched, and at last awakened with evidences of commencing apnoea in the color of the lips, tongue, and nails. When awake a few voluntary efforts to respire relieved him. These attacks became at last so frequent and perilous that a nurse sat by his bed and awakened him as soon as he began to breathe less and less deeply. As time went on the trouble increased, and whenever he fell asleep respiration ceased abruptly. He was finally worn out with loss of sleep, and died suddenly in one of these onsets of respiratory failure."

— Medical case studies, 1890, S. Weir Mitchell, M.D., LL.D Harv.,
Physician, 1829 - 1914 [1]

In Chapter 5 we concluded that although some sleep movements were characteristically different to wake movements, these characteristics were unable to reliably differentiate sleep from wake. This is due to the considerable variability of movement characteristics during sleep. For this reason, it is unlikely that movement characteristics will form an accurate predictor of sleep and wake regions. However, this analysis did not attempt to identify the source of these sleep

[1] Medical case studies and discussion detailed in [178].

movements. If these sleep movements are caused by a physiological process characteristic of sleep disorders (such as transient arousals, apnoeas and hypopneas), movements associated with sleep may capture signs indicative of sleep disorder severity. In this chapter we will explore the temporal association between movements during sleep, transient arousals and apnoeic events.

Arousals from sleep can occur spontaneously or as a response to some stimulus, such as an apnoeic event. During an arousal, the body is mobilised in an attempt to address a potential threat. It is therefore thought that body or limb movements may occur during arousal from sleep [179]. Indeed, historic American Sleep Disorders Association (ASDA) criteria for scoring arousals during REM sleep required leg muscle movement to accompany electroencephalogram (EEG) activity [51] (now optional [42]). Although this is not the case for apnoeic events, the cessation of, and return to, breathing may cause body and/or limb movement [59]. Consistent temporal association between movement and arousal or apnoea would suggest that actigraphy could predict these events and explain the origin of some movements during sleep. Since arousals during a sleep stage transition likely disturb sleep [52], and arousals to wake may not satisfy the wake scoring criteria (i.e. duration $< 50\%$ of an epoch), predicting arousals and/or apnoea would identify regions of sleep disturbance otherwise unidentifiable by the 30s wake estimates. In addition to predicting these events, the relationship with body and/or limb movements during sleep may negatively impact the wake predictive performance of actigraphy. Indeed, movement associated with these events will manifest during 'sleep' and may be seemingly misclassified as 'wake' with actigraphy. Identifying these movements as 'sleep disturbance', instead of 'wake', would improve the accuracy of sleep estimates (as compared to the gold standard polysomnography), and/or may provide an avenue for identifying signs indicative of sleep disorder severity.

There have been some studies that have investigated physiological markers for arousal detection using sensors that measure muscle activation (electromyogram (EMG)). Mograss et al. [180] analysed the occurrence of movement/arousals recorded with chin and arm EMG and standard polysomnography for 15 children presenting with OSA. Despite an extensive analysis on the prevalence of arousals and their relationship with apnoeas and hypopneas, they did not explore the relationship between arousals and specific movements. Further to this, Mograss et al. [180] only identified movement using chin and arm EMG, which records muscle activation, not necessarily specific movement types. Drinnan et al. [181] evaluated the ability of some physiological markers to detect transient arousals from sleep in 36 adults using raw wrist and ankle accelerometry and left and right tibia EMG. They found that limb movement is somewhat correlated with arousal events and therefore can identify the *occurrence* of arousals. Despite this, they noted poor sensitivity for specifically detecting arousals with movement. This may be due to their method of detecting movement during arousal. 'Arousal movement' was defined as any movement greater than what is 'standard' during sleep (assessed using raw accelerometry data). This definition assumes, possibly incorrectly, that arousals cannot cause small move-

ments, or movements similar to those that occur during sleep. In addition to this, they used uni-axial accelerometry to detect movements, which, as we saw in Chapter 4, cannot detect movements that occur orthogonal to the measurement axis. More recently, O'Driscoll et al. [100] evaluated the agreement between actigraphic measures of sleep fragmentation (identified as the percentage of high-activity sleep epochs and the percentage of surrounding immobility) and the polysomnographic arousal index for 130 children. O'Driscoll et al. [100] found that actigraphy was unable to accurately determine the level of sleep fragmentation when compared to the standard arousal index. However, this study used activity as scored by a commercial system on a 30s basis, and therefore can only detect arousals based on the occurrence or absence of large regions of movement.

There is evidence of an association between some sleep movements and arousal events in literature. However, it is unclear whether this association is specific to a certain subset of sleep movements, and whether these movements can be isolated from other sleep movements and/or from movements that occur during wake. It is also unclear if this association is extended to apnoeic events. Accurately identifying sleep-related movements that are *not* associated with arousals, apnoeas or wake would improve the accuracy of 'sleep' predictions.

This chapter indirectly addresses the second limitation of actigraphy, discussed in Chapter 1:

False positives: sleep epochs with observed movement are incorrectly identified as 'wake'.

This chapter will address the hypothesis:

> *Physiological and pathological events characteristic of sleep disorders (e.g. apnoeas, hypopneas and transient arousals) cause sleep movements that contribute to false wake detections.*

In this chapter we will first explore the temporal association between sleep movements, transient arousals and apnoea. We will explore the impact of arousals on actigraphy-derived estimates of wake, and whether actigraphy can predict these events.

This exploration will be performed in two sections:

Association between movement, arousal and apnoea
Section 6.1 aims to explore the temporal and spectral association between limb/body movements, transient arousal and apnoeic events. This section will explore the duration and percentage of

arousals and apnoeas/hypopneas that coincide with movement, and the duration and percentage of sleep movements that coincide with arousal.

Predicting arousals with movement

Section 6.2 aims to identify the ability of actigraphy to predict transient arousals, and conversely, the impact of arousals on actigraphy-derived wake scores. This section will compare the total wake after sleep onset of actigraphy-derived wake (considering arousal events) as compared to polysomnography scored wake.

6.1 Association between movement, arousal and apnoea[2]

In this section we will investigate the temporal and spectral association between limb/body movements, sleep-related arousal events and apnoea. There is evidence in literature that transient arousals may be accompanied by movement. However, these studies do not compare high-resolution information from all movements during sleep, arousal and wake for children. The literature also does not analyse movement from different locations on the body. We are able to address these limitations and perform these analyses with CMAS.

The specific aims of this section are to:

1. Determine the association between transient arousals and body/limb movement; and
2. Determine the temporal association between apnoea and body/limb movement.

6.1.1 Method

The full 30 participants detailed in Table 3.1.3 were used in this analysis. Each patient underwent the study procedure outlined in Section 3.1.4.

Analysis of association with transient arousal

The methodology for this analysis is summarised in Fig. 6.1. Movement segments were detected from the raw accelerometry for each accelerometer placement using the method outlined in Section 3.2.3. The manually scored arousal label and 30s sleep stage were extracted for each movement (described in detail below). After movements were segmented, blocks of missing accelerometry data within each segment (caused by missing wireless packets discussed in Section 3.1.4) were interpolated. This was performed post-movement detection to ensure that the interpolation procedure did not create artefacts. Segmented movements were stratified according to sleep stage (REM, NREM and wake) and arousal incidence.

Although the American Academy of Sleep Medicine (AASM) alterations to the ASDA criteria (summarised in Section 2.1.2) somewhat improved the inter-rater variability for scoring arousals [183, 50], there is still substantial variability. To account for any variability in scored arousal start and end times, a detected movement was considered associated with an arousal event if it occurred during or within a 2s window before or after the event. Movement was considered to occur during a specific sleep stage if more than 50% of the movement occurred during that stage. Arousals that occurred during a wake transition were not included in this analysis. Two metrics were developed to quantify the temporal association between limb and/or body movements and arousal events: event duration and prevalence. Event duration was assessed

[2]This work has been published in The Journal of Physiological Measurement: "Temporal association between arousal and body limb movement in children with obstructed sleep apnoea" [182]

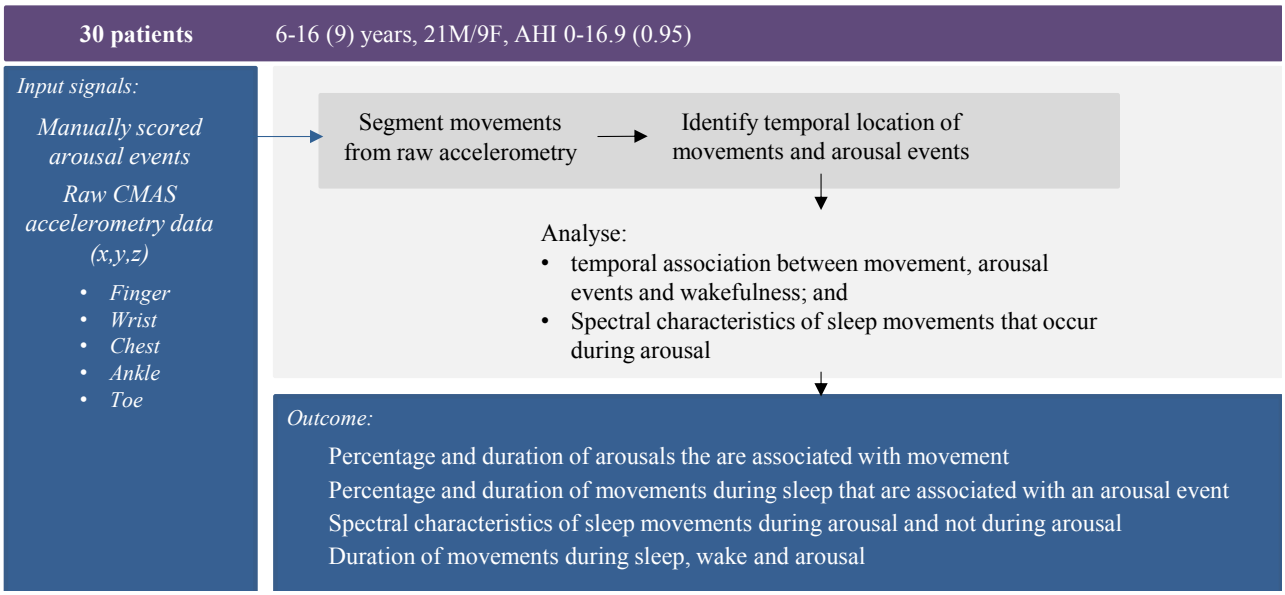


Figure 6.1: Methodology for exploring the characteristics of movement during different arousal events for 30 *patients. Movements and their corresponding arousal event are segmented and the duration and prevalence for each arousal event is explored.*

across the cohort by analysing the median duration for each patient, and the variation of these medians across the cohort (represented using the coefficient of variation). The prevalence of movements associated with arousal was defined as the percentage of movements that occurred during an arousal and also during sleep. Similarly, the prevalence of arousals with an associated movement was defined as the percentage of arousals that had a movement occur within the span of the arousal event. The spectral characteristics were derived for sleep movements that coincided with arousal and sleep movements that did not coincide with arousal using the approach outlined in Section 5.2.

This analysis was performed for the finger, wrist, upper thorax, ankle and toe accelerometers individually, and for all accelerometers combined; i.e. when movement was detected in *any* of the finger, wrist, upper thorax, ankle *or* toe recordings. Since the variability and number of events differed, Welchs t-test was used to compare the median duration distributions for movements that occurred during an arousal and those that did not have an associated arousal event for all patients. For the same reason, Welch's t-test was also used to assess the significance of the spectral characteristics. Finally, the association between arousal events, movement and AHI for this cohort was identified. The analysis determined whether a greater AHI corresponds to a greater number of arousals, or arousal-related movements, using Spearmans rank correlation coefficient.

Analysis of association with apnoea

The full methodology for this analysis is summarised in Fig. 6.3. The manually scored apnoeic events (central, mixed and obstructive apnoeas and hypopneas) were extracted from the

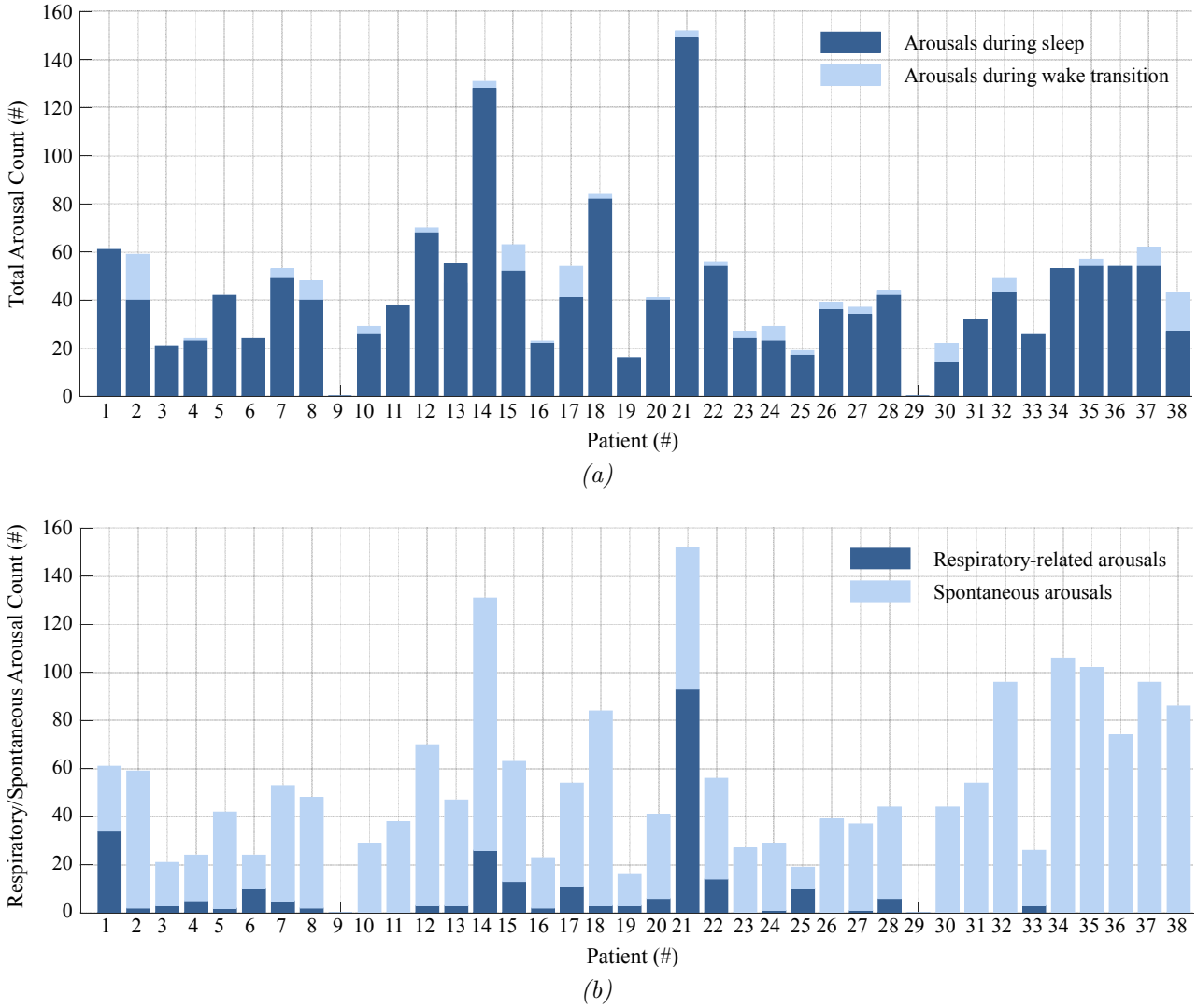


Figure 6.2: Number of arousals experienced by each patient: (a) during sleep (dark blue) and the transition to wake (light blue), and (b) respiratory-related arousals (dark blue) and spontaneous arousals (light blue).

polysomnography files. The associated regions in the raw accelerometry data were then identified for each apnoeic event and a surrounding ±2s region. This dilation in event scoring accounts for any inter-scorer variability in event start and end time. For each apnoeic event, movements within the event region were segmented using the process described in Section 3.2.3. Summary metrics were then determined for each event region (detailed below). The analysis in this section is limited by the number of apnoeic events from each patient (illustrated in Fig. 6.4). Since there is only a small number of apnoeic events (total of 402 across the 38 patients, excluding Patient 21, who dominated the distribution with 122 of the 402 events), this analysis can only be an exploration, and the temporal association between movement and apnoeic events cannot be accurately quantified.

Two metrics were derived to indicate the occurrence of movement within an apnoeic event: the total number of positional changes ΔDC of all movements for all limbs (6.1), and the total

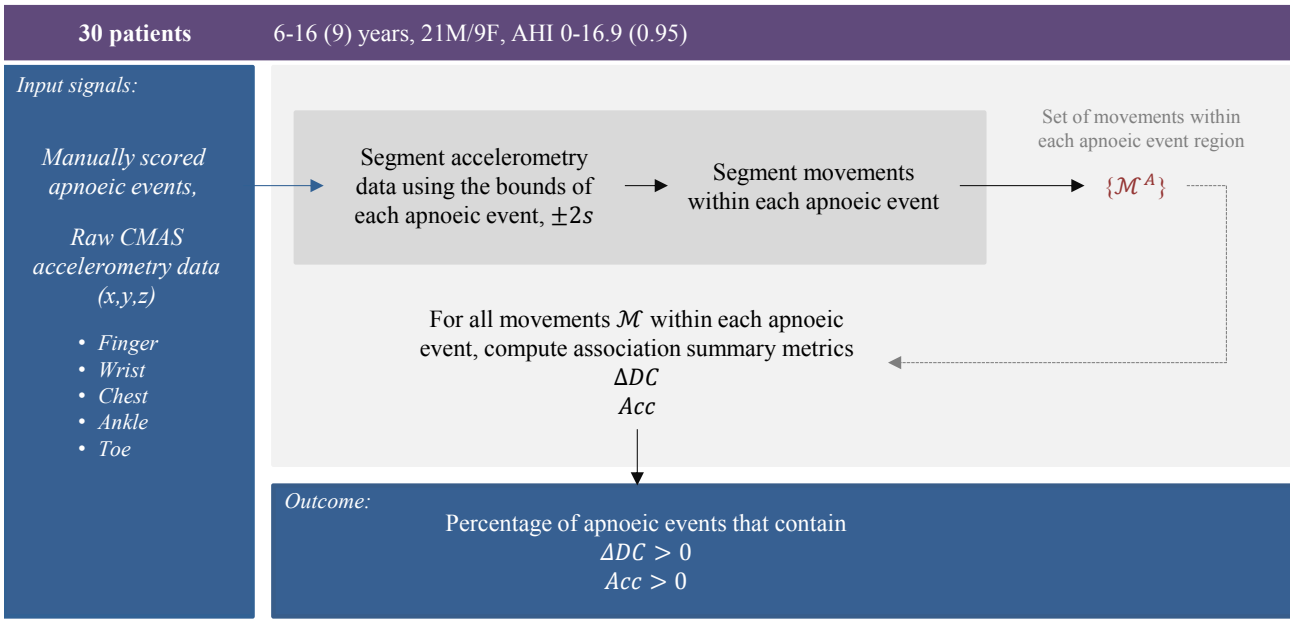


Figure 6.3: Methodology for analysing the association between any movement and an apnoeic event.

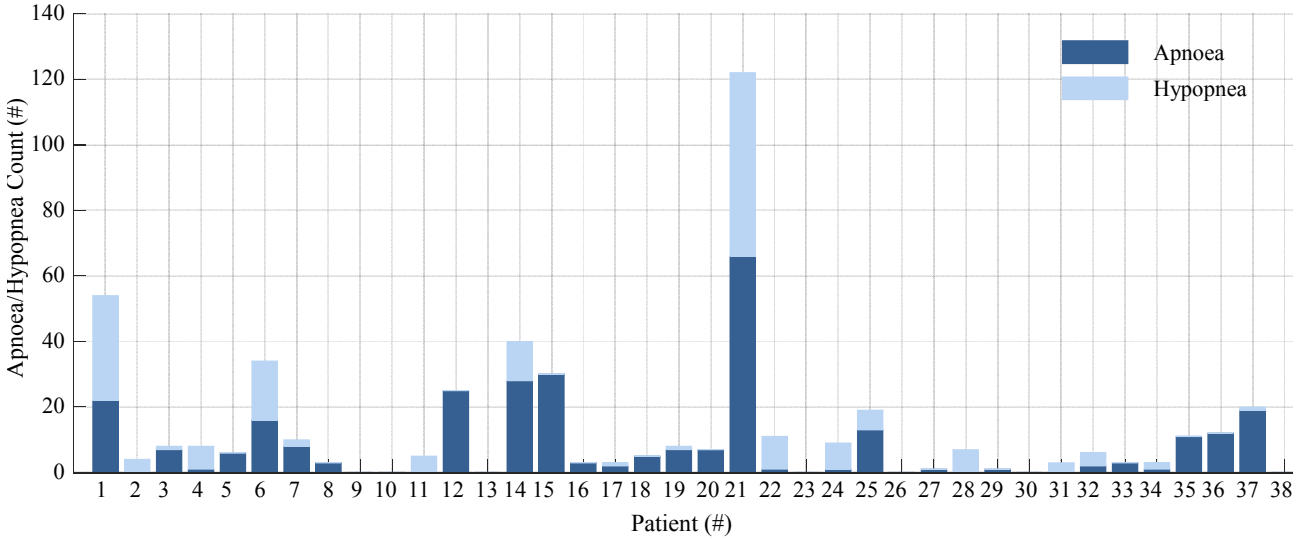


Figure 6.4: Total number of apnoeas (dark blue) and hypopnoeas (light blue) experienced by each patient.

acceleration Acc of all movements for all limbs (6.2) for each apnoeic event E. These values were normalised by the event duration t to ensure that longer events did not bias the metrics. A positional change was defined as occurring within a movement if the first 0.5s of a movement had a different (x, y, z) DC offset than the final 0.5s within a tolerance T. The total number of positional changes for each movement m within an apnoeic event E is described by,

$$\Delta DC = \frac{1}{t} \sum \mathbb{1}\{\|\hat{m}[0, \cdots, 0.5s] - \hat{m}[N - 0.5s, \cdots, N]\| > T\} \quad \forall m \in E, \tag{6.1}$$

where N is the length of the respective movement. The total acceleration for each movement in an event is described by,

$$Acc = \frac{1}{t}\sum_{n}^{N} m[n] \quad \forall m \in E \quad | \quad \|m[n] - \hat{m}\| > T. \tag{6.2}$$

The percentage of apnoeas and hypopneas that contain values of ΔDC and Acc greater than 0 was then analysed. This percentage indicates whether the event coincides with some movement.

6.1.2 Results

Association with transient arousal

On average, 67.5% of arousals were associated with wrist movement (see the median percentage of arousals that coincided with movement '%a' in Figure 6.5). The median percentage of movements that coincided with arousal, '%s' in Figure 6.5, shows that, on average, 17.5% of wrist movements were associated with an arousal. That is, 82.5% of wrist movements during sleep did not occur with an arousal. Other than gender, there was no distinguishable difference between the patient with the highest (patient 31, 96.9%) and lowest (patient 34, 1.89%) percentage of arousals coinciding with movement. Arousal incidence rate varied greatly across all patients, irrespective of AHI. This inter-patient variability was seen for all accelerometer placements, as summarised in Table 6.1 (also see Fig. E.1 through Fig. E.5 in Appendix E). Including movement segments from any sensor increased the percentage of arousals that coincided with movement to 89%. There was no difference in movement prevalence or duration between NREM and REM sleep.

As shown by the median movement duration in Table 6.1 and Fig. 6.7, the duration of both arousal-related wrist movements and wake movements were greater than sleep movements that were not associated with an arousal event (duration of 6.26s and 9.89s vs. 2.35s respectively, $p < 0.01$). As documented in Table 6.1, this was consistent across all accelerometer placements (also illustrated in Fig. E.1 through Fig. E.5 in Appendix E). Arousals with an associated wrist movement were generally longer than arousals without an associated wrist movement (median duration 12s vs. 9s respectively, $p < 0.01$), as shown in Fig. 6.8(left). However, arousals with a greater duration did not correspond with a longer associated wrist movement, as shown by the weak correlation in Fig. 6.8(right).

The spectral characteristics differed for longer-duration wrist movements ($2-10$s in duration) and chest movement ($5-10$s in duration). The greater wavelet coefficients are summarised for wrist movements in Fig. 6.9 and Fig. 6.10 and chest movements in Fig. 6.11. The greater coefficients (highlighted by the lighter blue in the spectrograms) indicate that the magnitude of sleep movements during arousal are generally greater than sleep movements that do not coincide with arousal. There were no significant differences for finger, ankle and toe movements.

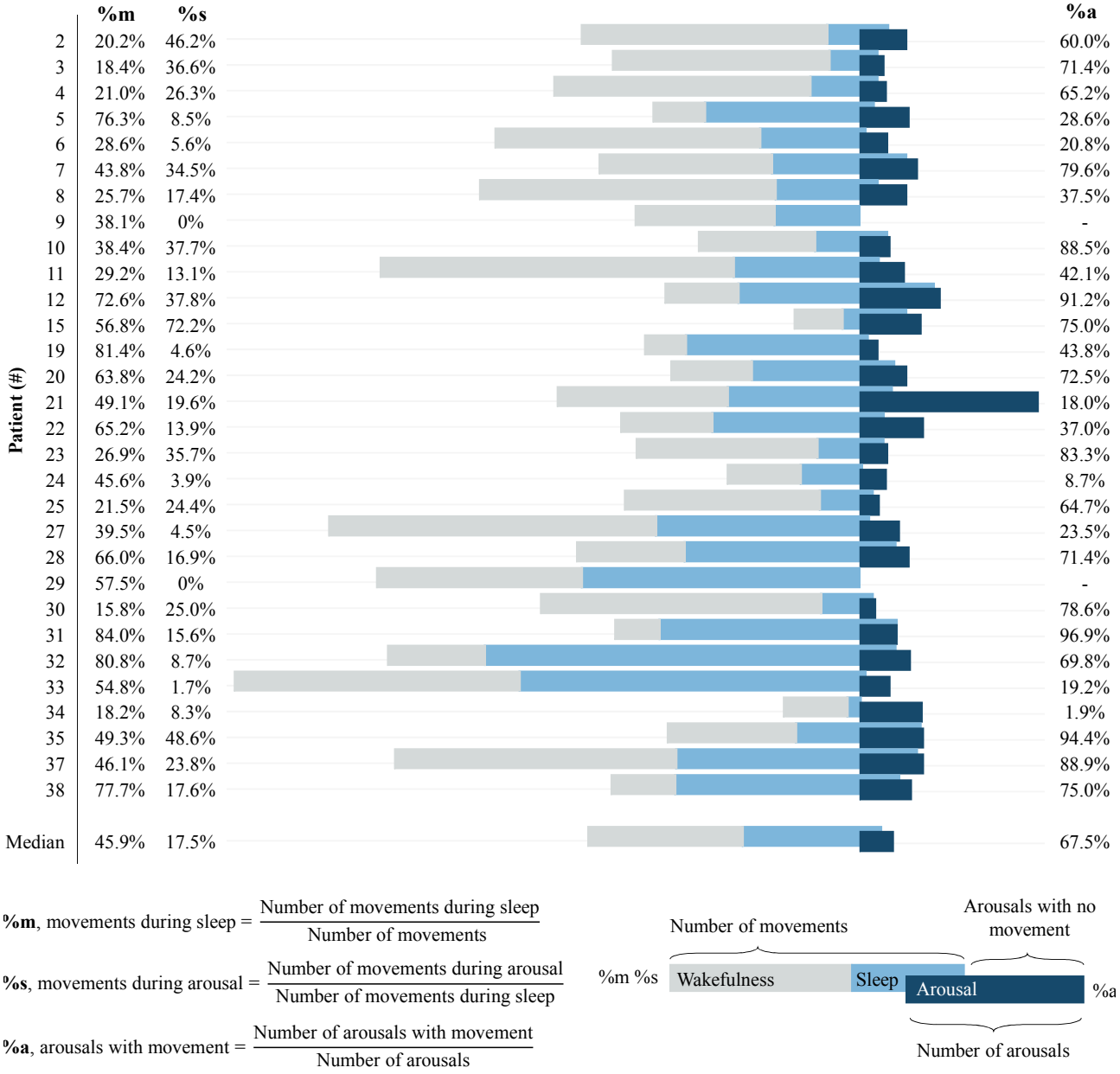


Figure 6.5: Number of wrist movements that occur during sleep, wake and an arousal event, and the total number of arousals across the night for each patient. The percentage of arousals that contain movement %a are indicated by the overlapping regions. The percentage of movements across the night that occur solely during sleep %m and during both sleep and arousal %s are also shown.

The total number of arousals were weakly correlated with AHI ($\rho = 0.19$) and respiratory-related arousals were moderately correlated with AHI ($\rho = 0.72$). Arousal-related wrist movements were mildly correlated with AHI ($\rho = 0.28$), as illustrated in Fig. 6.12. Patient 21, who had an AHI of 16.9, was removed because they were a significant outlier and influenced the correlation ($\rho = 0.27$ vs. 0.19). Patient 6, highlighted in Fig. 6.12, did not have a greater number of arousals, arousal-related movements or movements during sleep, despite having a moderate AHI.

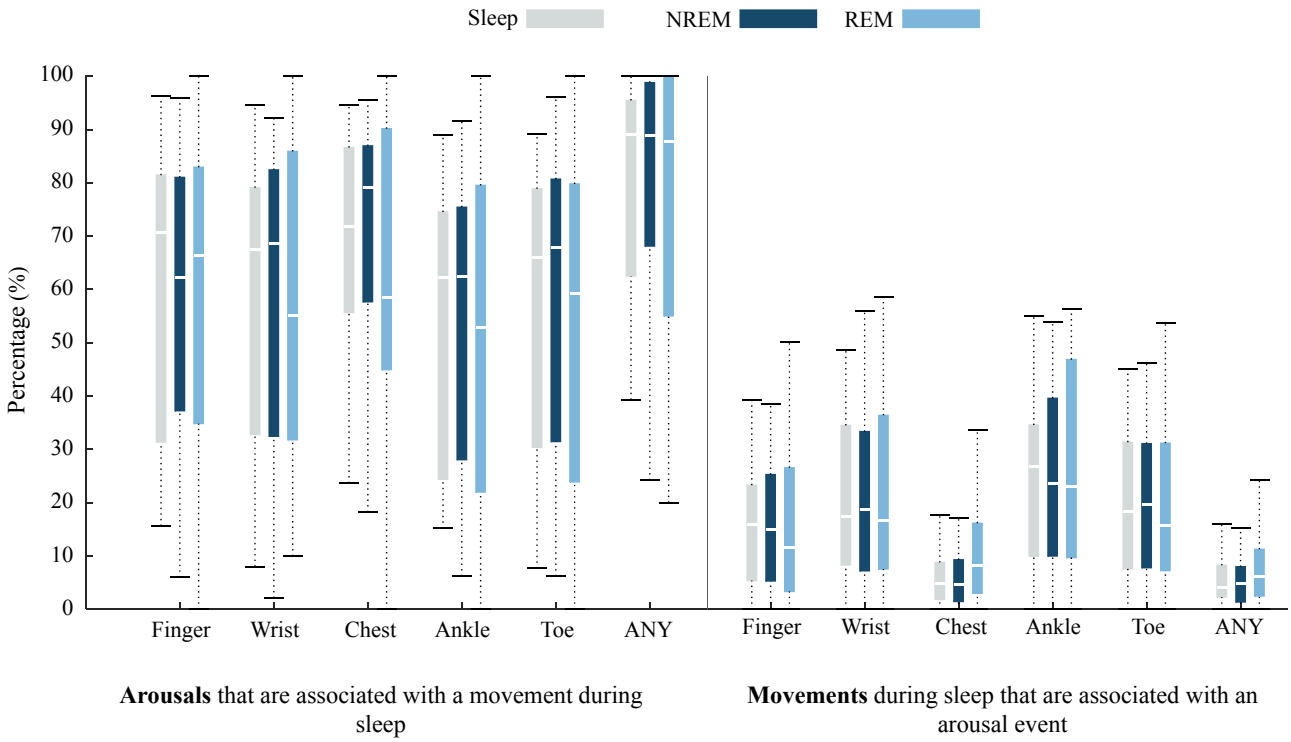


Figure 6.6: Percentage of arousals associated with a movement and movements associated with an arousal for each patient for finger, wrist, chest, ankle, toe and any movement during NREM, REM and sleep.

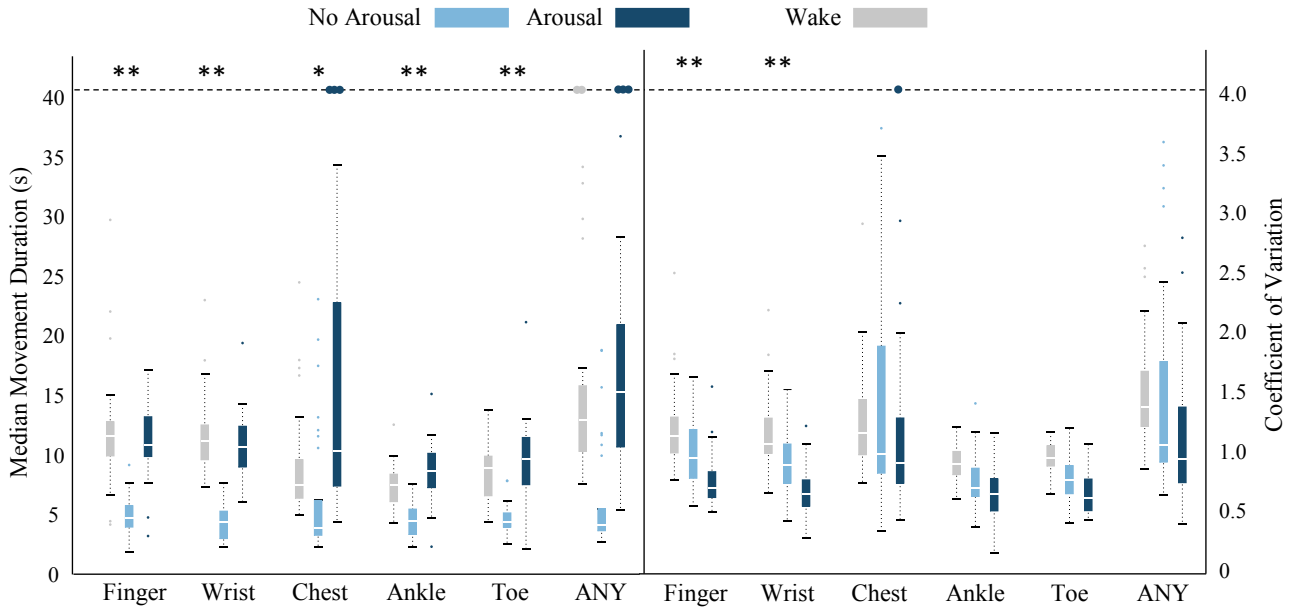


*Figure 6.7: Median duration and the coefficient of variation for the finger, wrist, chest, ankle and toe movement during sleep with and without arousal for all patients (significance shown as *$p < 0.01$, **$p < 0.05$).*

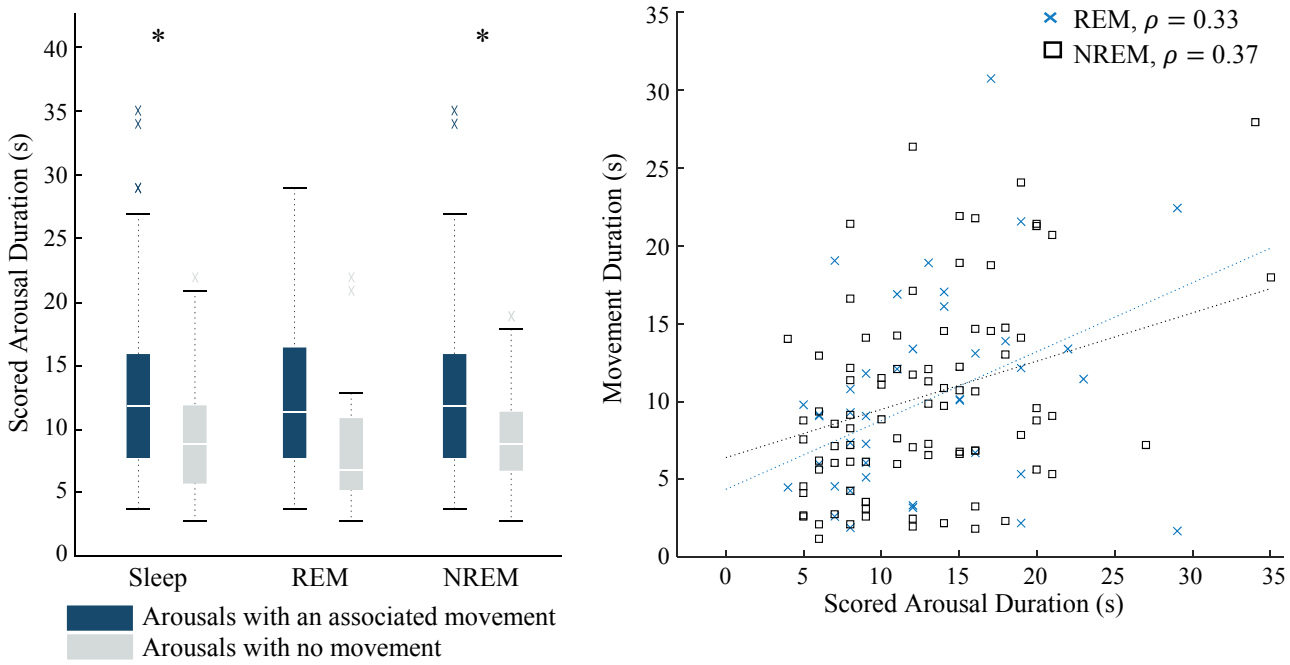


Figure 6.8: Arousal duration with and without an associated wrist movement (left) and the correlation between arousal duration and arousal-related wrist movement duration (right) and (significance shown as $^*p < 0.01$*).*

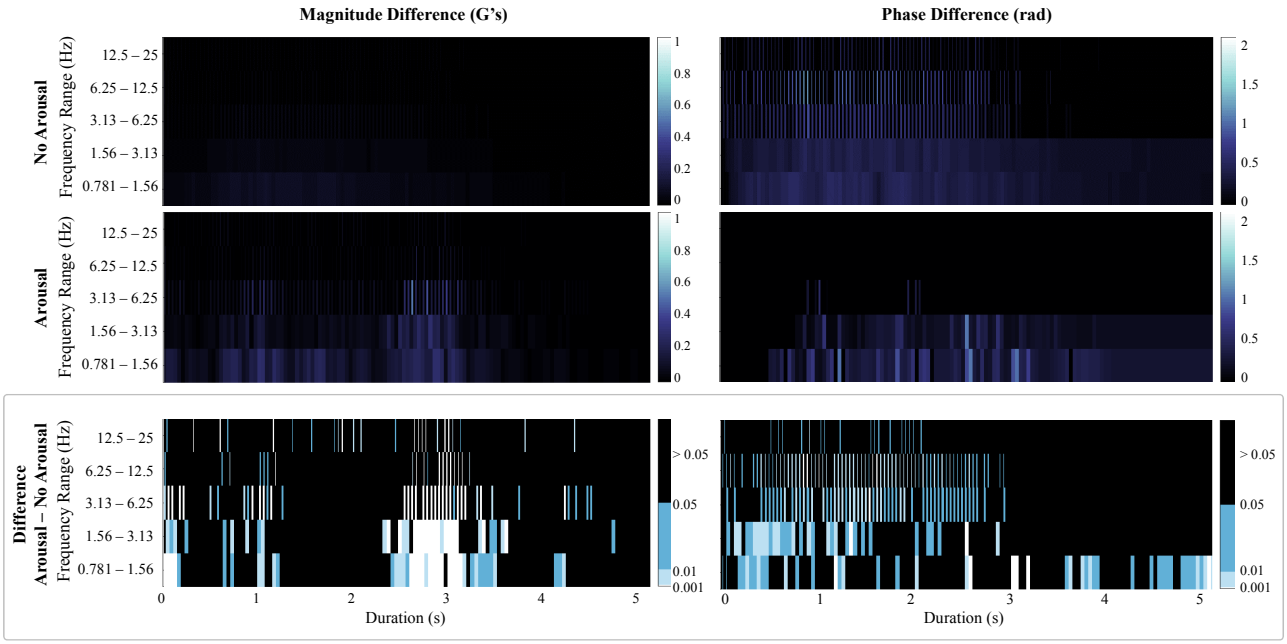


Figure 6.9: Spectral characteristics of wrist movements during sleep that coincide with arousal, and movements that do not coincide with arousal for movements $2 - 5s$ *in duration.*

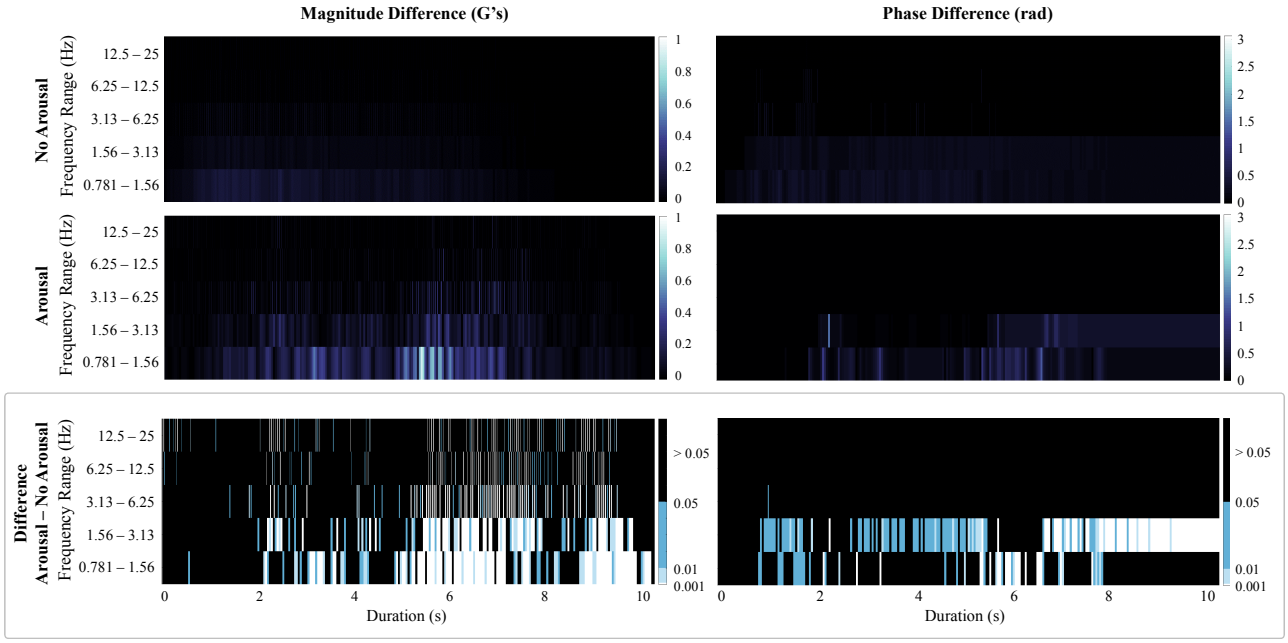


Figure 6.10: Spectral characteristics of wrist movements during sleep that coincide with arousal, and movements that do not coincide with arousal for movements $5 - 10s$ *in duration.*

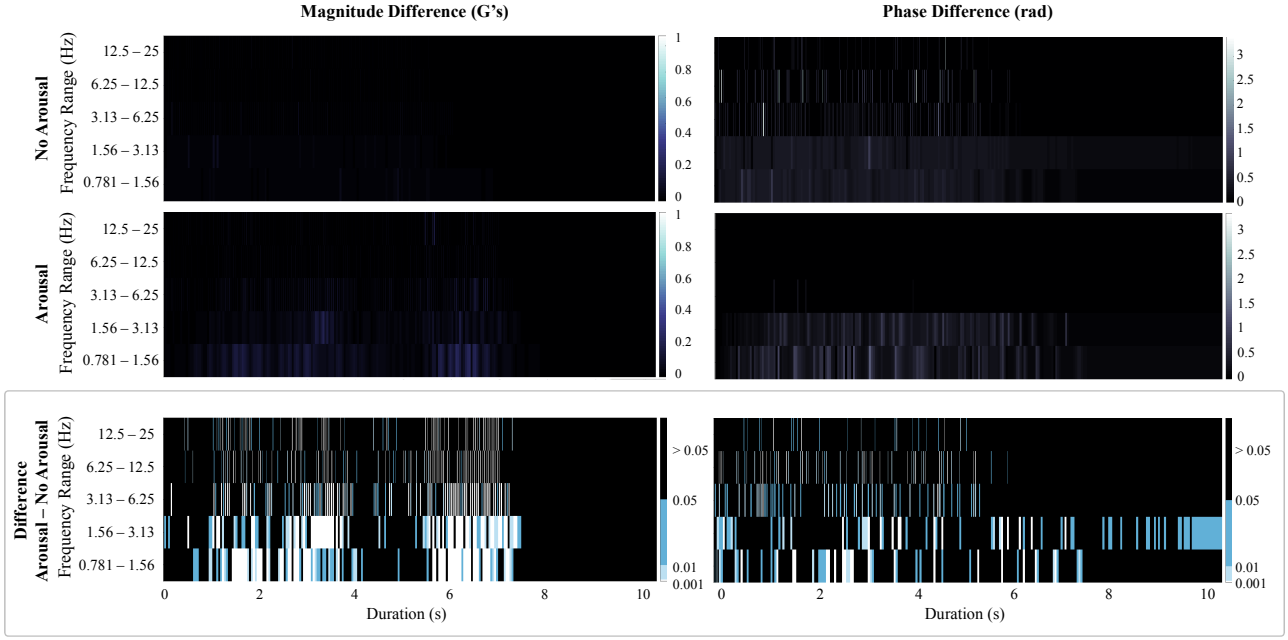


Figure 6.11: Spectral characteristics of chest movements during sleep that coincide with arousal, and movements that do not coincide with arousal for movements $5 - 10s$ *in duration.*

Table 6.1: Prevalence and duration of arousals and movements for the left finger, left wrist, chest, left ankle and left toe during wake, non-REM (NREM) and REM sleep stages

	Finger	**Wrist**	**Chest**	**Ankle**	**Toe**	**All**
Number of Movements per Hour of NREM/REM Sleep						
NREM	15.4 (21.4)	13.6 (18.0)	46.0 (79.8)	8.6 (13.5)	10.9 (17.0)	60.5 (103.1)
REM	13.3 (21.6)	14.0 (16.8)	48.0 (67.9)	5.1 (7.0)	4.7 (11.6)	65.9 (66.1)
Percentage of movements that occur during sleep that also occur during arousal						
Sleep	16.0 (17.9)	17.5 (26.2)	5.0 (6.9)	26.8 (24.8)	18.4 (23.8)	4.1 (6.0)
NREM	15.0 (20.0)	18.8 (26.1)	4.7 (7.8)	23.7 (29.6)	19.6 (23.3)	4.9 (6.6)
REM	11.5 (23.2)	16.7 (28.8)	8.2 (13.1)	23.1 (37.1)	15.8 (23.9)	6.2 (8.8)
Percentage of arousals that have an associated movement						
Sleep	70.7 (50.1)	67.5 (46.3)	71.8 (31.0)	62.3 (50.1)	65.9 (48.4)	89.0 (33.0)
NREM	62.3 (43.8)	68.7 (50.0)	79.1 (29.2)	62.5 (47.4)	67.8 (49.2)	88.9 (30.7)
REM	66.4 (48.1)	55.1 (54.1)	58.5 (45.2)	52.8 (57.5)	59.2 (55.9)	87.8 (45.0)
Median movement duration (s, CV)[b]						
Wake	6.92, 0.33	6.26, 0.26	4.35, 0.35	6.37, 0.89	4.96, 0.23	6.16, 0.38
Sleep, no arousal	2.70, 0.30	2.35, 0.31	2.29, 0.69	2.11, 0.35	2.96, 0.35	2.33, 0.55
Sleep, arousal	9.89[d], 0.41	9.56[d], 0.36	7.46[d], 0.61	5.91[d], 0.46	7.64[d], 0.38	11.03[d], 0.54
NREM, no arousal	2.77, 0.34	2.33, 0.39	2.30, 0.70	2.17, 0.30	2.97, 0.33	2.32, 0.57
NREM, arousal	9.22[d], 0.45	9.20[d], 0.44	7.49[d], 0.62	6.50[d], 0.44	6.72[d], 0.41	10.12[d], 0.57
REM, no arousal	2.72, 0.31	2.35, 0.39	2.22, 0.57	2.22, 0.49	3.38, 0.44	2.43, 0.32
REM, arousal	10.90[d], 0.60	10.23[d] 0.58	7.45[d] 1.02	7.40[d], 0.45	9.02[d], 0.34	12.40[d], 0.79
Median arousal duration (s, CV)[b]						
Sleep, no movement	8.50, 0.35	9.25, 0.29	9.00, 0.23	9.50, 0.28	9.50, 0.29	8.75, 0.15
Sleep, movement	12.5, 0.38	13.4, 0.56	13.0, 0.36	14.8, 0.48	12.5, 0.52	12.3, 0.40
NREM, no movement	9.00, 0.34	9.00, 0.36	10.5, 0.29	9.75, 0.32	9.25, 0.42	8.25, 0.34
NREM, movement	13.8, 0.42	13.8, 0.59	13.8, 0.45	15.0, 0.57	12.5, 0.62	13.0, 0.49
REM, no movement	7.75, 0.16	9.00, 0.35	6.25, 0.08	9.00, 0.32	9.00, 0.28	6.50, 0.36
REM, movement	14.3[c], 0.31	13.5, 0.51	12.3[c], 0.39	14.5, 0.36	15.3, 0.37	12.0, 0.42

[a] Values are shown as median (IQR). Values represent the percentage of movement, arousal, NREM and REM sleep that contain movement with arousal or arousal with movement.
The median prior probabilities for NREM/REM sleep are 78.7%/21.3% respectively.
The median (IQR) number of arousals per hour of NREM/REM sleep are 4.5(2.9)/6.2(3.9) respectively (p = 0.11).
The percentage of movement without arousal and arousal without movement (not shown) is approximately 100% minus movement with arousal and arousal with movement respectively.
[b] Values are shown as median across the cohort (i.e. the median of the individual patient medians) (s), coefficient of variation ($\frac{\sigma}{\mu}$).
[c] $p < 0.05$, [d] $p < 0.05$ Larger movement/arousal duration with arousal/movement compared to without arousal/movement.

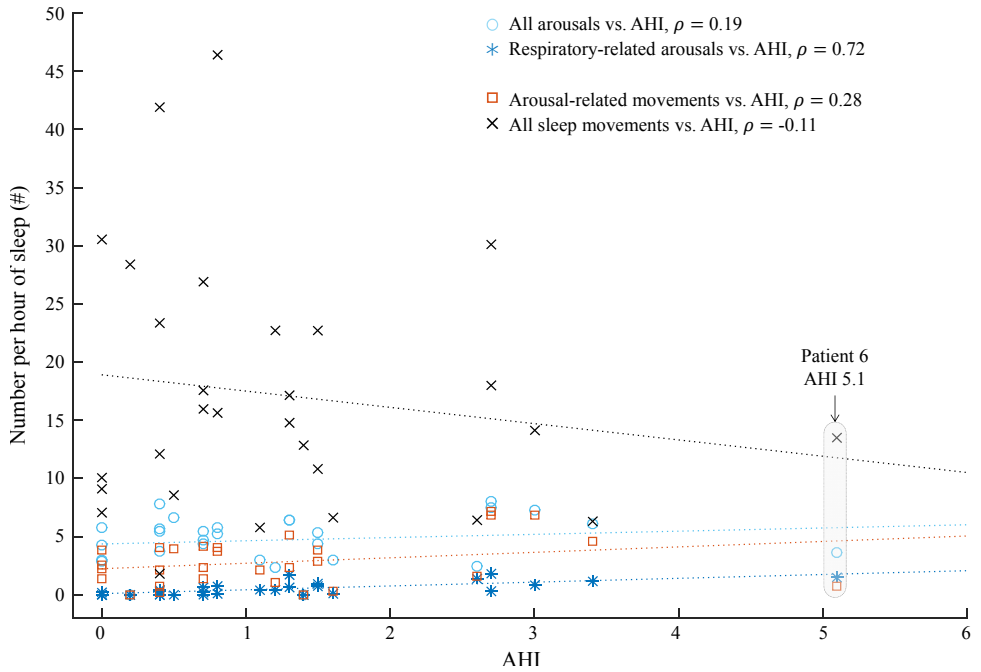


Figure 6.12: Correlation between number of arousals, arousal-related wrist movement and all sleep wrist movements and AHI.

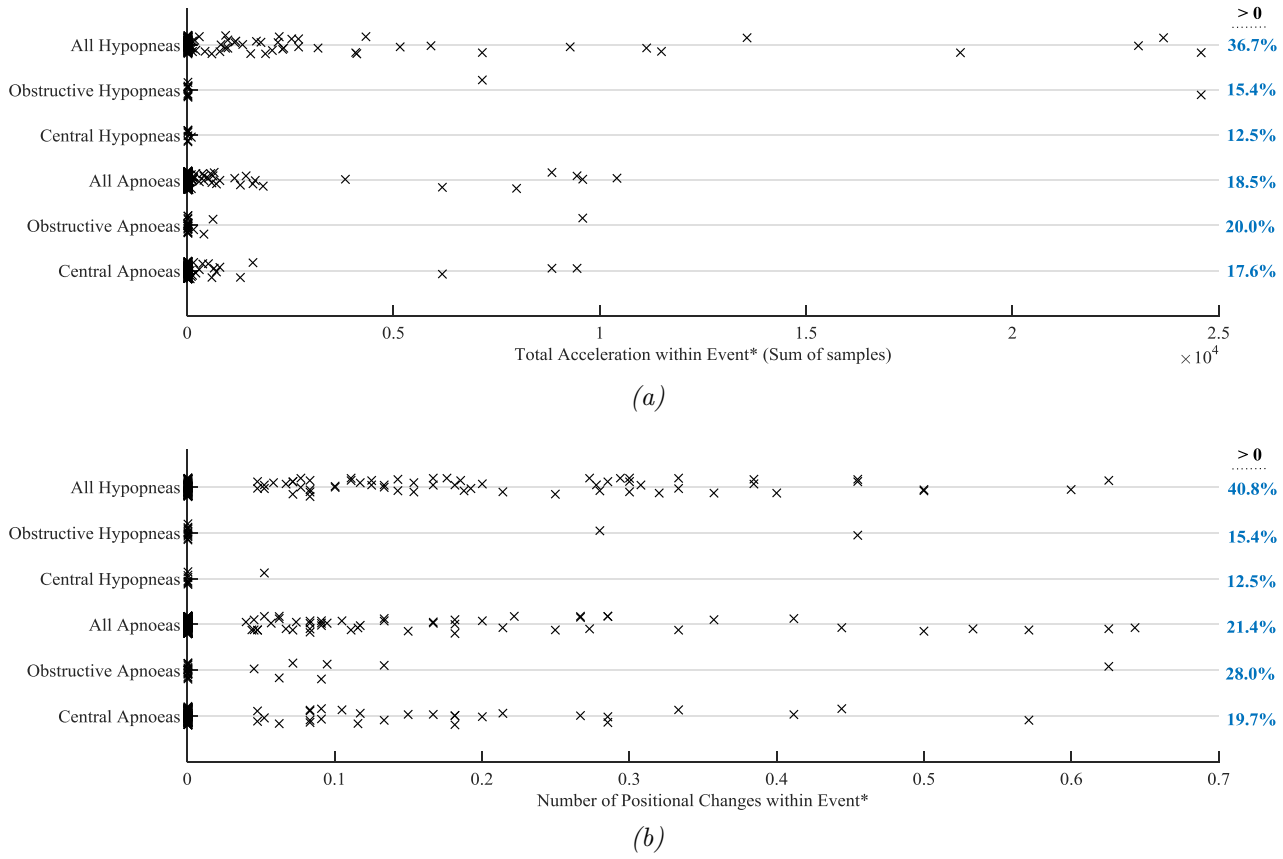


*Figure 6.13: Representation of (a) the total acceleration, and (b) the number of positional changes for each apnoeic event, *normalised by the event duration.*

Association with apnoea

As illustrated by the scatter plots in Fig. 6.13, on average, 36.7% of hypopneas contained vigorous movement and 40.8% contained postural changes. Obstructive apnoeas tended to coincide with more movement than central apnoeas (20.0% vs. 17.6% for vigorous movement during obstructive and central apnoeas respectively, and 28.0% vs. 19.7% for postural changes during obstructive and central apnoeas respectively). Hypopneas coincided more often with movement than apnoeas (36.7% vs. 18.5% for vigorous activity, and 40.8% vs. 21.4% for postural changes).

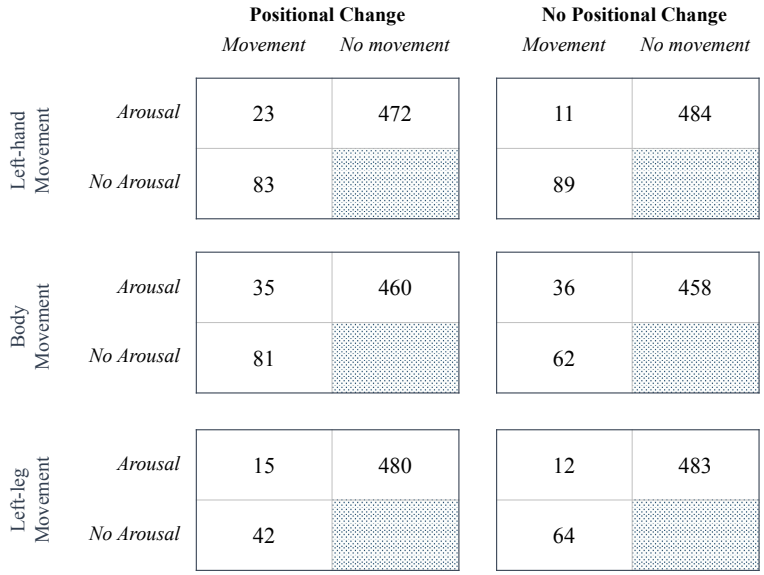
6.1.3 Discussion

This analysis sought to identify the temporal association between transient arousals, apnoeic events and movement to explore the impact on actigraphy-based estimates of sleep architecture. Given the moderate association between arousals and lengthy movements, arousals are indeed likely to confound wake predictions. While the greater duration of arousal-related movements suggests that actigraphy could differentiate arousal-related movements from other movements during sleep, on average 32.5% of arousals were not associated with a measured movement event. Actigraphy, on its own, cannot reliably detect all arousal events; however, since arousals with movement are longer than those without, actigraphy may detect the arousal events that have the greatest impact on sleep quality.

Movement and transient arousal

Body mobilisation, as an attempt to address a potential threat, is often considered a key element of the arousal response [4]. However, on average, only two thirds of arousals coincided with wrist movement. This increased to 89% when including movement detected from any accelerometer (summarised in Table 6.1 and illustrated in Fig. 6.6); however, these arousals only coincided with 4.1% of movements during sleep. Although actigraphy may differentiate some of these lengthier arousal-related movements from the other 95.9% of shorter-duration sleep movements (thereby identifying the corresponding regions of sleep disturbance), the inter-patient variability (illustrated by the large spread of prevalence in Fig. 6.6) suggests that it would only be effective for a subset of patients. Furthermore, this increased sensitivity to arousals with an associated movement requires additional accelerometers, which decreases the percentage of sleep movements associated with arousal. Additional accelerometers may also reduce the usability in a non-clinical setting, and increase the risk of artefacts from patient interference. Regardless, this is still promising and requires further exploration.

Comparing the manually labelled movements of the 10 patients that were used to derive the heuristic in Section 5.1 found that body movements without positional changes coincided the most with arousal. The confusion matrices for the labelled movements and arousal events for



30.2%\36.7% of body movements with\without positional change that coincided with arousal
vs. 21.7%\11% for hand movement with\without positional change; and
26.3%\15.8% for leg movement with\without positional change

7.0%\7.4% of arousals that coincided with body movements with\without positional change
vs. 4.6%\2.2% for hand movement with\without positional change; and
3.0%\2.4% for leg movement with\without positional change.

Figure 6.14: Confusion matrix of labelled movement types and arousal events for 10 *patients.*

the 10 patients are shown in Fig. 6.14. The association is low when compared to the association documented in Table 6.1. It is important to note that these movements are manually labelled and only from 10 of the 38 patients. Comparing the number of sleep movements that coincide with an arousal event to the heuristically identified sleep movements in Section 5.1 (i.e. the application of an observationally derived heuristic to identify sleep movements) across the cohort showed that the majority (88.2%) of arousal-related movements were identified by the heuristic. We saw in this chapter that short-duration movements during sleep were unlikely to occur during an arousal event. Removing the short-duration movement component from the heuristic further increased the percentage of arousal-related movements that were identified by the heuristic to 90.9%. This further supports the hypothesis that arousal events cause body and limb movement during sleep.

Arousals that did coincide with movement were generally longer than those that did not coincide with movement, as illustrated in Fig. 6.8. Similarly, the majority of movements during sleep were not associated with an arousal (only 17.5% of wrist movements). Therefore, in accordance with adult data presented by Drinnan et al. [181], the paediatric data does not support the theory that arousals consistently coincide with body mobilisation. Furthermore, only some movements that occur during sleep periods can be explained by arousals. Nonetheless, movements that were associated with an arousal were significantly longer than other movements during sleep (range of 6s $-$14s vs. 2s $-$7s for wrist movements, as indicated in Fig. 6.7).

Sleep movements that coincide with arousals also generally have greater spectral energy between 0.781Hz and 3.13Hz, particularly for movement magnitude (illustrated in Fig. 6.9). As such, our data does support that there is a sub-group of arousal events that are longer in duration and are associated with significant body/limb movements. In some cases, these extended events may satisfy the requirements of wake classification, but may overlap consecutive epochs, thereby falling short of the scoring requirement of occupying the majority of a single 30s epoch.

Spontaneous arousals have been found to be more common during NREM sleep than REM sleep in children (approximately 3 : 1) [52]. As such, it may be expected that arousal-related movements are also more common during NREM sleep. But, as shown by the similar percentage of arousals associated with movement in Table 6.1, we can observe no difference during NREM and REM sleep for this cohort. There is also no significant difference in the number of any movements per hour of NREM and REM sleep (60.5(103.1) vs. 65.9(66.1)), and the number of arousals per hour of NREM and REM sleep (4.5(2.9) vs. 6.2(3.9)). This finding is similar to Walter et al. [184], who found no significant differences between arousal occurrence within NREM and REM sleep (9.0(0.4) vs. 8.0(0.65)) for 51 children $3-5$ years of age with obstructive sleep apnoea.

Apnoeic events in children often terminate with arousal and/or movement [180], and consequently, children with obstructive sleep apnoea syndrome typically experience an increased number of arousals [6]. It was expected that movement prevalence would be higher for respiratory related arousals than spontaneous arousals. However, of the 118 respiratory-related arousal events experienced by the cohort (excluding patient 21, who was a significant outlier with 93 respiratory-related arousals), only 37% were associated with wrist movement. It was also expected that a greater AHI will coincide with a greater number of arousals, and, by extension, a greater number of sleep movements.

Similar to literature [6, 184], we saw in Fig. 6.12 that AHI was only somewhat correlated with arousals ($\rho = 0.19$ for all arousals and $\rho = 0.72$ for respiratory-related arousals) (Walter et al. [184] found $\rho = 0.5$). AHI was only moderately correlated with arousals that had an associated movement ($\rho = 0.28$). There was no correlation between the total number of movements during sleep and AHI, which is likely due to the large number of movements that are not associated with arousal. Therefore, without ignoring the sleep movements that are not associated with arousal, the total number of movements during sleep is likely a poor predictor of OSA severity. It is important to note that this cohort is biased towards low AHI, with only two patients presenting as moderate (patient 6, AHI= 5.1) or severe (patient 21, AHI = 16.9). As such, care should be taken when generalising these results to children with more severe sleep apnoea.

We have seen in this analysis that movements that were not associated with sleep disturbance (i.e. an arousal event or wake) tended to be shorter in duration (approximate range of 2s - 7s vs. 6s - 14s and 7s - 18s for wrist movements during sleep, arousal and wake respectively). As such, we can speculate that removing these short-duration movements from the accelerometry signal

would reduce the number of sleep epochs with a high activity count and consequently reduce false wake detections. We do not have a sufficient sample size to perform this analysis and the cohort is bias towards patients with symptoms of a sleep disorder; however, a preliminary exploration (detailed in Appendix B.1) found that removing short-duration movements as a pre-processing step on the raw accelerometry data increased the number of sleep epochs with a zero activity count, and decreased the number of sleep epochs with large activity counts. Although this is a promising result, care must be taken when interpreting and generalising the results because of the inherent pre-analysis bias.

Movement and apnoea

The identified relationship between apnoeic events and movement can, at best, provide an indication of the relationship because only 4 of the 38 patients that undertook a sleep study had an AHI greater than 5 (moderate AHI), and only 31 of the 38 patients had scored apnoeic events. Of these patients, the number of events were heavily dominated by Patient 21, who had 30% of the events (122 of the total 402). The distribution of events across the patients is illustrated in Figure 6.4. We can see from the scatter plots in Fig. 6.13 that less than half of apnoeic events coincided with movement. Interestingly, hypopneas coincided more with movement than apnoeas. As these are regions of shallow breathing (as opposed to a cessation of breathing during an apnoea [7]) this might suggest that body movements occur as a response to continued respiratory effort, rather than a distinct cessation of, and return to, breathing. Actigraphy cannot predict all apnoeic events because less than only half of the events contained some form of movement (illustrated by the scatter plots in Fig. 6.13). Therefore actigraphy, on its own, cannot be used to reliably predict apnoeas and/or hypopneas.

6.2 Predicting arousals with movement

In the previous section, we saw that extended arousals coincided with movement. These lengthier arousals may indicate regions of sleep disturbance, where the shorter-duration arousals may not. For this reason, identifying these lengthier events may provide a more detailed representation of sleep quality than the conventional 30s wake scores. In addition to this, in the previous analysis we saw that arousals are likely to confound actigraphy-derived estimates of wake.

In this section we will:

- Identify arousal duration above which actigraphy can accurately predict events; and
- Explore the extent that arousals confound actigraphy-derived estimates of wake.

6.2.1 Method

The full 30 participants detailed in Table 3.1.3 were used in this analysis. Each patient underwent the study procedure outlined in Section 3.1.4. The cohort dataset was separated into a training and test dataset: 15 patients ($2-10$, 15, 21, 27, and $30-32$ in Table 3.1.3, 10 male, median 9 years, range $6-16$ years) formed the training dataset, and the other 15 patients ($11-12$, $19-20$, 22, $23-25$, $28-29$, $33-35$ and $37-38$ in Table 3.1.3, 12 male, median 9 years, range $6-15$ years), formed the test dataset. The test dataset was used to assess detection accuracy in a leave-one-out cross-validation design on patients.

Arousal detection using actigraphy

Actigraphy's ability to detect arousals of different durations was analysed by applying a threshold to a representative metric (i.e. movement duration or the zero-crossing summary value) of segmented tri-axial movements that occur within 2s of an arousal event. Movement duration and magnitude were used to summarise movement information because we identified in the previous analysis that these metrics differ between movements that coincide with arousal and movements that do not coincide with arousal. The process is summarised in Fig. 6.15. The training dataset was used to identify a generalisable threshold that maximised the precision and recall of detecting arousal events greater than a specified duration t_a (between 1s and 30s). Since the arousal events vary in duration, we cannot identify a true-negative event; unlike wake as scored on a 30s basis, there can be no discrete region for which a negative arousal event can be defined. For this reason, the recall and precision is analysed to identify the optimal threshold [185] (these metrics do not require the number of true negatives). Recall is defined as:

$$\text{Recall} = \frac{TP}{TP + FN}, \tag{6.3}$$

where TP is the number of true positives (i.e. arousals that are correctly predicted), and FN is the number of false negatives. The number of false negatives is identified as the number of arousals that do not coincide with a movement segment; 'no movement' (i.e. the value of the summary metric is approximately 0) cannot identify the event as an arousal, since arousals are detected by the representative metric value occurring above a threshold. Precision is defined as:

$$\text{Precision} = \frac{TP}{TP + FP}, \tag{6.4}$$

where FP is the number of false positives. The number of false positives is defined as the number of movement segments that are falsely identified as coinciding with an arousal event.

We will assess the ability to predict arousal events using both movement duration and a summary zero-crossing activity count ZC_M^A for each segmented movement associated with an

arousal. This summary activity value was defined as the activity count ZC_{TRI} normalised for the duration Δt of the movement segment,

$$ZC_M^A = \frac{ZC_{TRI}}{\Delta t}, \tag{6.5}$$

where ZC_M^A represents the zero-crossing value for each movement during arousal. A threshold T_n (varied between 0 and the maximum movement duration or maximum ZC_M^A) was applied to the movement duration and the summarised activity counts ZC_M^A to predict arousal events. These predicted events were then compared to the manually scored arousal events from

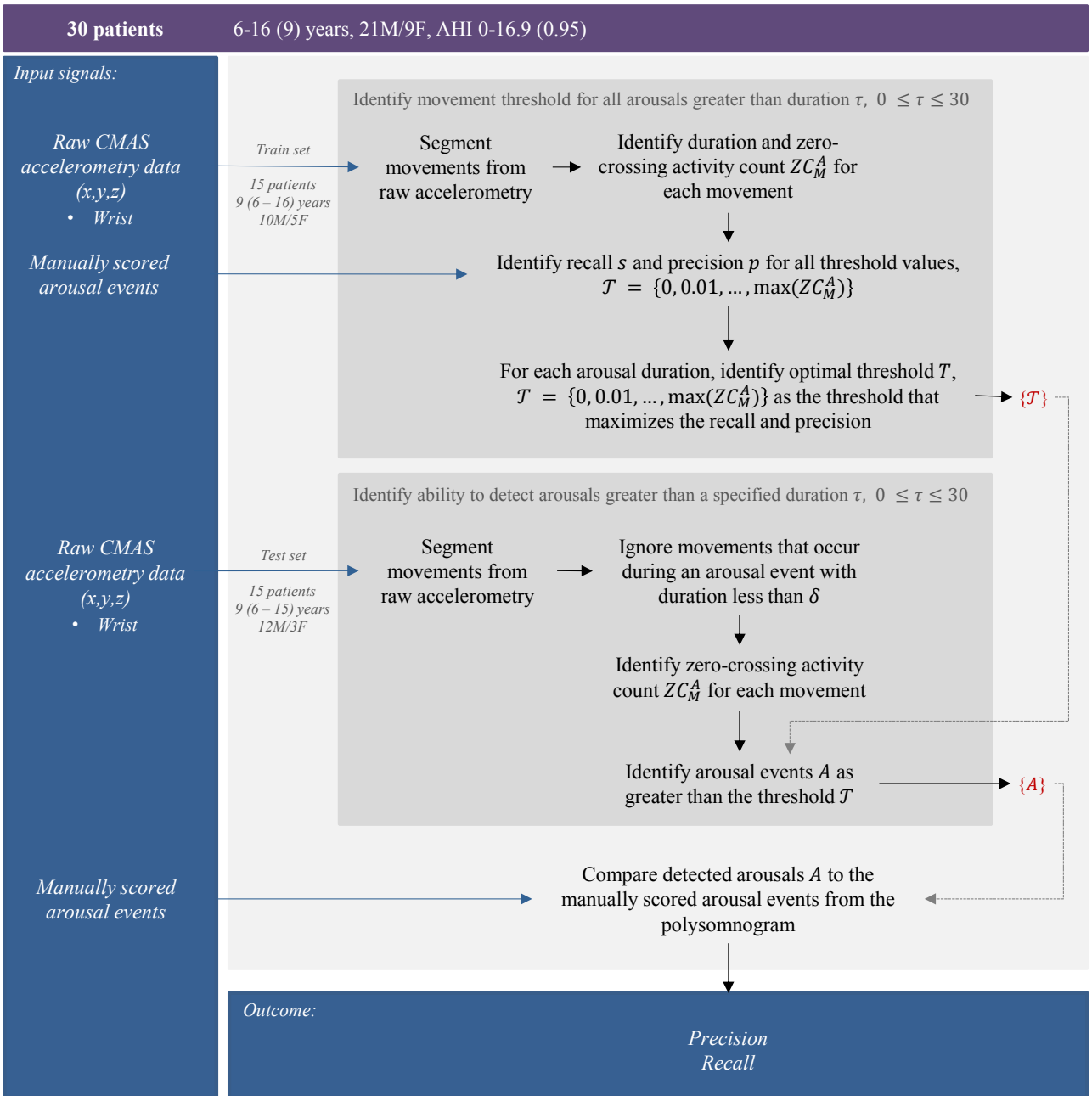


Figure 6.15: Methodology for identifying the ability of actigraphy to predict arousals on the basis of movement duration.

polysomnography. The threshold T_n for each arousal duration was identified as the threshold $T_n \in \{T_1, \ldots, T_N\}$ that gave precision p_n and recall s_n closest to $(1, 1)$ in the precision-recall space. This objective is described by:

$$\mathcal{X}(T_n) = \|[1 - p_n, 1 - s_n]\|. \tag{6.6}$$

The optimal threshold T_n for each movement representation (duration and ZC_M^A) was identified by finding the threshold that corresponds to the minima of $\mathcal{X}$, when $\mathcal{X}$ is applied to all thresholds in $\{T_1, \ldots, T_N\}$. The optimal threshold T_n was then applied to the test data to identify the accuracy of detecting arousal events. This was compared using the precision and recall of detecting arousal events for varying arousal durations. In this analysis, arousal events that are scored on a wake transition, and movements that occur during wake are ignored, since we are only attempting to identify arousal events during sleep.

Impact of arousal on actigraphy-derived wake scores

Given that the extended arousal events are associated with lengthier movements (as seen in Section 6.1), it is likely that these arousals confound actigraphy-derived estimates of sleep and wake when compared to polysomnography. This was assessed by comparing the total duration of polysomnography and actigraphy derived wake after sleep onset (WASO) and arousal duration for each patient. As documented in Fig. 6.16, arousal events $\mathcal{A}$ were first detected from the zero-crossing activity counts of segmented wrist movements ZC_M^A. All arousals were detected using the corresponding threshold in training. The movements that were associated with an arousal were set to NaN in the raw accelerometry data, prior to generating the 30s zero-crossing activity counts ZC_M^S for wake detection. ZC_M^S is defined similarly to ZC_M^A; however, it represents the zero-crossing value for each movement during sleep and wake.

Regions of wake $\mathcal{W}$ were identified as any ZC_M^S greater than a threshold T_S. The threshold was identified as the threshold that gave the greatest Kappa agreement in ROC analysis on the training dataset. The conventional zero-crossing smoothing filter (see Section 2.3.1) was applied to the data, prior to ROC analaysis. The total duration of arousal events τ_{acti}^A and WASO τ_{acti}^W was then identified by summing the duration (indicated here as $\mathcal{T}(\cdot)$) of all detected arousals $a[i]$ and detected wake scores $w[i]$ respectively. That is,

$$\tau_{acti}^A = \sum_{i=1}^{N_A} \mathcal{T}(a[i]), \qquad \mathcal{A} = \{a[i], a[i+1], ..., a[i+N_A]\} \tag{6.7}$$

and

$$\tau_{acti}^W = \sum_{i=1}^{N_W} \mathcal{T}(w[i]), \qquad \mathcal{W} = \{w[i], w[i+1], ..., w[i+N_W]\} \tag{6.8}$$

where N_A and N_W are the total number of detected arousals and wake scores respectively. The total actigraphy derived arousal and WASO, τ^A_{acti} and τ^W_{acti}, were then compared to the polysomnography derived arousal and WASO, τ^A_{PSG} and τ^W_{PSG}. These were derived from the manual scores, documented in the polysomnography files. Once derived, the effect of arousals on actigraphy scores, relative to the polysomnography, was analysed by comparing WASO to polysomnography,

$$\tau^W_{acti} \simeq \tau^W_{PSG}.$$

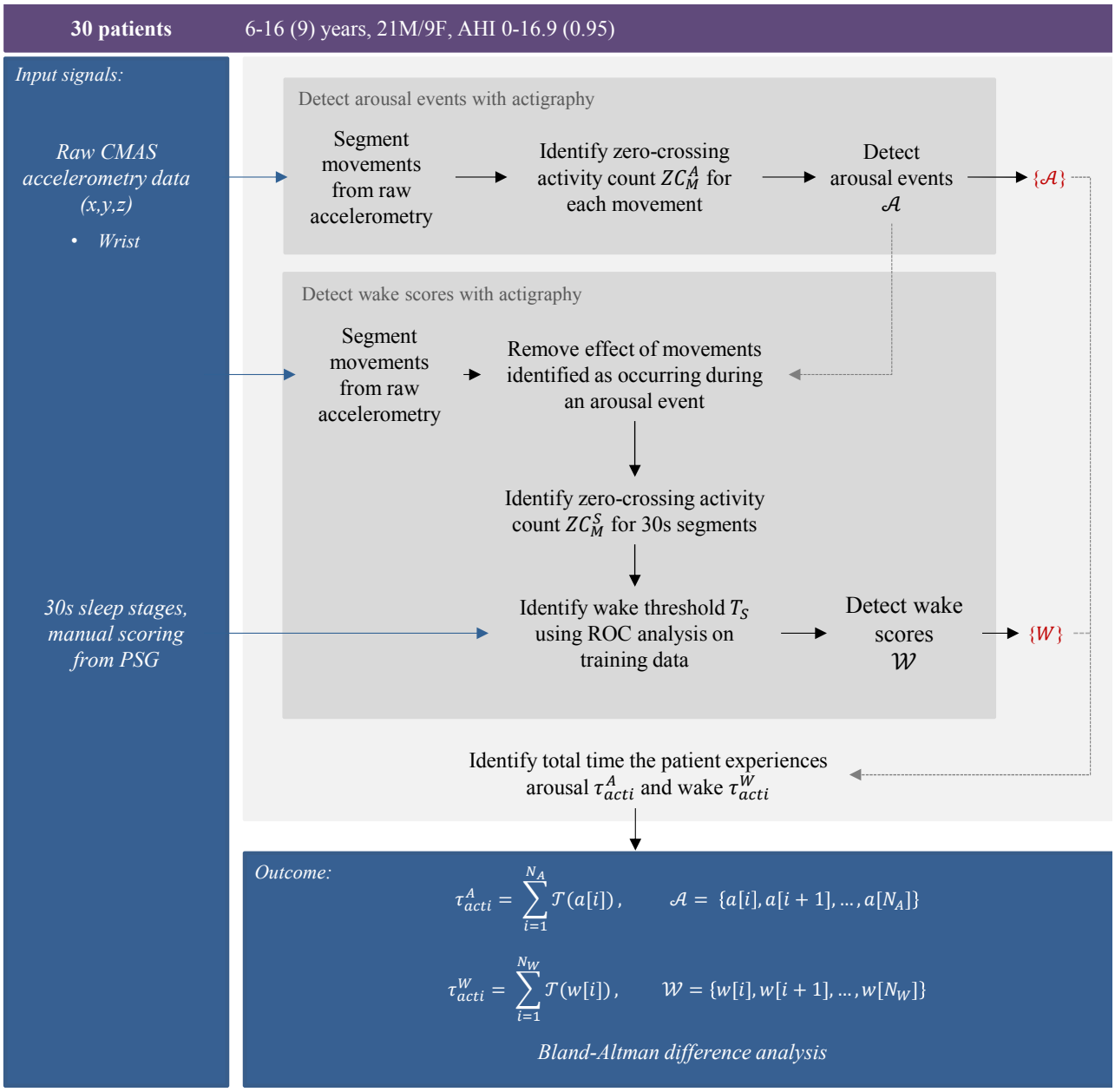


Figure 6.16: Methodology for identifying the total arousal and wake duration as detected with actigraphy to the events that are manually scored using polysomnography.

This was compared to the combination of actigraphy derived arousal duration and WASO,

$$\tau^{A}_{acti} + \tau^{W}_{acti} \simeq \tau^{W}_{PSG}.$$

The polysomnography and actigraphy-derived WASO and arousal total duration were normally distributed and had similar standard deviation. For this reason, the two-sample t-test was used to assess the significance of the estimation difference from the Bland-Altman plot.

6.2.2 Results

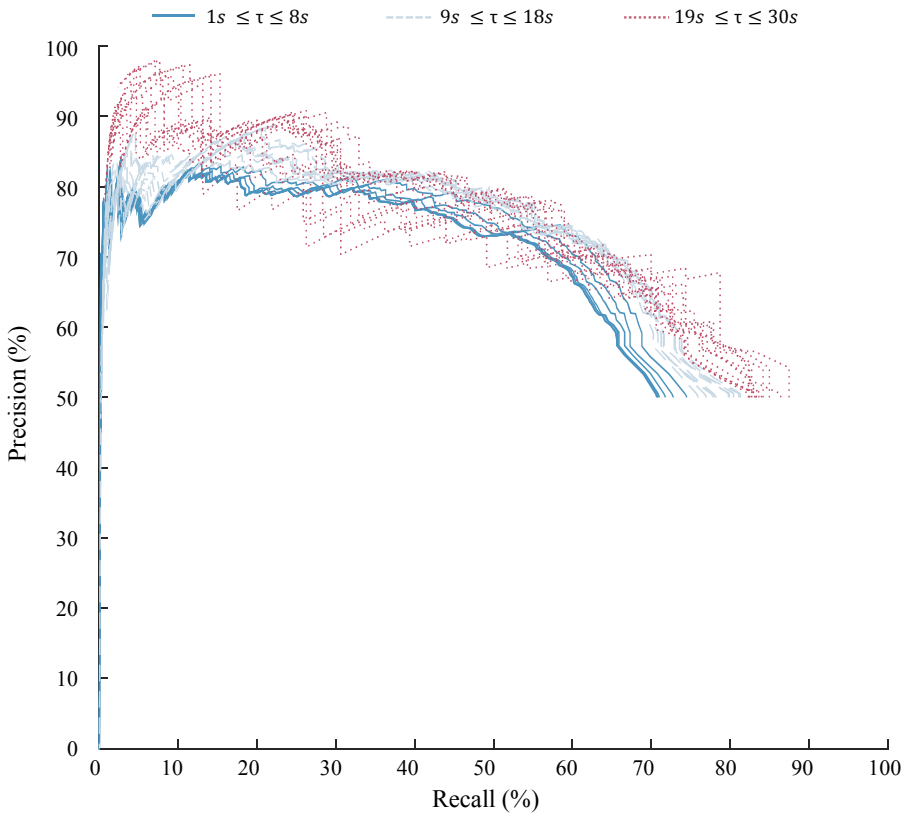
The ability of actigraphy to identify arousal events *greater than* a specified duration is shown in Fig. 6.17a for varying threshold values of movement duration and Fig. 6.17b for varying threshold values using the zero-crossing summaries ZC^{A}_{M}; threshold values to the left of the figure are larger than those to the right of the figure. Movement duration achieved approximately 80% precision; however, this was at a low recall (approximately $10-20\%$, illustrated in Fig. 6.17a. The precision decreased to 50% with an increase in recall. The zero-crossing summary representation achieved approximately 70% precision for up to 60% recall (as illustrated in Fig. 6.17b).

The zero-crossing movement summary was better able to predict arousals greater than a specified duration (illustrated by the comparative recall plots in Fig. 6.18a and Fig. 6.18b). Movement duration had a greater recall; however, the precision was less than the zero-crossing summary representation. The recall and precision of arousal detection in Fig. 6.18 illustrates that movement information was best able to detect arousals of duration between 8s and 16s. After approximately 16s, the precision consistently decreased.

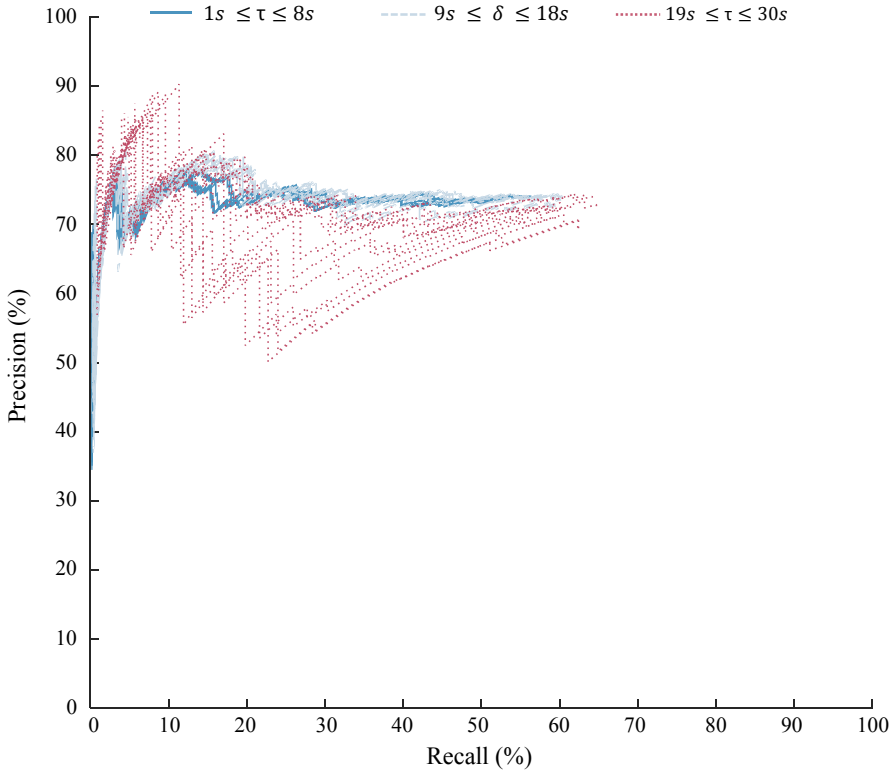
The difference between polysomnography derived estimates of total arousal duration and estimates derived from actigraphy are illustrated in Fig. 6.19a. Including actigraphy-derived arousal events with actigraphy-derived WASO in Fig. 6.19b tended to decrease the difference with polysomnography WASO estimates (from 12.58 mins, $p = 0.76$, to 3.84 mins, $p = 0.31$).

6.2.3 Discussion

In general for all arousal events, the predictive performance increased as we decreased the threshold (illustrated in Fig. 6.17b). This indicates that actigraphy is most effective with high sensitivity to arousals. Actigraphy was able to effectively detect the arousal events that coincided with movement (as evidenced by the 70% recall in Fig. 6.17b). Indeed, we saw in the previous analysis that only 67.5% of arousal events (averaged across the cohort) coincided with wrist movement (illustrated in Fig. 6.5); 32.5% of arousals did not coincide with movement. Actigraphy will not be able to identify these arousals, resulting in false negatives that will impact the recall.

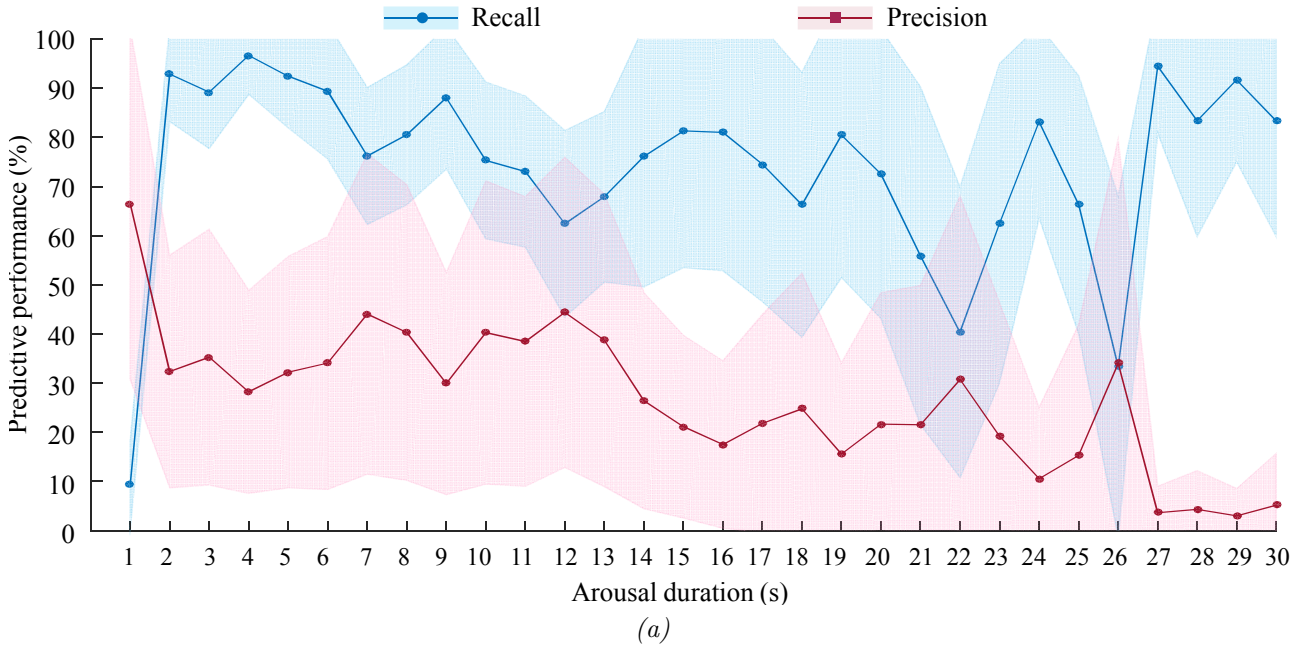


(a)

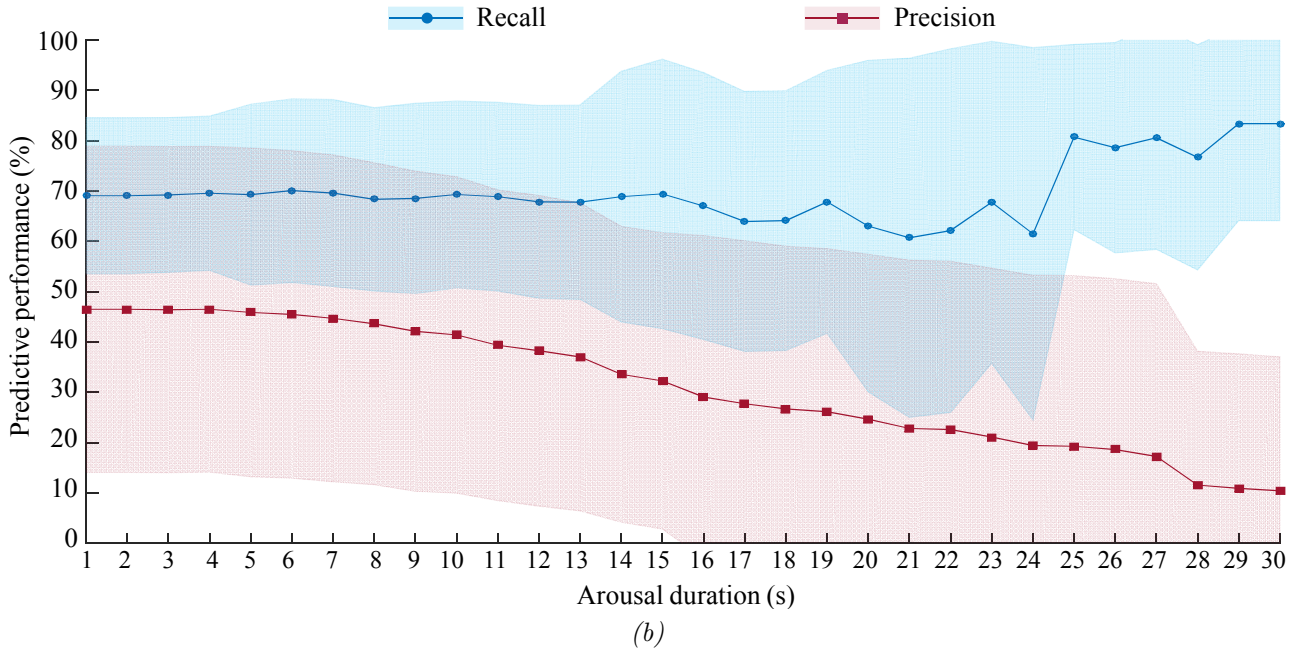


(b)

Figure 6.17: Ability to detect arousal events greater than a specified duration for varying threshold values using (a) movement duration, and (b) zero-crossing summary of movement. Dark blue lines represent short-duration arousals (< 8s), light blue lines represent medium-duration arousals (between 9s and 18s) and red lines represent lengthy arousals (> 19s).

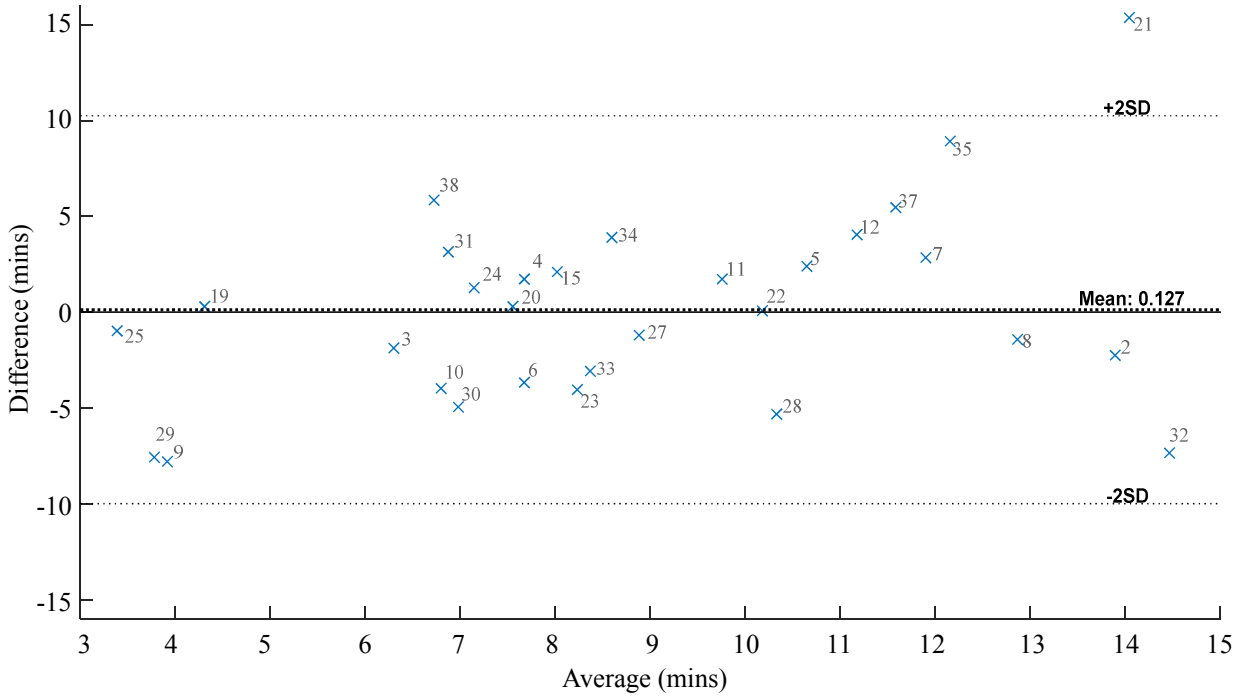


(a)

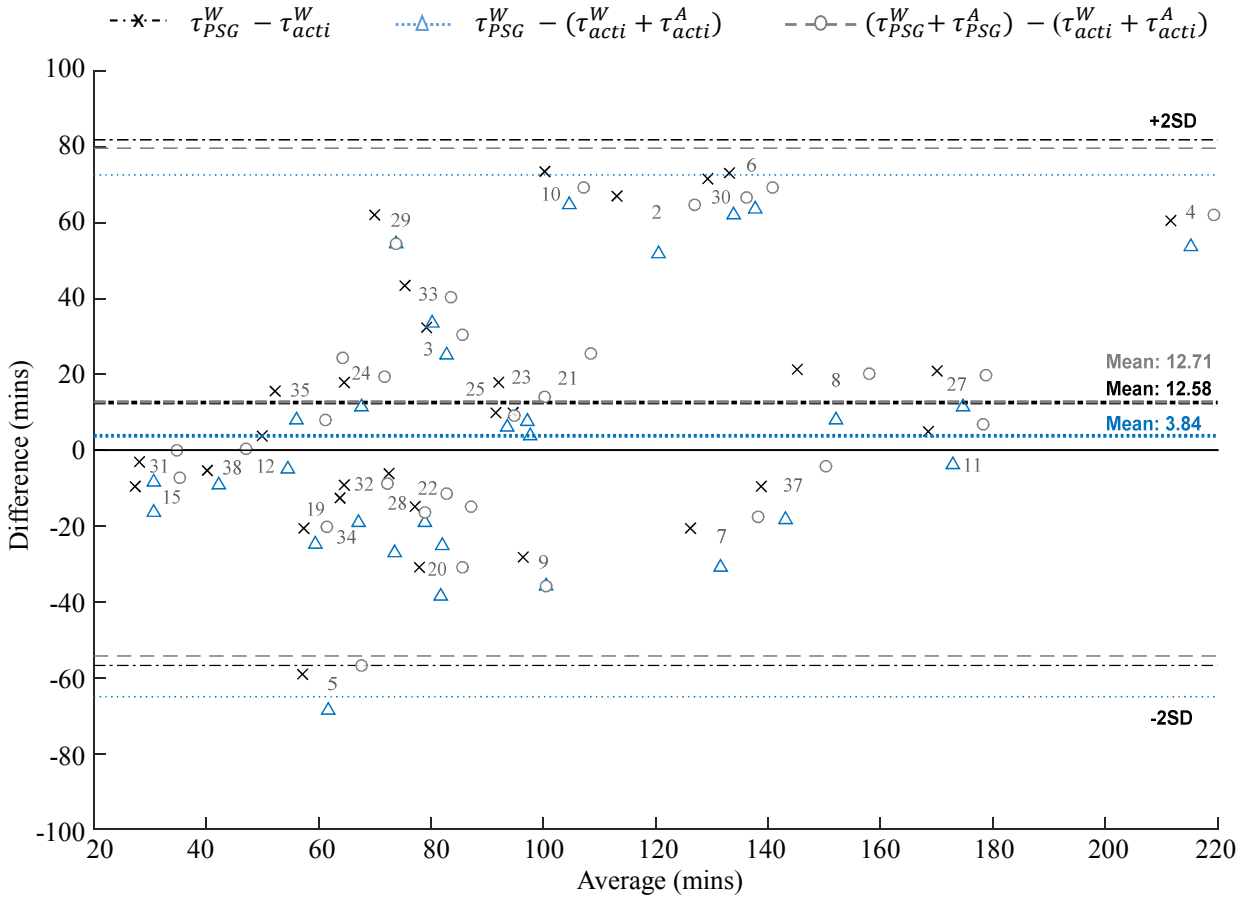


(b)

Figure 6.18: Mean recall (blue) and precision (red) of detecting arousals greater than a specific duration (between 1s and 30s), using (a) movement duration, and (b) zero-crossing summary of movement. The shaded region represents ±1 standard deviation.



(a)



(b)

Figure 6.19: Bland-Altman plot showing (a) the difference between the total duration of polysomnography derived arousal events and actigraphy derived arousal events for each patient, and (b) the difference between the polysomnography derived WASO, actigraphy derived WASO and the combination of polysomnography derived WASO and arousal events, and actigraphy derived WASO and arousal events.

Actigraphy predicted all arousal events better than when restricting the prediction to only longer arousals. This can be seen by the poor performance for arousals greater than 19s in Fig. 6.17b (illustrated by the varying red region). We can also see that the thresholds performed better for arousals greater than 8s until approximately 16s. The average recall for arousal detection was fairly consistent across all arousal durations until 25s, where it increased by approximately 10% (shown in Fig. 6.18b). The increased recall indicates that there were less false negative detections when predicting arousals greater than 25s in duration. Supporting what we saw in Section 6.1, this suggests that arousal events greater than 25s coincide more with movement than shorter duration arousals; however, the precision was low $(10 - 20\%)$ for these arousals, indicating a larger number of false positives than shorter arousal events. This is as expected because there are significantly less arousals of longer duration (summarised in Fig. 6.20), but there are still the same number of movements during sleep; the ratio of arousal-related movements to sleep movements will reduce as we move the arousal duration threshold towards longer arousals.

Although the average arousal duration difference was low across the cohort (approximately

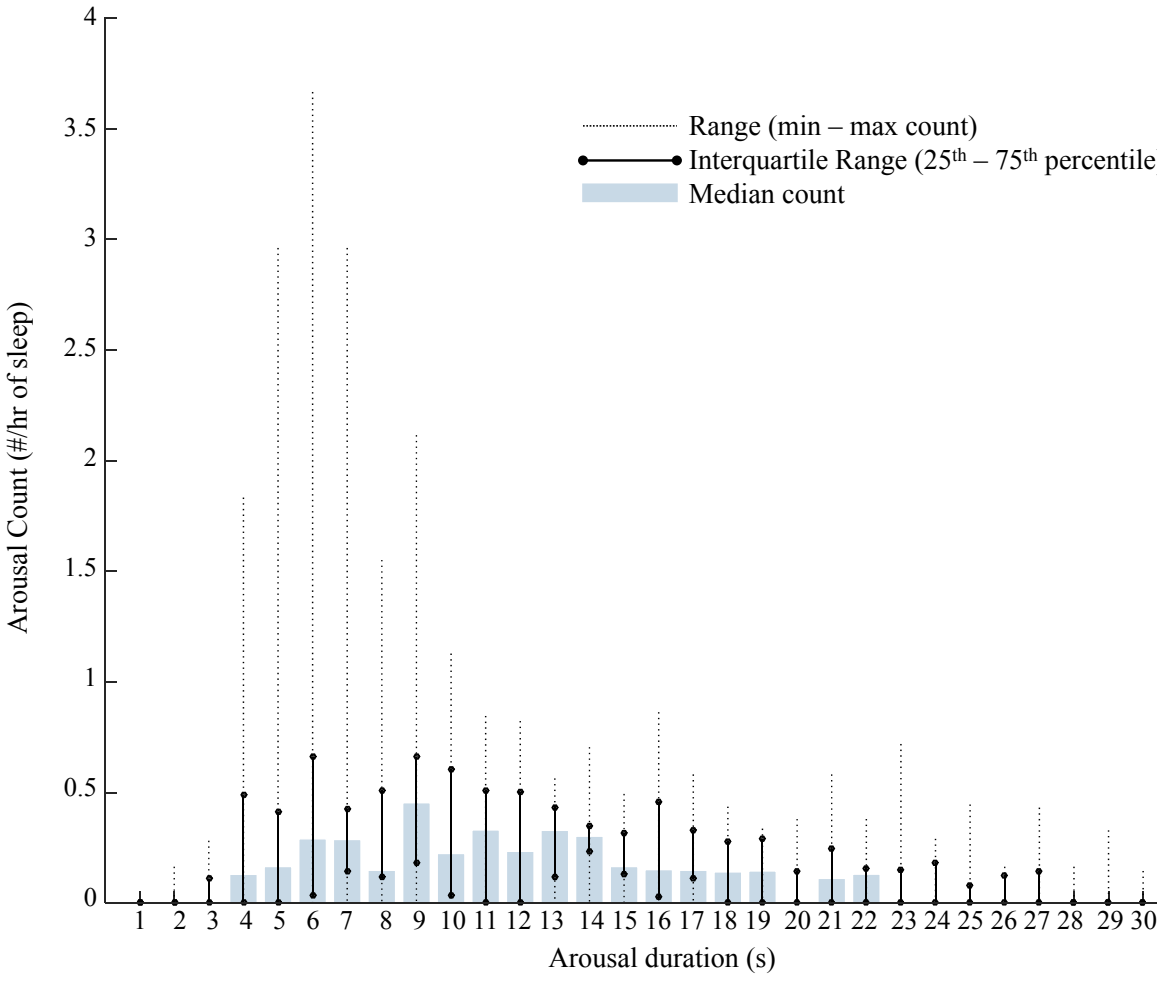


Figure 6.20: Median number of arousals of each duration ($0 \leq \tau_{\leq} 30s$) across all patients.

0.127 mins or 7.6s, illustrated in Fig. 6.19a), the estimates varied considerably between patients (with a difference standard deviation of 5 mins), particularly when considering that the average total duration of arousal events was $7 - 11$ mins for the cohort. The large variance is in accordance with the limitations of wake detections; actigraphy is unable to accurately estimate the *duration* of specific events; events derived from movement information cannot detect regions without movement unless estimated with a smoothing filter. However, as we saw in the previous analysis, actigraphy can effectively detect the *occurrence* of an event. Including actigraphy-derived estimates of arousal duration with WASO estimates did improve the difference between polysomnography and actigraphy derived WASO estimates (from an average of 12.6 mins to 3.8 mins, as illustrated in the Bland-Altman plot in Fig. 6.19b). This improvement is likely due to increased regions of apparent 'wake'; often wake will occur with no activity, and these will be misclassified as 'sleep', reducing the total estimated wake time. Therefore, including movements that normally occur during sleep, but are not large enough to register on a 30s epoch scale, increases the total estimated wake time. This processes acts similarly to the moving average filters discussed throughout this thesis (detailed in Section 2.3.1).

6.3 Summary

One limitation of actigraphy is its poor sensitivity for detecting wake [84]. Section 6.1 quantified the temporal association between apnoea, arousals and body and/or limb movements highlights two mechanisms leading to this limitation. We saw that not all apnoeic events can be predicted with movement information; less than half of the apnoeic events coincided with any movement. We also saw that lengthy arousal events are associated with lengthy movements. These movements are likely to contribute to a large activity count in the associated epoch, and consequently may be misidentified as wake using the standard movement quantification techniques employed in commercial actigraphy [84]. However, these incorrectly identified epochs of wake are being correctly identified as periods of sleep disturbance. From this we can conclude that the typical epoch-by-epoch agreement rates between polysomnography and actigraphy are likely a pessimistic estimate of actigraphys ability to quantify sleep disturbance. The second mechanism that may lead to actigraphys poor sensitivity is the numerous short movements during sleep that are not associated with an arousal event. These movements may cumulatively contribute to a significant activity count, which would be misidentified as wake using conventional actigraphy. But, the data demonstrates that these movements are significantly shorter than arousal-related movements. Therefore, developing a scoring algorithm that explicitly removes these short movements from analysis may significantly improve the limitations discussed above, and consequently improve the accuracy of sleep/wake scoring using actigraphy.

A key objective here is to determine whether actigraphy can detect arousal events. Although the greater duration of arousal-related movements suggests that actigraphy may be capable of

differentiating these movements from other movements during sleep (and thereby identify the corresponding arousal events), the number of arousal events with an associated movement varied greatly across all patients. As such, actigraphy cannot be used to reliably predict all arousal events. However, as arousal events without an associated movement tend to be shorter, it may be possible to detect lengthy arousal events and thereby estimate the severity of sleep disturbance. Indeed, literature suggests that longer arousals are a significant contributor to sleep disturbance [5], and that the incidence rate or length of arousal likely determines the extent of daytime dysfunction caused by sleep disruption [186].

The performance of predicting arousal events of different durations and the accuracy of estimating total arousal and wake duration was explored in Section 6.1. We saw that although we can detect the arousal events that coincide with movement with high recall, there will be many false arousal detections; we can only discriminate arousal-related movements from other sleep movements with, at best, 50% accuracy (summarised by the precision in Fig. 6.17b). We also cannot estimate the duration of arousal events. Incorporating arousal duration with wake scores does improve estimates of total time spent awake, however this is likely acting in a similar way to standard smoothing filters. This is likely due to re-identifying what was classed as 'sleep' movements to 'wake/arousal' movements.

In this chapter we saw that:

- On average, approximately 67% of arousal events coincided with wrist movement for this cohort, indicating that actigraphy cannot identify all arousal events;
- The number of arousal events varied greatly across the cohort, indicating that actigraphy cannot reliably predict arousal events in a generalised context;
- Arousals that did coincide with movement were generally longer in duration;
- Movements that coincided with wake or arousal were longer in duration than movements that occurred during sleep and did not coincide with an arousal;
- Actigraphy was able to predict the arousal events that coincided with movement;
- A representation of movement intensity predicted arousals better than associated movement duration; and
- The temporal association between apnoeic events and movement was inconsistent.

7

General Discussion and Conclusions

"There is a time for words, and there is a time for sleep."

— Odysseus, Homer, *Poet, c. 750 BC*

The objective of this thesis was to improve the utility of actigraphy for home-based assessment of paediatric sleep disorders by addressing the two key error types in the conventional actigraphy framework:

Error 1. False negatives: wake epochs with no measured movement are incorrectly identified as sleep.

Error 2. False positives: sleep epochs with measured movement are incorrectly identified as wake.

To address these error types, we investigated three hypotheses:

I Uni-axial accelerometry measured solely at the wrist limits sleep and wake prediction accuracy because movements orthogonal to the measurement axis, or occurring elsewhere on the body, cannot be detected, and consequently using tri-axial multisite accelerometry will improve the performance of actigraphy;

II Movement characteristics can differentiate sleep from wake because the physiological nature of these movements differ; and

III Physiological and pathological events characteristic of sleep disorders (e.g. apnoeas, hypopneas and transient arousals) cause sleep movements that contribute to false wake detections.

The key findings from investigating these hypotheses were:

I Summaries of movement that exploit tri-axial information significantly improve sleep and wake predictions relative to the conventional uni-axial representations, and incorporating movement information from multiple accelerometer locations with the conventional wrist placement further improves sleep and wake predictions.

Representing activity with three-dimensional vector techniques significantly improves sleep and wake prediction accuracy when compared to the same one-dimensional representation (Kappa agreement in Chapter 4: 0.402(0.141) vs. 0.268(0.210), $p < 0.05$ for tri-axial and uni-axial respectively). Tri-axial accelerometers can detect a greater range of movement and isolate physiological characteristics with greater accuracy because they measure two additional axes of motion. Unlike a single axis, measuring three axes of movement can summarise both vector magnitudes and phase changes. We saw in Chapter 4 that activity counts that summarise vector magnitudes have a greater resolution of activity than those that summarise phase changes. Vector magnitude appears a better descriptor of movements during sleep and wake than postural changes. This was verified in Chapter 5, where we saw (visually, by examining videos; and analytically, by examining spectral characteristics) that wake movements are generally more vigorous (i.e. greater vector magnitudes), and share similar postural spectral characteristics with sleep movements. Therefore, the most effective representation of movement for sleep and wake discrimination summarises the magnitude of movements.

Measuring movement of multiple limbs improves sleep and wake prediction accuracy by capturing additional movements that predominantly occur during wake. Incorporating ankle and toe movement into the conventional classification framework with wrist measurements in Chapter 4 gave the greatest performance improvement in terms of additional measurements (Kappa: 0.565(0.231) vs. 0.488(0.257), $p < 0.05$ for multisite and wrist accelerometry respectively). Additional accelerometers improve the first error type by increasing the number of detected movements during wake; however, they also increase system complexity and the potential for patient discomfort. The benefits to classification performance of additional accelerometers will need to be considered against these limitations for each clinical application. Conforming to the conventional framework, the

wrist was verified as the most effective single accelerometer placement for sleep and wake prediction.

The greatest impact on predictive performance throughout the analyses in this thesis was consistently seen when applying a smoothing filter (direct comparison shown in Appendix B.2). The conventional filter design smooths regions of large activity values into surrounding epochs of no activity, and, depending on the filter coefficients, attenuates single epochs of activity. Considering that wake corresponds with regions of high activity, the smoothing filter temporally extends this activity and consequently reduces false sleep detections. Similarly, the smoothing filter reduces false wake detections by attenuating short regions of high activity epochs during sleep. However, the smoothing filter is not always effective because there are extended regions of high activity during sleep (such as positional changes) and short regions of low activity during wake (such as limb movements). Applying the smoothing filter to these regions will worsen the effect on sleep and wake predictive performance. An approach that considers the likelihood of high-activity epochs occurring during sleep or wake may improve this; however, the variability in movement prevalence and physiological characteristics between patients suggests that this may not be feasible in a generalised context.

II Movement characteristics differ between sleep and wake, although they are not able to accurately predict sleep and wake for this cohort.

Applying a heuristic to remove movements specific to sleep as a pre-processing procedure on the raw multisite accelerometry data effectively improved the second error type, resulting in improved sleep and wake classification accuracy (Kappa: 0.630(0.292) vs. 0.565(0.231) for post- and pre-heuristic respectively). Where additional accelerometers reduce false sleep detections by increasing the number of detected wake movements, the heuristic reduces false wake detections by reducing the number of sleep movements. Since applying the heuristic requires minimal computation and no additional hardware, pre-processing the raw data is an effective improvement that can easily be incorporated into the existing actigraphy framework. However, the heuristic applied in this thesis was formulated by manually analysing movements from a subset of the patient cohort. As a consequence of the relatively small patient sample size, these sleep movements may not represent a generalised cohort. Therefore, to incorporate this into conventional sleep actigraphy assessment, a specific heuristic will need to be identified with a different study design.

Short-duration movements were consistently identified as occurring more often during sleep in Chapter 5 and again in Chapter 6. Summarising movement within a large epoch results in many short-duration movements appearing as a single long movement, despite being

distinctly different movements. Misrepresenting movements within these low-temporal resolution activity counts can cause false wake detections because the resulting activity value of cumulative short-duration movements is similar to wake epochs. From this, we hypothesised that removing short-duration movements from analysis would reduce false wake detections (thereby addressing the second error type) by minimising high-activity sleep epochs. However, pre-processing the raw data to specifically remove these short-duration movements (see Appendix B.1) did not significantly improve the predictive performance, and including movement duration in the regression model in Chapter 5 showed no significant effect. Although short-duration movements are generally specific to sleep, this information alone cannot discriminate sleep from wake for this cohort. Incorporating these movements with others during sleep (as done with the heuristic in Chapter 5) is effective at significantly improving sleep and wake discrimination; however, this requires multiple accelerometers, which introduces the previously discussed practical limitations.

Localised spectral characteristics of sleep and wake movements significantly differed for movements $2-5$s in duration for a subset of patients. We saw in Chapter 5 that wake movements were generally more vigorous, had low-frequency spectral content characteristic of positional changes and were more consistent across the cohort. The time-invariant epochs (commonly 30s in existing commercial systems) cancel any sleep/wake discrimination potential of these spectral characteristics and movement duration because they are summarised within this extended period of time. However, the sleep and wake predictive performance was not improved when incorporating these characteristics into a movement-by-movement classification model, relative to the conventional epoch-by-epoch summaries (AUC: 63.9(6.7)% vs. 69.7(7.9)% for movement-by-movement classification and the conventional epoch summary respectively). The similar predictive performance of these distinctly different classification models verifies that movement characteristics cannot *accurately* differentiate sleep and wake movements for a generalised cohort. Therefore, some sleep movements will always cause false wake detections in actigraphy-derived sleep and wake estimates; actigraphy is only suitable for indicating an approximation of sleep quality.

III Lengthy transient arousals correlate with the longer sleep movements that are most likely to cause false wake detections.

The prevalence and characteristics of sleep movements vary considerably across the cohort, which impacts the ability to detect sleep movements and accurately describe their physiology. As a consequence of this variability, movement information alone cannot accurately predict sleep and wake. However, we saw in Chapter 6 that the longer confounding sleep movements correlate with lengthy transient arousals that are characteristic of sleep disturbance caused by sleep disorders. Therefore, sleep movements that confound wake es-

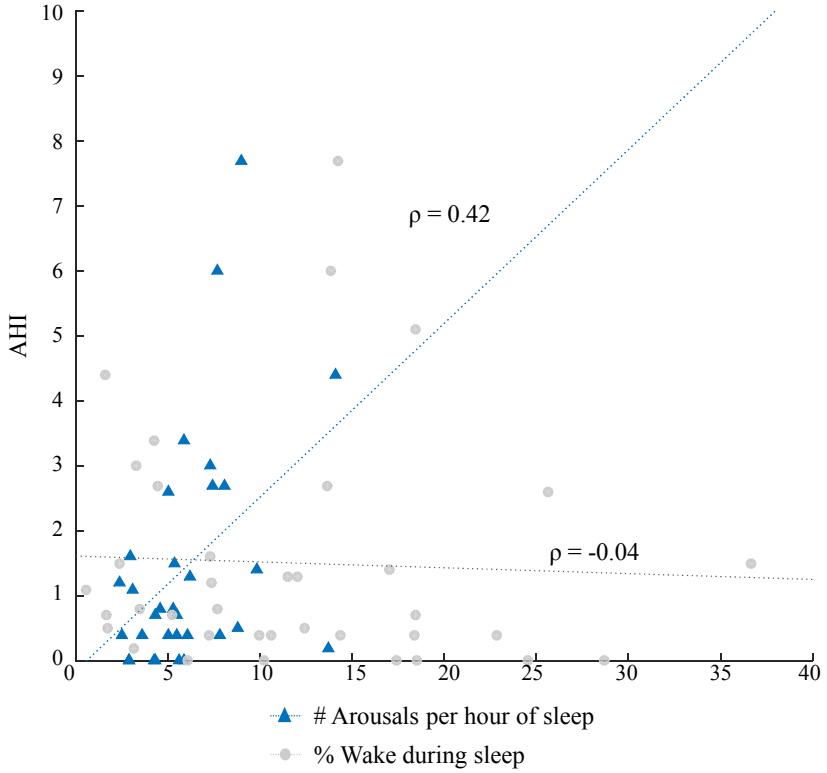


Figure 7.1: Correlation between AHI and the number of arousals per hour of sleep and percentage of time spent awake after sleep onset, as defined by polysomnography. Patient 21 *was removed from the diagram is it was a significant outlier (AHI =* 16.9*).*

timates are actually indicative of regions of sleep disturbance. Incorporating these lengthier sleep movements with the conventional wake scores may summarise sleep disturbance and thereby provide a metric that is correlated with sleep disorder severity. Indeed, we can see from Fig. 7.1 that the number of arousals per hour of sleep is more correlated with AHI (indicative of OSA) than the percentage of wake after sleep onset for this cohort. We can hypothesise that combining these measures would provide a metric that is better correlated with AHI. Actigraphy is likely able to achieve this because it can predict both wake and lengthy transient arousals (duration between approximately 8s and 16s, seen in Chapter 6). However, this exploration requires a different clinical study design with a larger set of patients (control and moderate-to-severe sleep disorder symptomatology).

7.1 Conclusions

We can conclude from the analyses in this thesis that:

- Measuring movement with tri-axial accelerometry detects a larger range of movements that, when summarised with tri-axial techniques, improve sleep and wake predictions relative to conventional uni-axial accelerometry;

- Epoch summaries of movement that capture magnitude information are more effective for sleep and wake classification than summaries that capture postural information;
- Incorporating ankle and toe movement into the scoring routine improves sleep and wake classification by increasing activity during wake epochs, which reduces false sleep detections;
- Identifying and heuristically removing movements that are generally sleep-specific as a pre-processing step on the raw accelerometry data improves the sleep and wake discrimination ability of activity counts;
- Localised spectral characteristics differ between sleep and wake movements, although not in a generalised context;
- The technique (temporal or spectral, movement-based or epoch-based) for summarising movement generally does not affect the sleep and wake predictive performance;
- Movement information alone cannot accurately estimate sleep quality; and
- Lengthy and high-intensity sleep movements are associated with the lengthier transient arousals that are characteristic of sleep disturbance.

The utility of actigraphy for sleep assessment may be improved by combining detection of the transient arousals characteristic of sleep disturbance with the standard wake scores to capture signs indicative of sleep disorder severity.

7.2 Contributions

There were three main contributions to paediatric sleep assessment with actigraphy:

1. Previous and existing commercial actigraphy systems in sleep assessment cannot concurrently analyse uni-axial and tri-axial data from the same device, simultaneously analyse movement from different locations on the body, and temporally synchronise with the gold standard polysomnography. In this thesis, we have directly compared the sleep and wake predictive performance of the conventional activity count summary techniques, uni-axial and tri-axial accelerometry using the same device, and movement information from five locations (i.e. left-finger, left-wrist, upper thorax, left-ankle and left-toe movements). These analyses were performed using high-resolution raw accelerometry data that was synchronised to within 0.1s of polysomnography;

2. There have been no known attempts in literature to analyse specific localised spectral characteristics of movements that occur during sleep and wake, or that predict sleep and wake on a time-varying basis. We have segmented high-resolution sleep and wake movements and compared the localised spectral information. We have also analysed sleep and wake predictive performance on a movement-by-movement basis; and

3. There have been no known analysis of the high-resolution temporal association between apnoeic events, transient arousals and movements during sleep for children. We have segmented raw movements during sleep and temporally compared these to manually scored apnoeic events and transient arousals. We have also analysed the ability of sleep movements to predict transient arousals.

7.3 Implications for clinical practice

Representing movement with summaries that incorporate information from tri-axial accelerometry will yield benefits to sleep and wake classification with minimal additional cost, and no increase in system complexity or invasiveness. This is likely to be the most attractive clinical improvement for a device targeted at home-based sleep assessment. Incorporating multisite accelerometry will further improve sleep estimates, particularly when applying a heuristic that preliminarily removes movement types specific to sleep. However, additional accelerometers increase system complexity and patient discomfort. This may seem unimportant for diagnostic studies, where the numerous sensors for full polysomnography are already required. However, in practice, we regularly observed that children often play with the accelerometer units, not only interfering with the data, but at times inadvertently removing the units altogether. Minimising the number of required sensors mitigates this problem. In addition to patient interference, a more complex system requires expert support for system configuration, which is not ideal for home-based assessment. Therefore, the benefits to sleep/wake classification performance of data from additional accelerometers will need to be weighed against these practical limitations.

Capturing sleep disturbance severity by incorporating the detection of lengthy and high-intensity sleep movements into the conventional classification framework requires no additional hardware modifications, and negligible additional on-board processing. Provided that the actigraphy modules store raw tri-axial accelerometry data, movement segmentation and sleep disturbance detection could be implemented within the standard computer-based actigraphy scoring software. Although this analysis could not be performed in this thesis (due to inadequate patient sample size and disorder prevalence), incorporating lengthy and high-intensity sleep movement detection with the standard wake scores may provide a useful metric for non-invasively assessing signs indicative of sleep disorder severity in a non-laboratory setting.

7.4 Limitations

There were a number of limitations of the analyses performed in this thesis:

1. The cohort comprised of children symptomatic of sleep-related breathing disorders, was male dominated, had a large range in age, and was heavily biased towards low-AHI (with only 2 in the 30 usable studies in the moderate to severe AHI categories). The results are consequently unlikely to be representative of those observed in severe OSA patients, and care must be taken when generalising the results to other cohorts;

2. The sleep stages and respiratory events used to validate the accuracy of all analyses were manually scored by one technician. Although there are well documented rules for scoring these events, the inter-scorer variability is often high in literature. Because of this, there may be some inaccurately scored events used as a 'ground truth' in these analyses;

3. Data were collected in a sleep laboratory, which may present different data than that collected in a more typical sleeping environment because of the invasive nature of polysomnography and the general laboratory environment;

4. For the multisite analysis, accelerometers were only placed on the left-side limbs and the chest. Although movement from the right-side of the body and head may be of interest, practical limitations restrict placing accelerometers on all areas of interest on the body; and

5. The number of transient arousals, apnoeas and hypopneas varied greatly across the cohort, with some patients experiencing none, and others dominating the distribution. This variability limits the generalisability and statistical power of the association analysis between sleep movements and these physiological and pathological events.

7.5 Recommendations for future research

- In Chapter 5, we found localised spectral characteristics that differed between sleep and wake movements, but were not significant across the cohort when incorporated into a classification model. Since airway morphology changes with age, gender and hormones, movement characteristics may significantly differ in a generalised cohort between age, gender and pre- and post-puberty. It would be necessary to expand the study with a larger number of children with severe obstructive sleep apnoea and a larger range of age and gender to better analyse these relationships.

- Chapter 6 found that longer and high-intensity sleep movements are temporally associated with lengthy transient arousals. Given that transient arousals are characteristic of sleep disturbance associated with sleep disorders, actigraphy could detect the arousals that are detrimental to sleep quality. The standard 30s wake scores cannot summarise these sleep disturbances because of the poor temporal resolution. Incorporating these lengthier arousals into the actigraphy routine could capture additional aspects of sleep disturbance associated with sleep disorders and may provide a novel index of sleep apnoea severity. This analysis requires a larger sample of children with moderate to severe sleep disorder, and a control group of otherwise healthy children.

- The analyses throughout this thesis were performed on a paediatric cohort. Although there are known pathological differences between children and adults, it would be interesting to determine if the results can be extended to adults.

A

Heuristic: Removal of restless sleep

The analysis in Chapter 5 applies a heuristic to remove movements that were identified as predominantly occurring during sleep. The pseudo-code algorithm for the heuristic is detailed below.

Data: Raw (x, y, z)100Hz accelerometry data
Result: Raw (x, y, z)100Hz accelerometry data
repeat
  if *movement is detected within* $1s$ *and there is no more movement for another* $7s$ **then**
    segment movement s;
    if *(any wrist*$(x, y, z) > T$ *and any chest*$(x, y, z) > T$ *and any ankle*$(x, y, z) > T$ *and no chest DC offset) ;* `/* Body shift */`
    or (any wrist$(x, y, z) > T$ *and wrist DC offset) ;* `/* Hand positional shift */`
    or (duration of $s < 2s$*) ;* `/* 'Bursty' movement */`
    then
      $s(x, y, z) \leftarrow [0, 0, 0]$; `/* Set movement segment to 0 */`
    else
      Ignore;
    end
  end
until *all data points are checked*;

B

Pilot Analysis: Exploring wake quantification with actigraphy

The analyses in Chapter 5 and Chapter 6 found that short-duration movements occur most often during sleep. This short analysis will explore the impact of removing these movements as a preliminary processing step on the raw accelerometry data, prior to generating the conventional activity counts. We also saw throughout the thesis that post-processing activity counts using a smoothing filter appeared to have the greatest impact on sleep and wake detection accuracy. This analysis will also briefly compare the effect of this process on sleep and wake detection accuracy using a smoothing filter with constant coefficients, and a regression model that incorporates different epoch information.

B.1 Removal of short-duration movements

From Section 6.1 we found that short-duration movements mostly occur during sleep and are generally not associated with transient arousal or wake. From this, we can conclude that short-duration movements are likely not indicative of sleep disturbance. In this exploratory analysis we will remove these short-duration movements prior to deriving activity counts and assess if this improves the accuracy of sleep scores.

The duration threshold for each location was first identified. The threshold for each patient was identified as the threshold with the maximum Kappa in ROC analysis for the duration of all identified movements. The threshold for the location was found by taking the median of these thresholds across all patients. The final movement duration thresholds for differentiating sleep from wake for each accelerometer placement were identified as:

Placement	Threshold
Finger accelerometer	5.82s
Wrist accelerometer	2.88s
Chest accelerometer	2.73s
Ankle accelerometer	3.00s
Toe accelerometer	2.75s

The raw data was then pre-processed to remove all detected movements that were less than the identified threshold. Raw data were then segmented into 30s regions to correspond with the manual sleep/wake scoring from polysomnography. Activity counts were then generated for the processed data using the tri-axial zero-crossing method (detailed in Section 4.1.1). The discrimination ability for sleep and wake scoring of the activity counts was compared to activity counts derived from raw tri-axial accelerometry prior to removing the short-duration movements. The differentiation ability was analysed for each individual accelerometer placement and also when combining movement information from all placements. This was combined by summing the individual activity counts for all placements. A weighted moving average filter was applied to the activity counts prior to ROC analysis.

Removing the short-duration sleep movements increased the number of sleep epochs with a zero activity count, and decreased the number of sleep epochs with high values of activity counts, as shown in Fig B.1. However, removing these short-duration movements did not significantly improve sleep and wake predictive performance. As expected, the maximum Kappa operating point tended to move towards reducing the sensitivity (89.2(10.2)% vs. 82.7(25.6)% for pre- and post-removal of short-duration wrist movements respectively) for improved specificity (62.8(18.7)% vs. 65.4(19.9)% for pre- and post-removal of short-duration wrist move-

Table B.1: Sleep/wake predictive performance for the wrist accelerometer placement and all placements combined prior- and post-removal of short-duration movements.

	Se (at 70% Sp) (%)	**Maximum κ Operating Point**		
		Se (%)	**Sp (%)**	κ
Wrist accelerometer placement				
Pre-removal	77.5 (17.4)	89.2 (10.2)	62.8 (18.7)	0.521 (0.159)
Post-removal	76.0 (18.0)	82.7 (25.6)	65.4 (19.9)	0.537 (0.166)
All accelerometer placements				
Pre-removal	83.1 (16.9)	89.6 (7.9)	65.2 (19.3)	0.523 (0.177)
Post-removal	83.7 (15.7)	90.6 (7.3)	66.7 (17.7)	0.570 (0.153)

Values are shown as mean (SD).

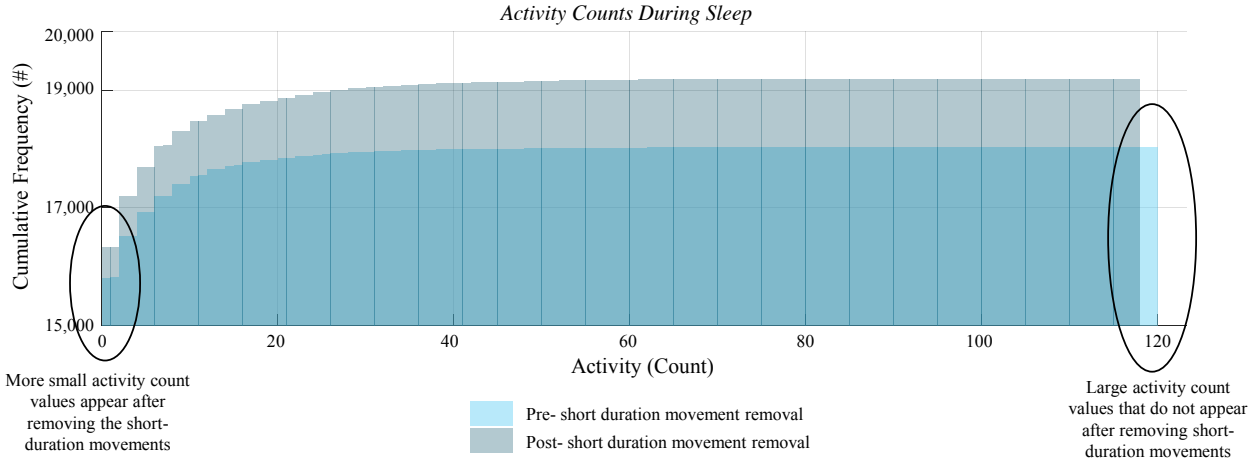


Figure B.1: Histogram of activity counts during sleep for all accelerometer placements pre- and post-removal of the short-duration movements.

ments respectively). This also resulted in slightly higher Kappa agreement (0.527(0.159) vs. 0.537(0.166) for pre- and post-removal of short-duration wrist movements respectively). We can also see from the average ROC curves in Fig. B.2 that removing the short-duration movements improved the sleep and wake discrimination ability for some operating points for wrist movement (shown in Fig. B.2a) and for most operating points for all movements combined (shown in Fig. B.2b). Considering that these were not significant effects, short-duration movements do not have a great effect on sleep and wake predictive accuracy for this cohort.

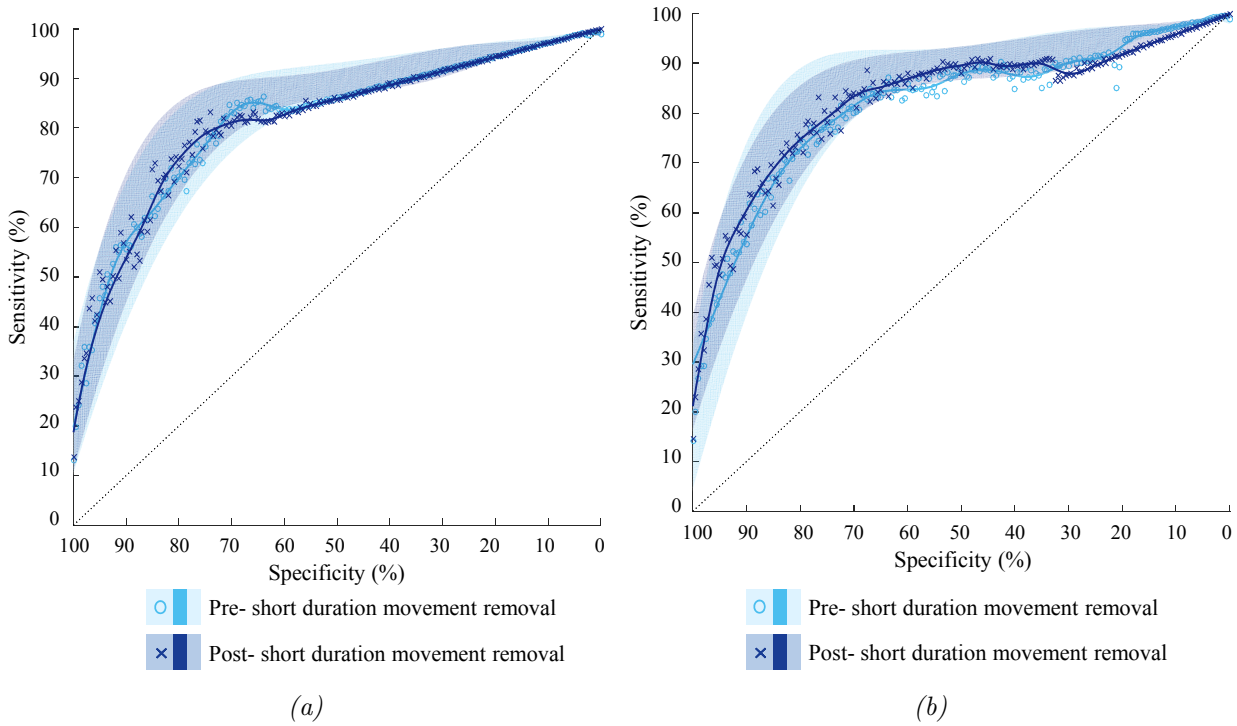


Figure B.2: Median ROC curves for the full population used in this analysis for removing short-duration movements for (a) wrist movement, and (b) any movement.

B.2 Differentiating inactive wake from sleep

Sadeh et al. [187] derived a wake re-scoring algorithm using discriminant analysis with different summary variables (summarised in Table B.3). Although this technique smooths the data by combining information from surrounding epochs, the main advantage of this technique is that it uses different summaries of movement for each epoch of activity. A single summary value likely has inadequate resolution for accurately representing sleep and wake. Indeed, we saw this in 4 where there was no significant differences between the different time-series representations, and again in 5 with the spectral characteristics. From this, we can hypothesise that multiple summary values for each epoch is significantly more accurate than a single summary value. However, Sadeh et al. [187] found only a minor difference in sleep and wake detection when applying their algorithm.

Domingues et al. [18] successfully identified statistical models with high discrimination ability of sleep and wake movements. Previous samples and coefficients of previous sample models were combined to discriminate sleep from wake. Although Domingues et al. [18] were able to differentiate sleep movements from wake movements, they performed no validation studies and the analysis was only performed on a small set of patients. Furthermore, this method does not address the limitation of no movement while awake. Autoregressive models use a feedback component in the model; the current prediction is dependent on previous predictions. These systems are more complex than standard smoothing models and consequently require more data than the standard smoothing models. Further to this, the smoothing functions use information that summarise future epochs. For these reasons, we will not be analysing autoregressive models.

Linear regression was used to identify a model by fitting variables to 'wake'. Half of the 30 patients were used to train the model, and the other half were used to test the effect of the model

Table B.2: Summary of patient characteristics for the 15 *patients used for training the regression model and* 15 *used for testing the model in this analysis.*

	Training	Testing
N	16	16
Age (years)	9 (6 - 16)	9 (6 - 16)
Gender (M/F)	12/4	10/6
AHI	1.15 (0 - 16.9)	0.55 (0 - 6)
Sleep Efficiency (%)	82.5 (10.0)	84.0 (8.5)
Total Sleep Time (mins)	428.2 (68.6)	443.8 (52.9)
REM (%)	21.5 (7.2)	22.5 (5.5)

Excluding gender, values are shown as mean (sd) where normally distributed, and median (range) otherwise.

Table B.3: Variables used in the regression analysis.

Variable	Description
	Sadeh et al. [187] variables
LOG-Act	The natural logarithm of the current epoch $E[0]$ plus 1. $$ln(E[0]+1)$$
Mean-W 5 min	The average activity 5 minutes before and after the current epoch of activity inclusive. $$\frac{1}{21}\sum_{n=-10}^{+10} E[n]$$
NAT	The number of times activity counts have a value between 50 and 100 within the epochs 5 minutes before and after the current epoch, inclusive. $$\sum_{n=-10}^{+10} \mathbb{1}\{50 < E[n] < 100\}$$
SD-last 6 min	The standard deviation σ of the activity within the 5 minutes prior to the current epoch, inclusive. $$\sigma_{E[-10,\cdots,0]}$$
	Other variables
Act_n	Raw zero-crossing activity count in epoch n, for all 5 epochs prior and post the current epoch. $$E[n], \forall -5 \leq n \leq 5$$
NumMovements	The number of detected movements within the current epoch, using the detection method discussed in Section 3.2.3.
MeanDuration	The mean duration of detected movements within the current epoch, using the detection method discussed in Section 3.2.3.
MeanDuration_n	The mean duration of detected movements (using the detection method discussed in Section 3.2.3) within epoch n, for all 5 epochs prior and post the current epoch.

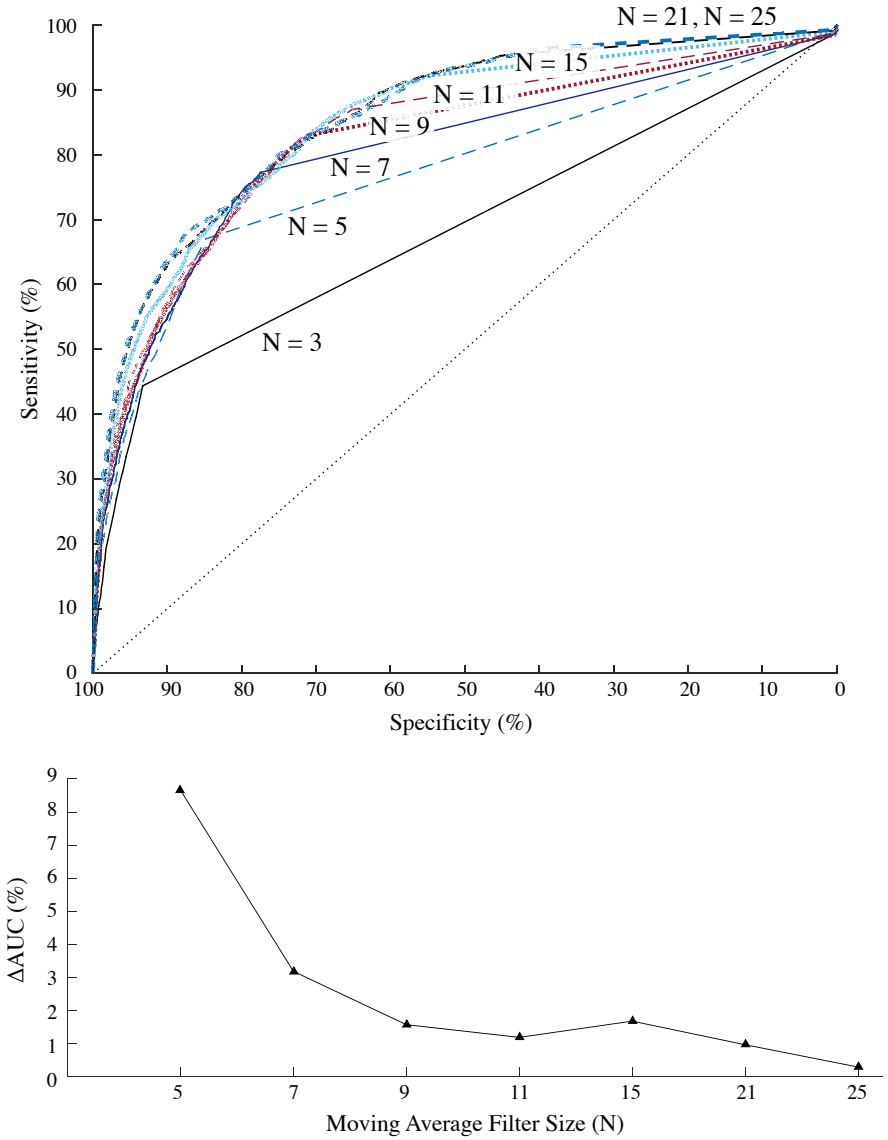


Figure B.3: Comparison of different lengths of moving average filter.

on wake detection (patient characteristics summarised in Table B.2). Variables included in the regression were based on Sadeh et al. [187] work. These variables represent longer-duration movements and are consequently more suited for wake detection. We have seen throughout the analyses that short-duration movements are more common to sleep. Unfortunately, the conventional activity counts do not summarise movement duration. For the purpose of this analysis, two activity count metrics were derived that represent the:

- number of detected movements within a 30s epoch; and
- mean detected movement duration within a 30s epoch.

In addition to the Sadeh et al. [187] variables, zero-crossing activity counts surrounding the

current epoch (±5 epochs) and average movement duration surrounding the current epoch (±5 epochs) were included. The additional variables are summarised in Table B.3.

To analyse the effect of the different variables on wake detection, the data were standardised prior to generating the regression model. The standardised coefficients β and $p-$values were analysed to identify the variables with the greatest influence on the wake model. The significance of the individual coefficients is summarised by the t statistic,

$$t = \frac{c}{SE}, \tag{B.1}$$

where c is the coefficient and SE is the standard error of that coefficient.

The unstandardised data was used to generate the prediction model using stepwise regression. Unstandardised data was used in this case so that the final equation could directly be applied to the raw zero-crossing activity counts. Both Akaike information criterion (AIC) and adjusted R^2 was used as the criterion for the stepwise model. This was done to compare the fit for maximum information (AIC) and explained variability of the predicted outcome in the training procedure (adjusted R^2). The model was then used to predict sleep and wake for the test set of patients. The agreement (standard and Kappa), AUC and sensitivity and specificity were analysed to assess the performance of the model against raw activity counts, and a standard 9−epoch (4.5 minute) moving average filter. A filter of 9 epochs was chosen because the increase in AUC was comparatively small for higher order filters (as shown in Fig. B.3(bottom), which shows the change in AUC between consecutive increments in filter order). The thresholds for prediction were identified as the point in the training ROC curve that gave the maximum Kappa agreement. The predictive performance was compared using the Wilcoxon rank sum test where the performance metrics were not normally distributed.

The significant standardised β coefficients are shown in Table B.4. LOG-Act had the greatest effect on the model when compared to the other coefficients ($\beta = 0.184$ vs. $\beta < 0.1$). The logarithm of activity counts seems a better representation of activity than raw activity counts. Therefore, the logarithm of previous and future epochs (±2.5 minutes) was added to the stepwise model. The significant β coefficients for the stepwise analysis are compared in Table B.5. Excluding Act_{-5}, both methods chose the same variables as the significant effects on the model. These variables formed the final regression model,

$$\begin{aligned}\hat{\text{Wake}} = {} & 0.111\log(A_0+1) + 0.088\log(A_{-5}+1) + 0.049\log(A_{-3}+1) \\ & + 0.034\log(A_{-1}+1) + 0.011\text{Mean-W 5 min} - 0.006\end{aligned} \tag{B.2}$$

The model explains 97.3% of the variation ($R^2 = 0.973$) and improves prediction when compared to a constant model (F statistic of $88,500$, $p < 0.001$). For model validation, the sleep and wake classification thresholds were found using ROC analysis:

Method	Threshold
No smoothing	1
Moving average filter	3.759
Regression model	0.231

The predictive performance of the models is summarised in Table B.6. Compared to no smoothing and the standard moving average filter, the regression model provided significantly better sleep and wake discrimination ability and sensitivity at 75% specificity.

Table B.4: Significant β coefficients for the linear regression model using Sadeh et al. [187] variables and the surrounding 2.5 minutes of activity.

Variable	β	t Statistic
Intercept	0.124	-8.03
LOG-Act	0.184	257.36
Mean-W 5 min	0.089	76.08
Act_0	-0.043	-80.76
Act_{-5}	0.033	52.79
Act_{-4}	0.027	53.0
Act_{-3}	0.025	44.27
Act_{-1}	0.021	45.66
NAT	0.019	26.04
Act_{-2}	0.016	30.93
Act_1	0.012	25.62
SD-last 6 min	-0.025	-22.80

Table B.5: Significant β coefficients for the stepwise regression model including logarithmic activity.

	AIC		R^2	
Variable	β	t Statistic	β	t Statistic
Intercept	0.198	62.82	0.210	72.25
Mean-W 5 min	0.110	18.84	0.124	26.34
LOG-Act	0.066	16.58	0.063	17.39
LOG-Act_{-5}	0.052	13.05	0.107	17.07
LOG-Act_{-3}	0.045	11.59	0.043	12.29
LOG-Act_{-1}	0.043	11.38	0.043	12.49
Act_{-5}	-	-	-0.060	-9.95

Table B.6: The predictive performance after applying the regression model to zero-crossing activity counts.

	No smoothing	Moving Average Filter	Regression Model
Agreement (%)	88.3 (74.5 - 94.1)	86.5 (77.3 - 94.3)	88.1 (77.3 - 93.6)
Kappa	0.388 (-0.034 - 0.691)	0.501 (-0.040 - 0.721)	0.494 (-0.068 - 0.765)
Sensitivity (%)	92.6 (17.5 - 98.5)	93.4 (84.3 - 99.6)	88.0 (17.3 - 95.6)
Specificity (%)	41.5 (8.2 - 87.0)	55.5 (6.8 - 86.9)	58.8 (8.2 - 94.5)
AUC (%)	68.8 (48.7 - 85.5)	83.3 (46.5 - 95.2)	89.8 (41.8 - 97.4)*
Sensitivity at 75% Specificity (%)	41.5 (3.6 - 74.9)	79.1 (17.8 - 95.9)	89.2 (11.5 - 98.8)*

Values are shown as median (range) where a non-normal distribution is observed.
* Regression model significantly greater performance than no smoothing, $p < 0.05$.

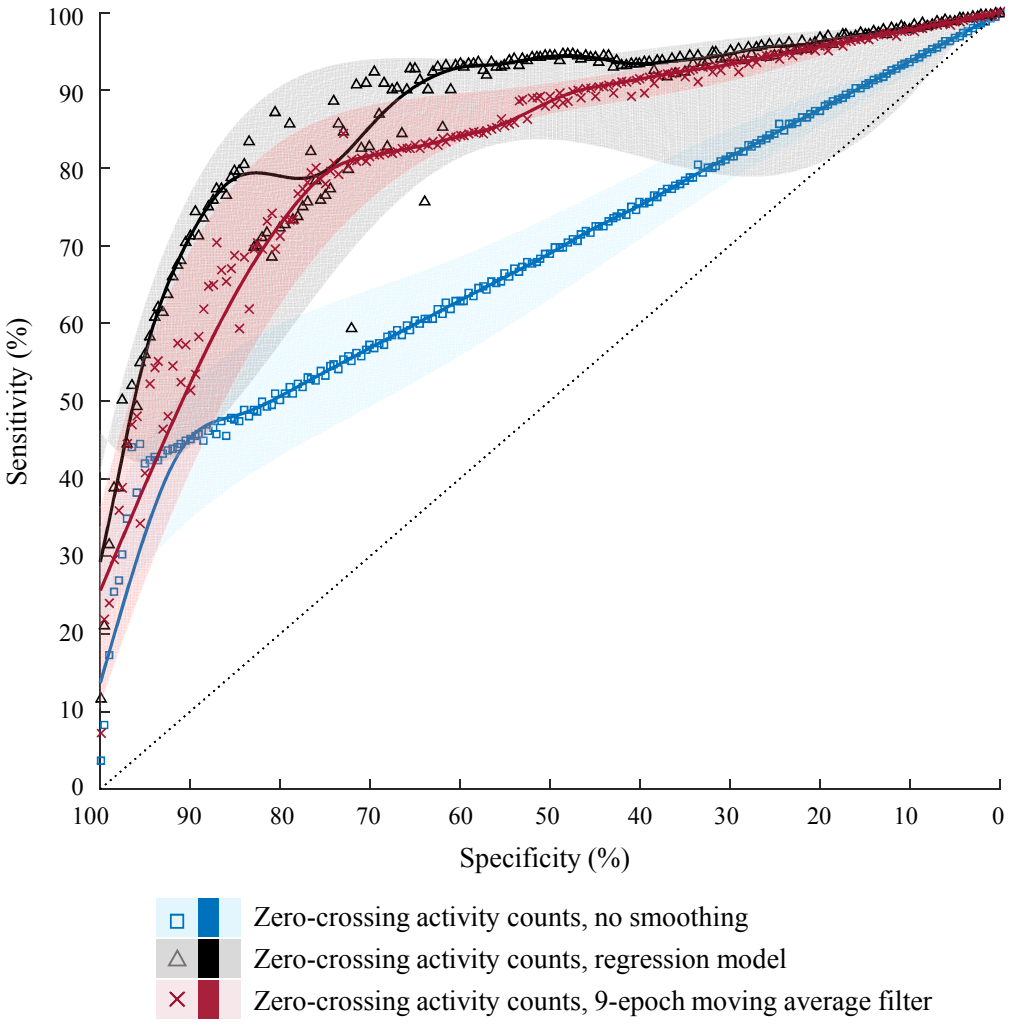


Figure B.4: ROC of the test patients with no smoothing, the regression model and a standard 9-epoch moving average filter. Shown is the median curve and the bounds represent the inter-quartile range of values.

The Sadeh et al. [187] variables all had a significant effect on the regression model. The largest effect was the logarithm of the raw activity count. This is likely due to the increments and/or distribution of activity because the activity count distribution is skewed towards low values. A logarithmic transform consequently improves the discrimination ability between these activity values. The adjusted R^2 criterion chose Act_{-5} as an additional variable. However, this variable had a negative effect on the model and therefore may not contribute additional information. This could explain why the AIC criterion did not chose Act_{-5} as a significant variable. The regression model had greater performance than the standard moving average smoothing filter (comparative ROC curves illustrated in Fig. B.4). The additional information summarised by the regression model improves sleep and wake discrimination.

C

Software Module: Labelling movements

A MATLAB graphical user interface (GUI) software module was developed to aid with manually labelling movements during sleep and wake, for each patient. The software module requires a patient study .m file containing the raw CMAS data from all accelerometer modules (finger, wrist, thorax, ankle and toe), the manually scored hypnogram from polysomnography and the video file from the polysomnography montage. The software also assumes that the raw accelerometry data is synchronised with the polysomonogram scoring (see Section 3.2.1). The importing and synchronisation procedures are performed in the *import_data.m* and *synchronise_accel.m* files.

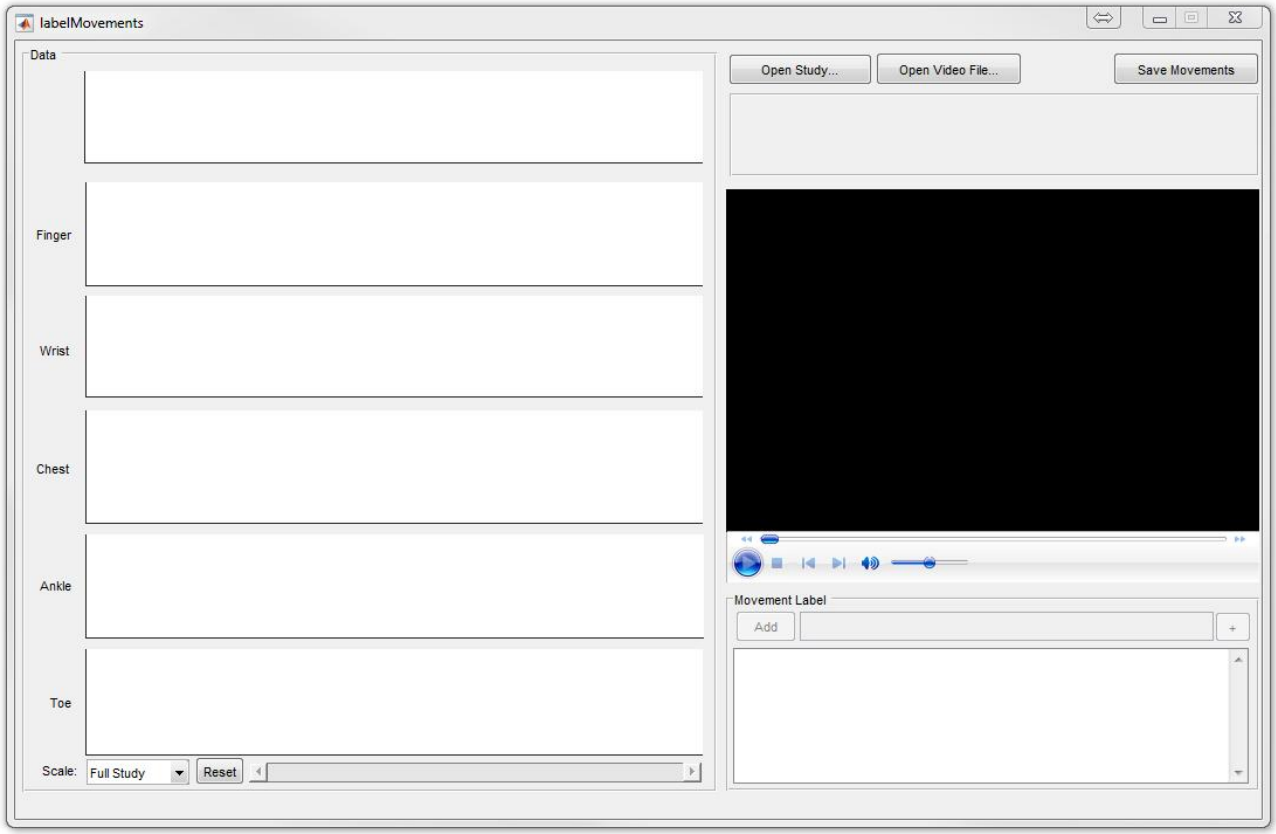
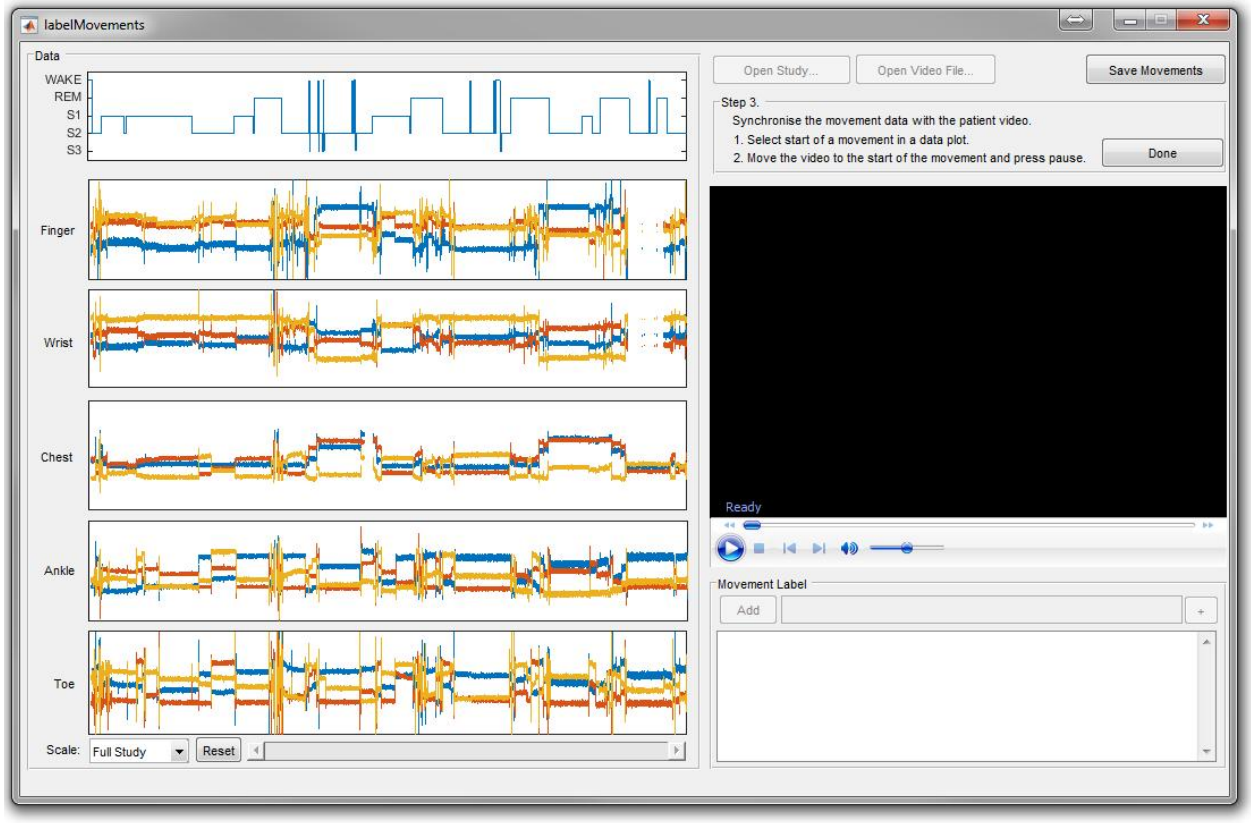


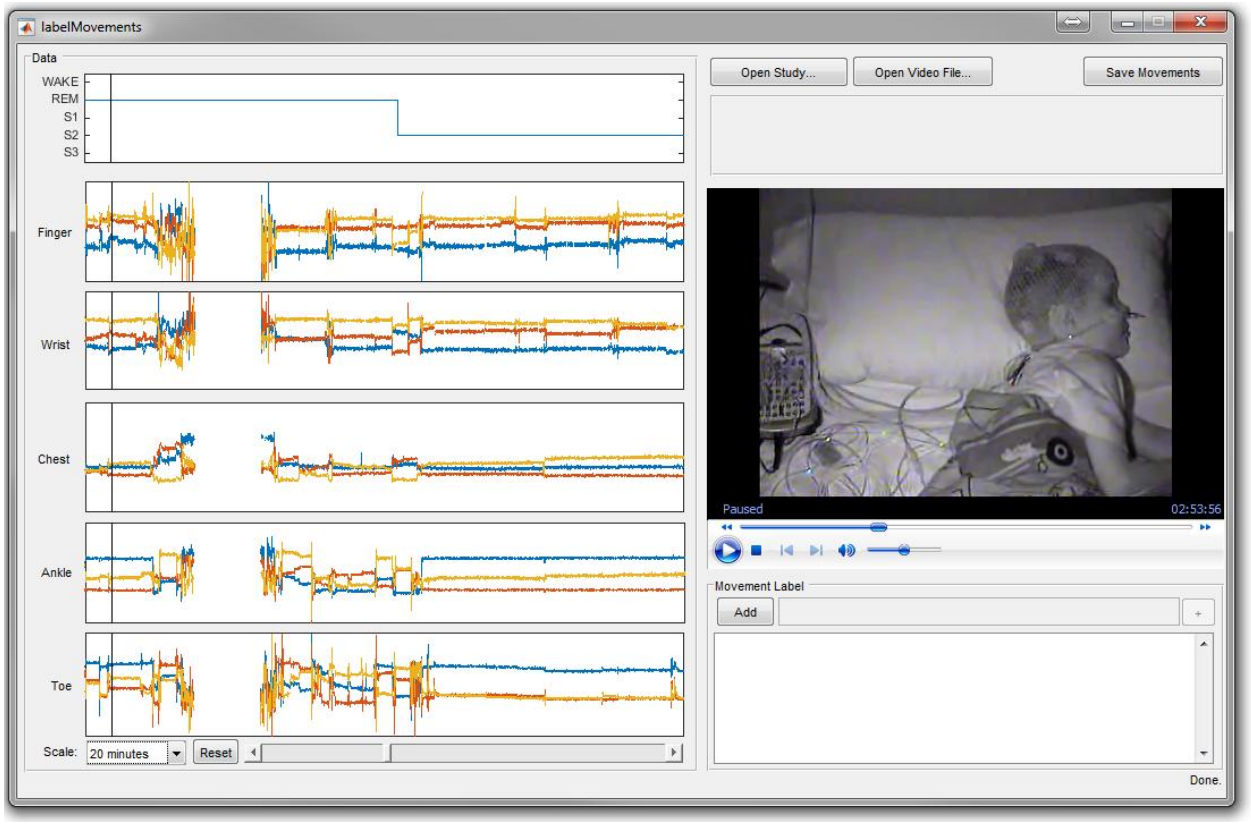
Figure C.1: Custom software for manually labelling patient movements.

The user is first asked to load the patient's .m and video files (shown in Fig. C.2). Once these are loaded into the GUI, the user is required to ensure that the video is synchronised with the accelerometry data by selecting the start of a movement in the accelerometry plots and identifying the exact temporal location of the movement in the video. After the 'Done' button is selected, the accelerometry data is synchronised with the polysomnogram video. As the video plays, a cursor will move at the same position on the accelerometry plots.

Once the data is synchronised with the video, the user can label movements by positioning the movement of interest in the centre of the plot view and selecting the 'Add' option on the right. This will pause the video and the movement region will be highlighted. The bounds of this highlighted region can be altered by clicking near the edge of the region (towards or further away from the movement). An example of segmenting movements is shown in Fig. C.3. Once the user is satisfied with the highlighted region, they can type in a description for the movement and select '+'. The movement will be added to the list of movements for that patient. When the user has labelled all movements, selecting the 'Save Movements' button will save the list of movements within the patient's .m study file as a cell matrix with columns: description, sleep stage, temporal indices, and (x, y, z) raw data.



(a)



(b)

Figure C.2: Custom software for manually labelling patient movements with (a) accelerometry data loaded and (b) patient video loaded.

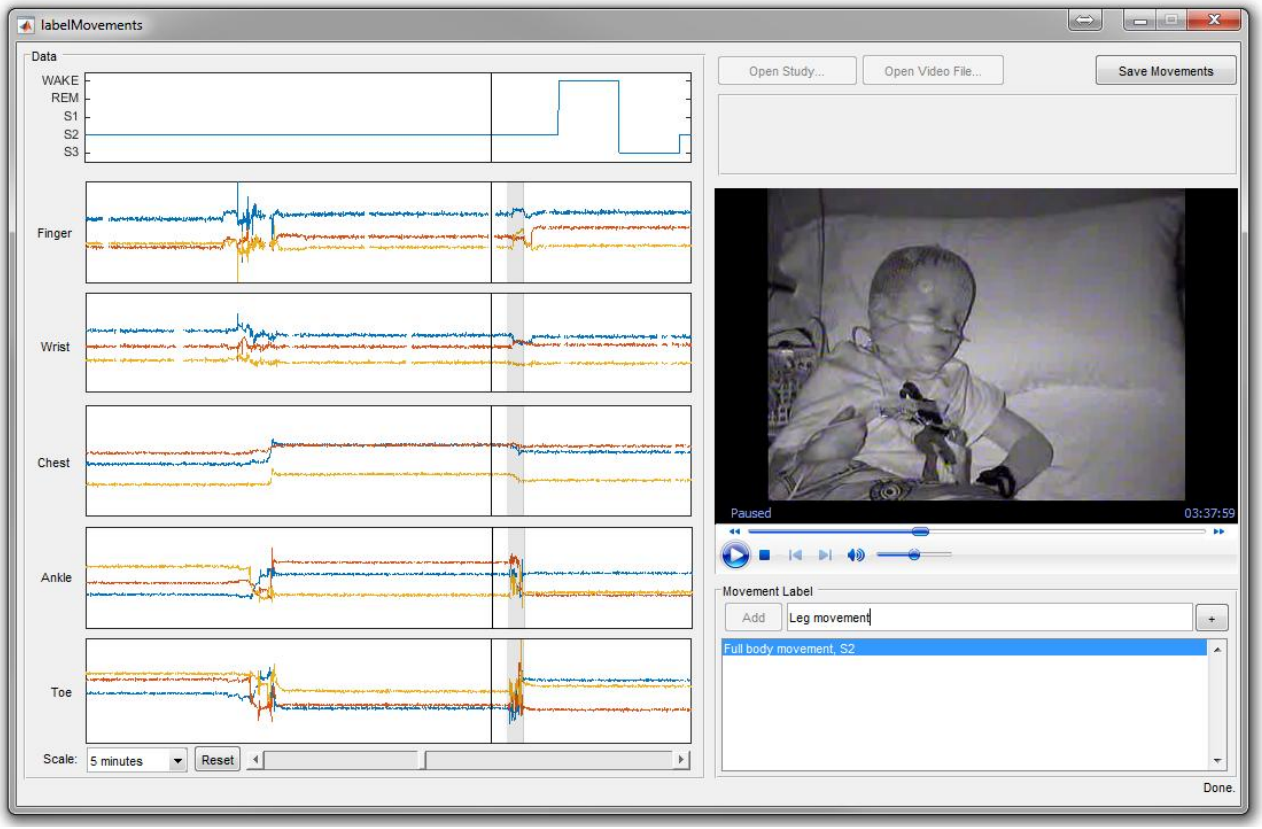


Figure C.3: Segmenting movements with the custom software.

D

Software Module: Validating the automated movement detection algorithm

A custom MATLAB GUI was developed to assess the accuracy of the automated movement detection algorithm in Chapter 5. This software allows the user to manually count the number of movements that were detected by the algorithm, but did not contain movement (false positives), and the number of actual movements that were not detected by the algorithm (false negatives). The GUI displays 120s of raw accelerometry data overlaid with a binary plot indicating movement detection (*high* indicates detected movement, *low* indicates no detected movement). The user can move the data forwards or backwards temporally by 120s increments. Within each displayed window, the user counts the number of correctly and falsely detected movements and updates this in the counter (i.e. the up and down arrows for FPR and FNR) in the GUI. The user can load different patient data by specifying the patient number and pressing the 'Load' button at the bottom of the GUI. The GUI stores the total number of false positives and false negatives for all tested patients, and outputs the rates accordingly once the module is closed.

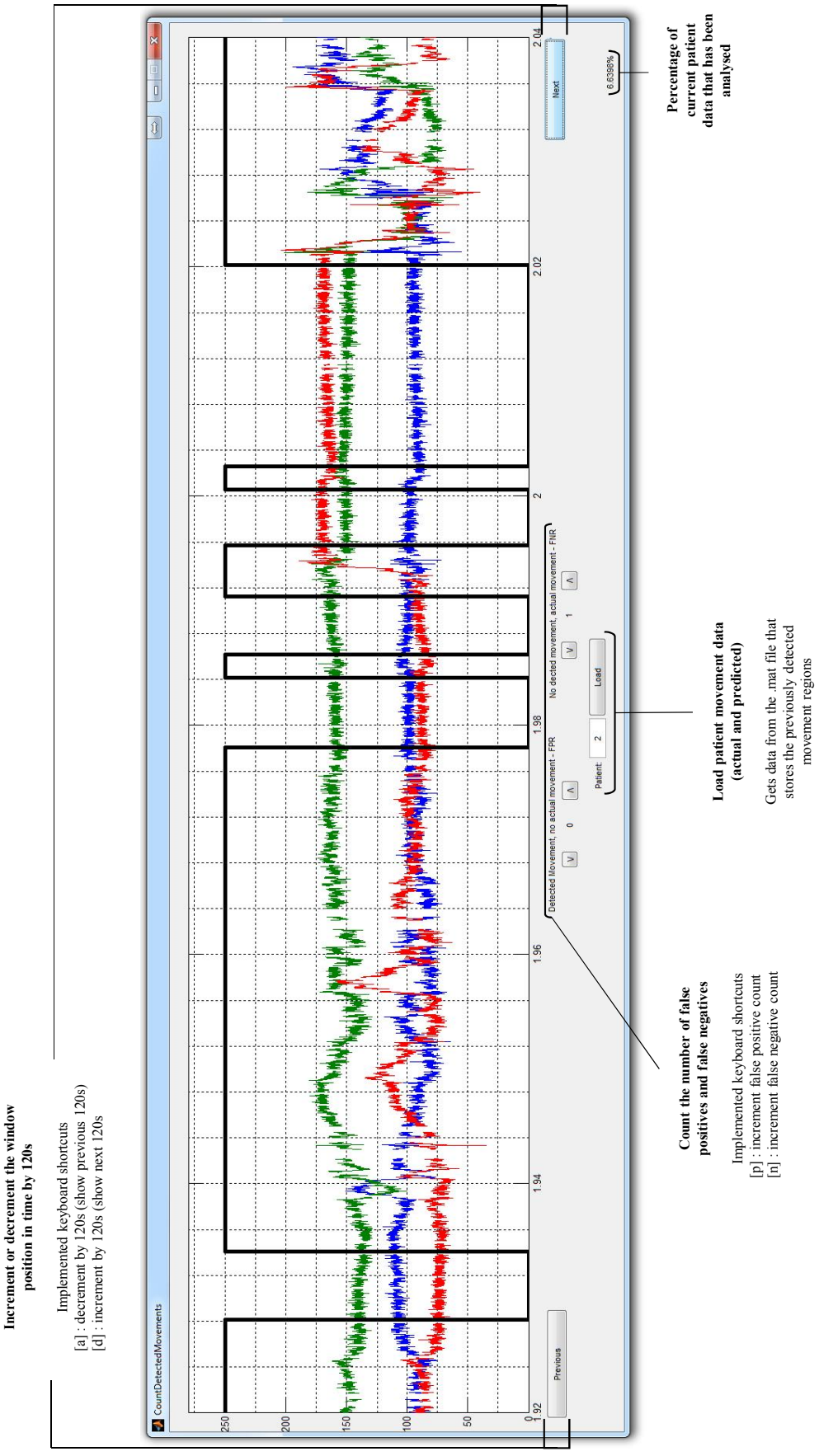


Figure D.1: Custom MATLAB graphical user interface software for manually assessing the accuracy of automated movement detection algorithm.

E

Additional Figures: Prevalence of arousal-related movements for different sensor placements

The number of movements that occur during sleep and wake for the finger, upper thorax, ankle, toe and all locations combined, and the number of arousals are shown below. Movement during sleep are shown as light blue, wake as grey and the total number of arousals are shown as dark blue. The percentage of arousals that coincide with movement $\%a$ are indicated by the overlapping regions. The percentage of movements across the night that occur during sleep $\%m$ and the percentage of movements that occur during both sleep and arousal $\%s$ are also shown.

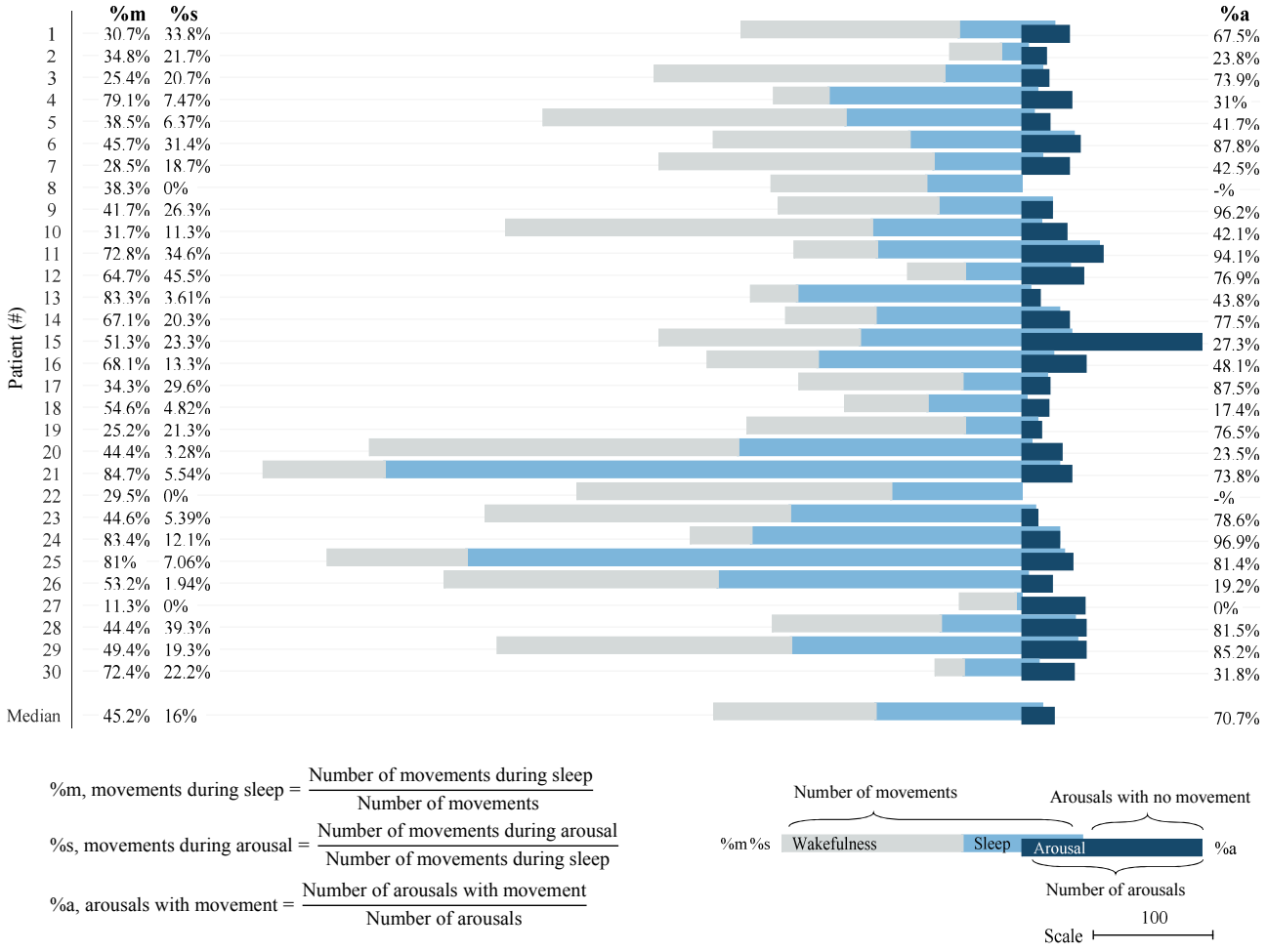


Figure E.1: Prevalence of movement during wake, sleep and arousal events for finger movement.

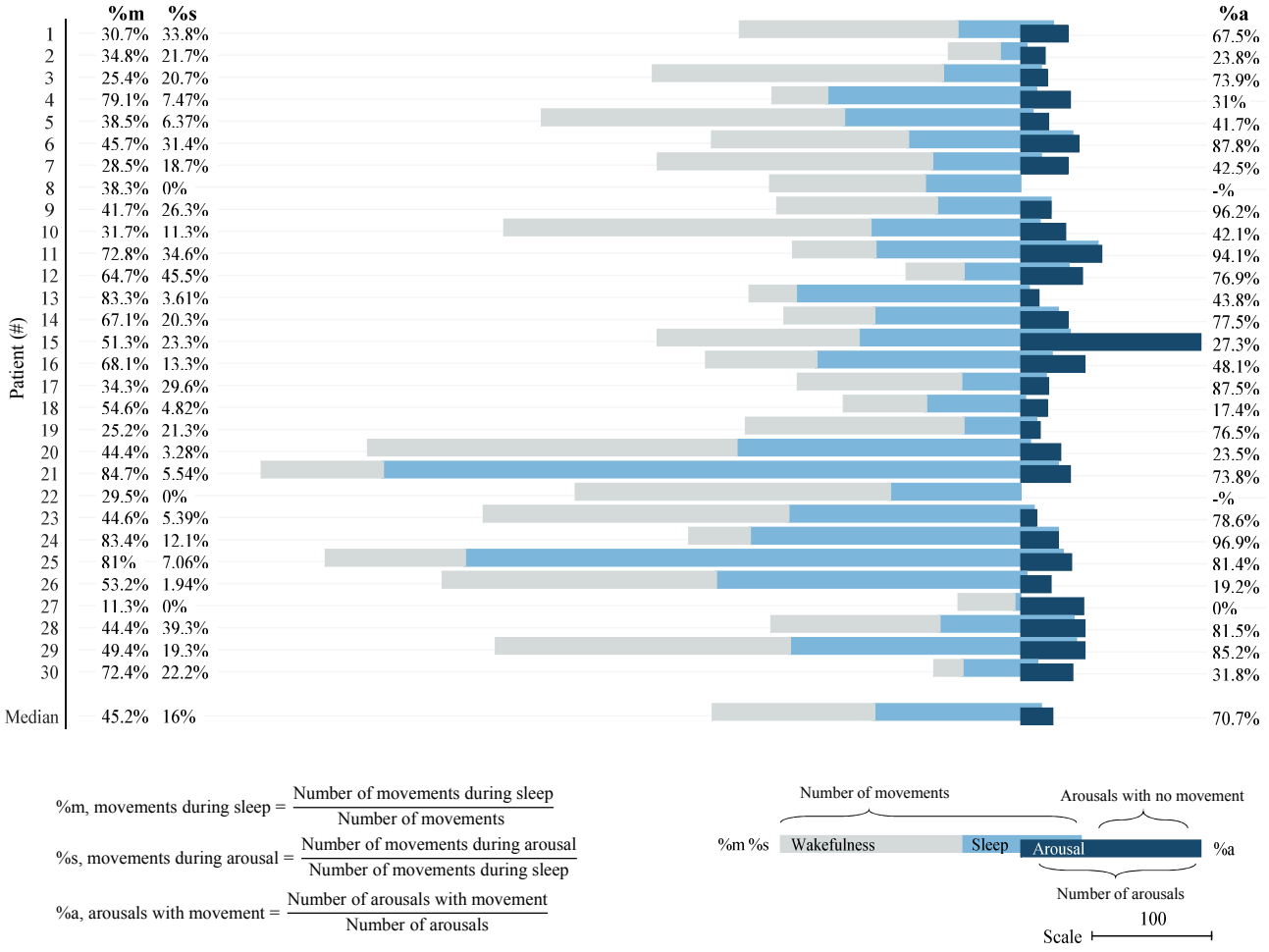


Figure E.2: Prevalence of movement during wake, sleep and arousal events for chest movement.

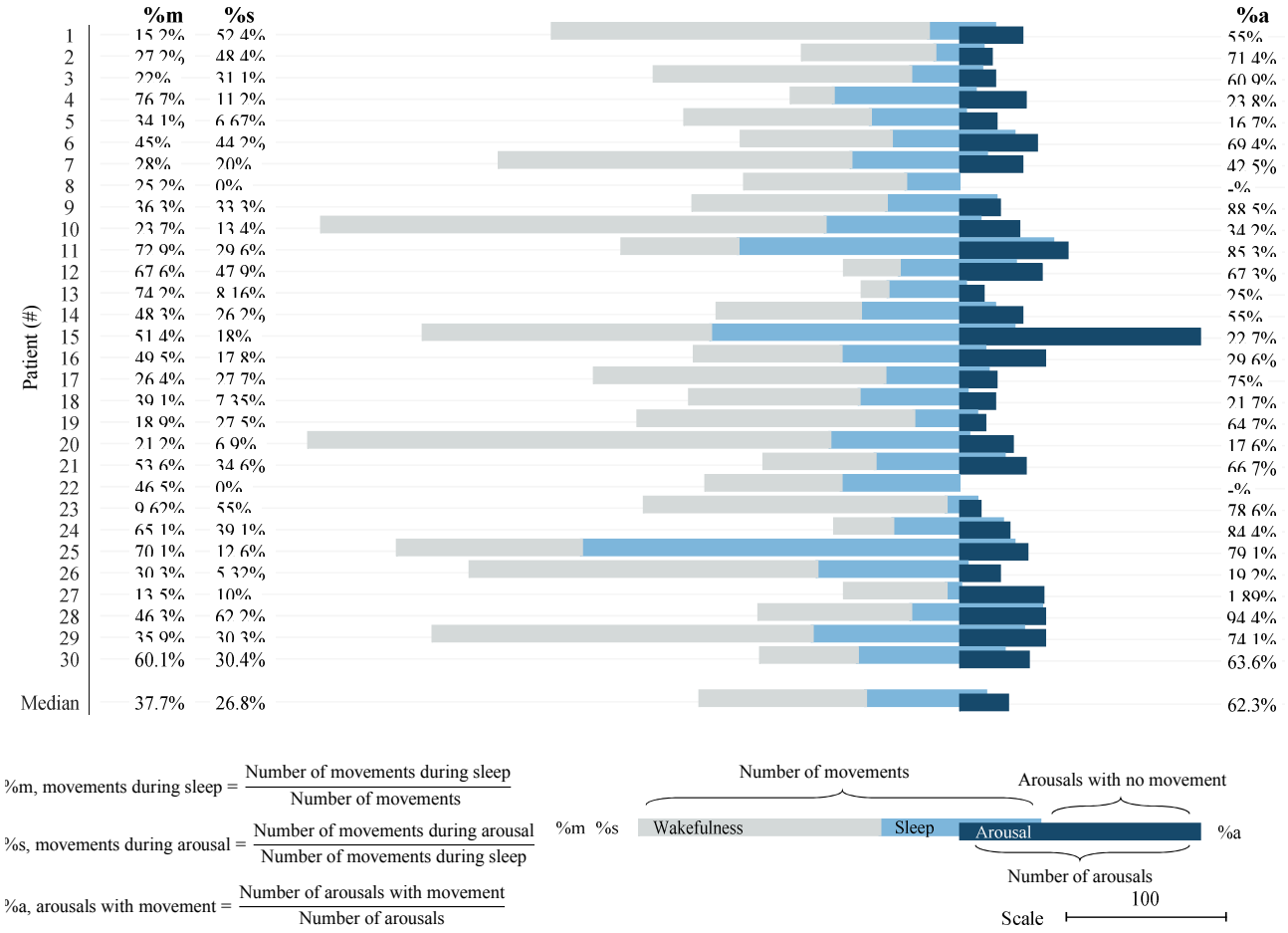


Figure E.3: Prevalence of movement during wake, sleep and arousal events for ankle movement.

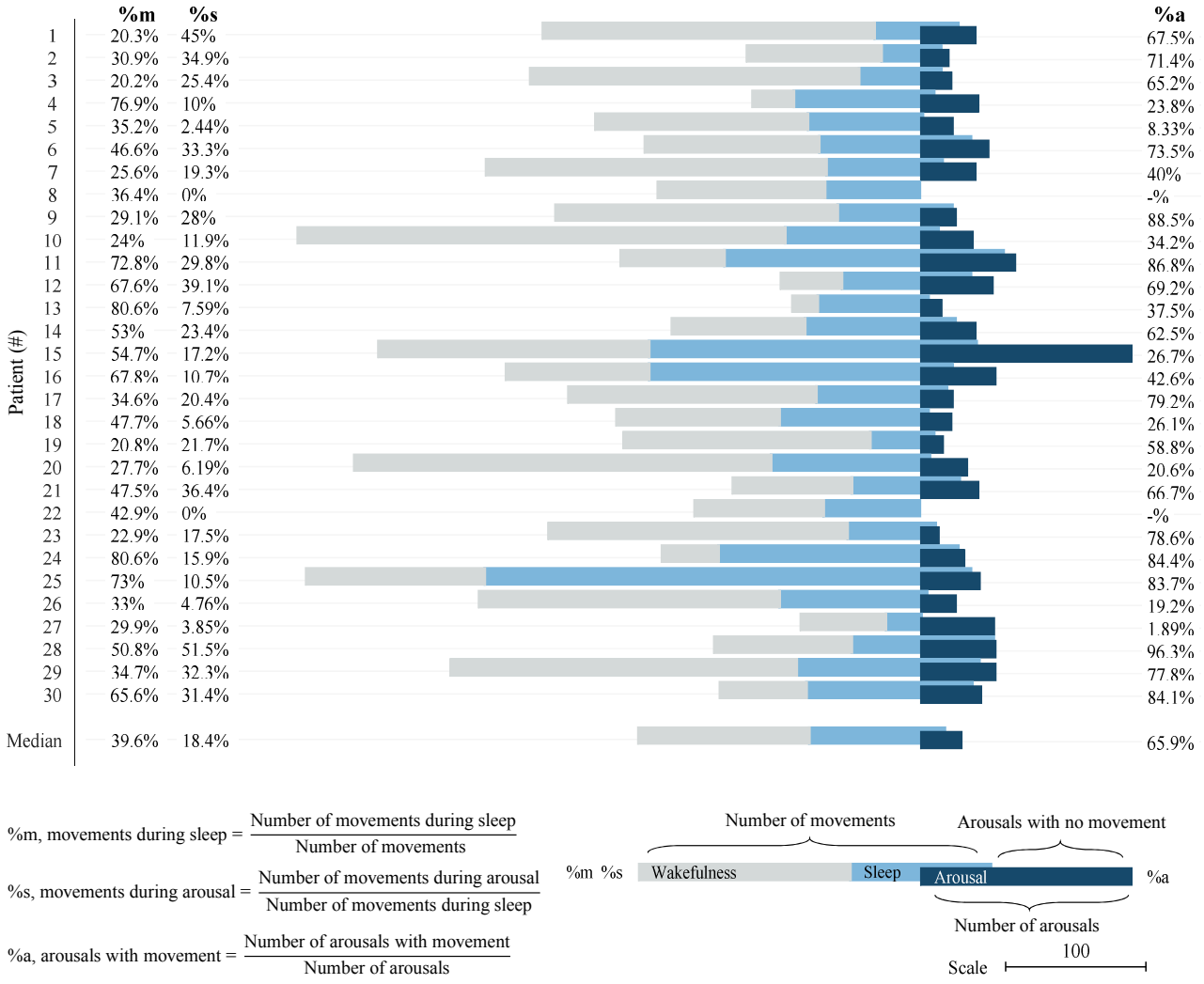


Figure E.4: Prevalence of movement during wake, sleep and arousal events for toe movement.

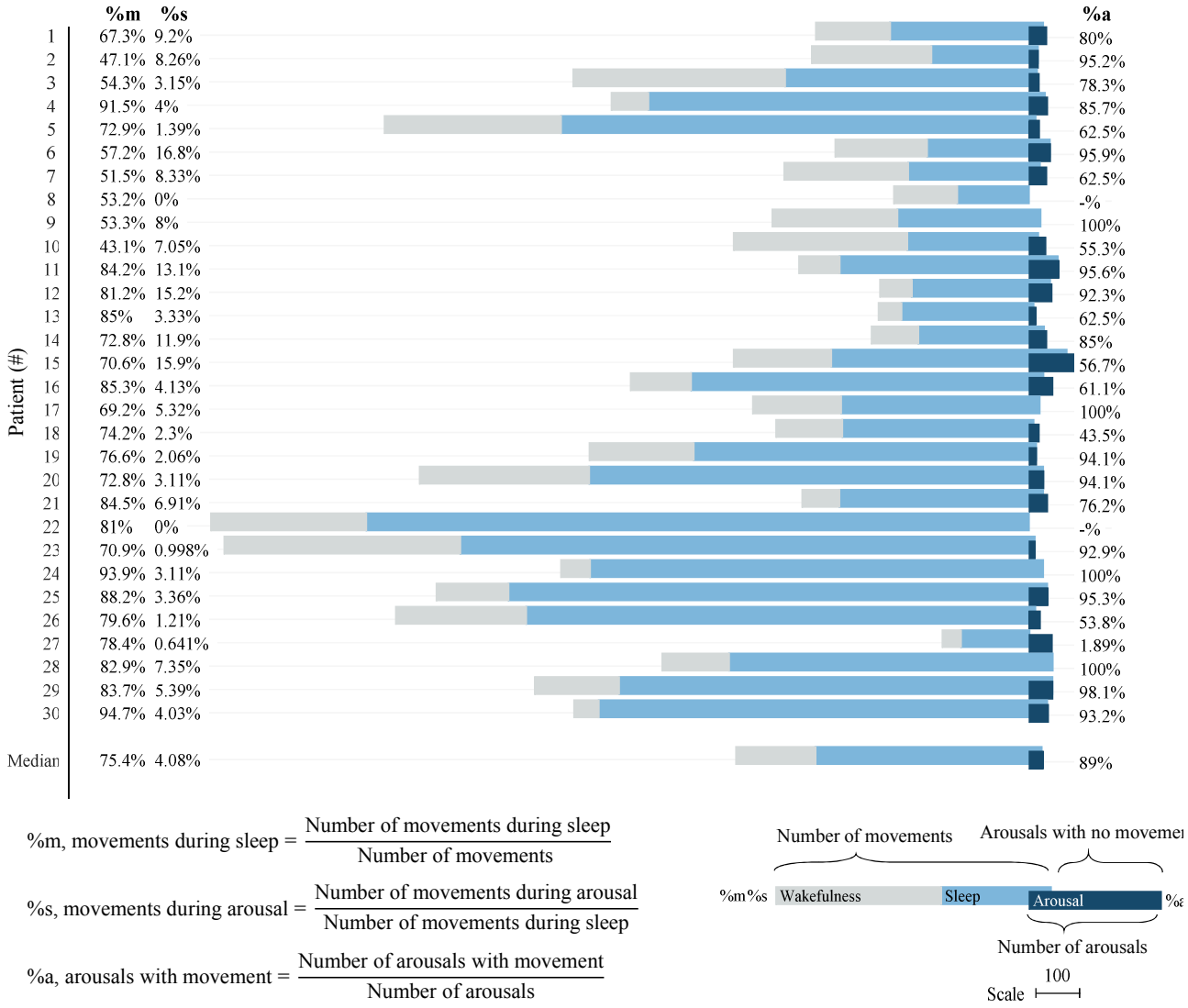
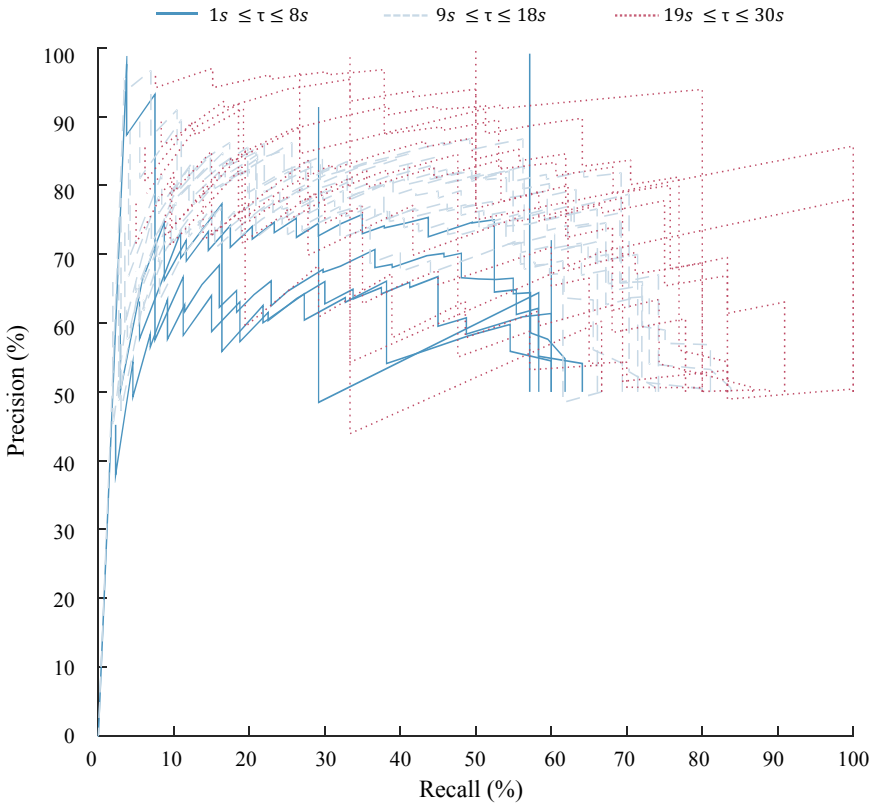


Figure E.5: Prevalence of movement during wake, sleep and arousal events for any (finger, wrist, chest, ankle or toe) movement.

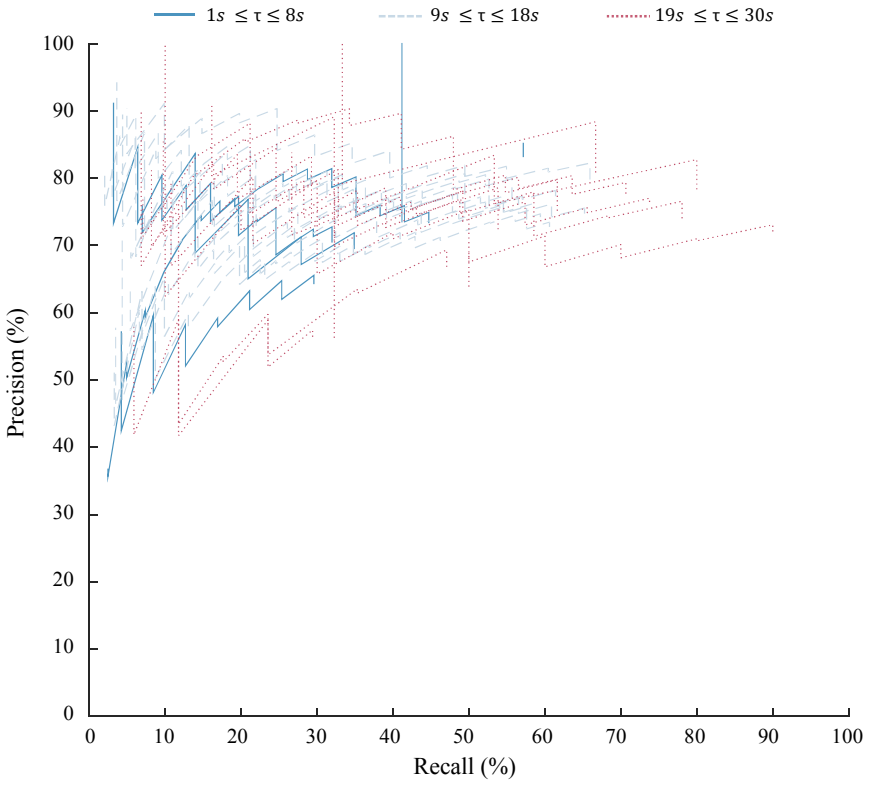
F

Additional Figures: Predicting arousal events of specific durations

Chapter 6 predicted arousal events greater than a specific duration. Using the same methodology as Section 6.2, this appendix extends that analysis to predict arousals of a specific duration. The following figures summarise the predictive performance. Note that the spread of arousal event duration across the cohort is fairly Gaussian; there are not many short-duration and long-duration arousal events relative to moderate-duration.

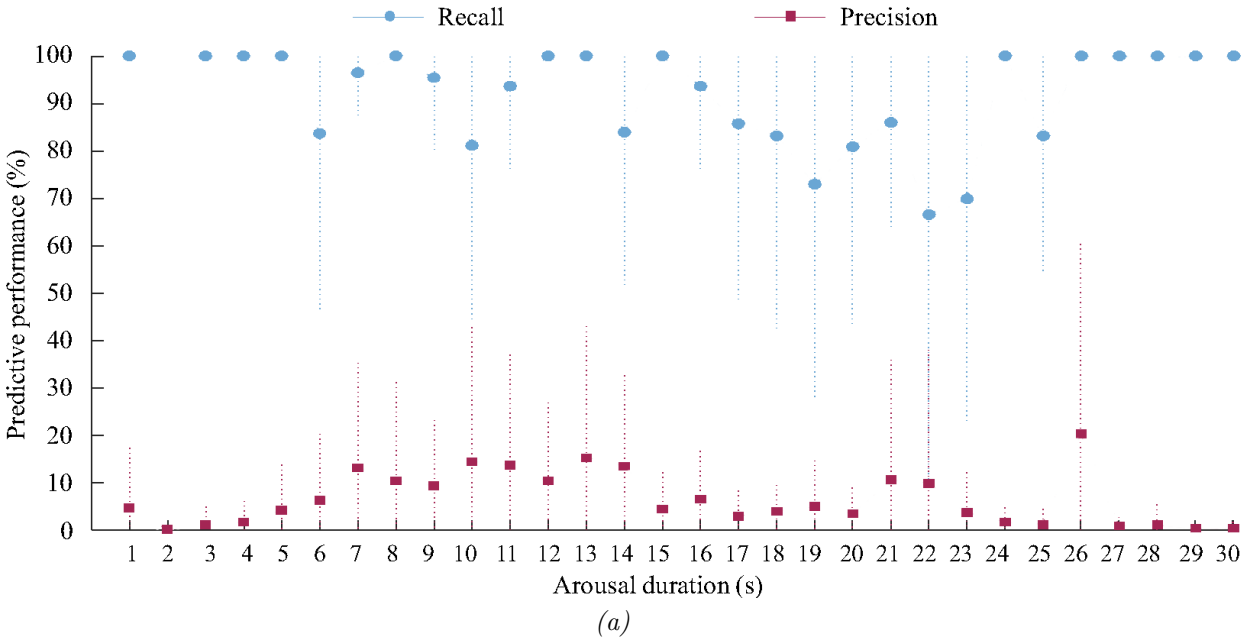


(a)

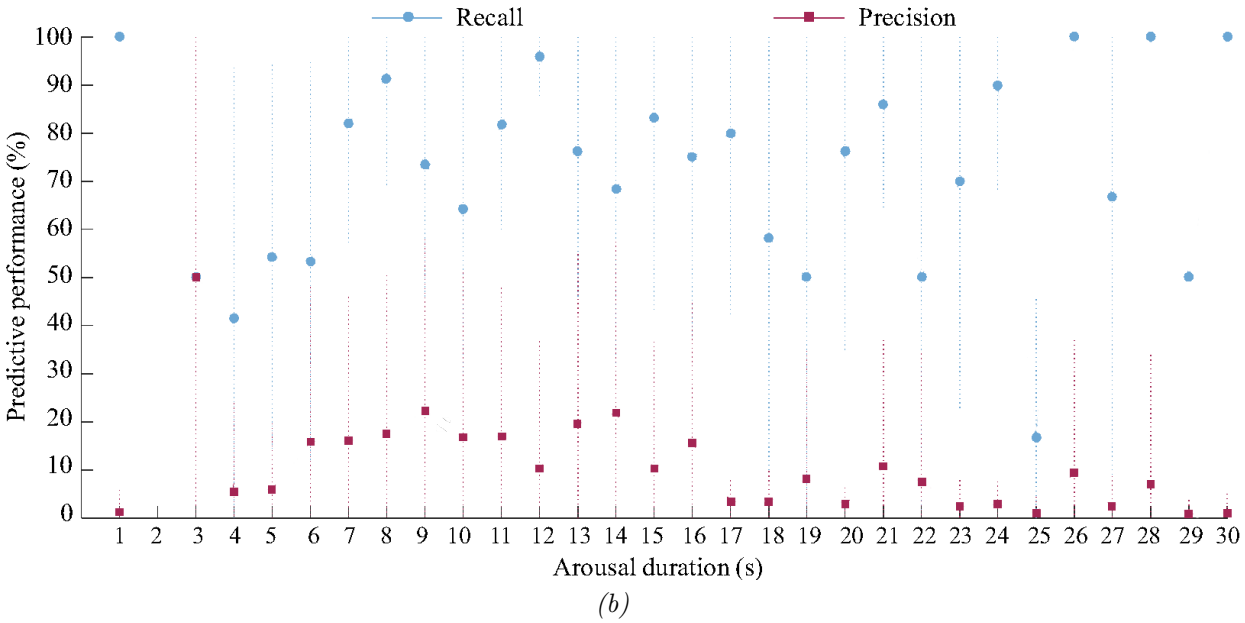


(b)

Figure F.1: Ability to detect arousal events of a specific duration for varying threshold values using (a) movement duration, and (b) zero-crossing summary of movement. Dark blue lines represent short-duration arousals (< 8s), light blue lines represent medium-duration arousals (between 9s and 18s) and red lines represent lengthy arousals (> 19s).



(a)



(b)

Figure F.2: Mean recall (blue) and precision (red) of detecting arousals greater than a specific duration (between 1s and 30s), using (a) movement duration, and (b) zero-crossing summary of movement. The vertical lines represent ±1 standard deviation.

Bibliography

[1] Aristoteles and W. A. Hammond, *Aristotle's psychology: a treatise on the principle of life.* S, Sonnenschein; The Macmillan, 1902.

[2] J. a. Mindell, "Sleep disorders in children." *Health psychology : official journal of the Division of Health Psychology, American Psychological Association*, vol. 12, no. 2, pp. 151–162, 1993.

[3] E. J. Stepanski, "The effect of sleep fragmentation on daytime function," *Sleep*, vol. 25, no. 3, pp. 268–278, 2002.

[4] E. S. Katz, "Arousal: Ontology, Functional Anatomy, Methodology, and Consequences," in *Sleep Disordered Breathing in Children.* Springer, 2012, pp. 105–119.

[5] A. Roebuck, V. Monasterio, E. Gederi, M. Osipov, J. Behar, A. Malhotra, T. Penzel, and G. D. Clifford, "A review of signals used in sleep analysis," *Physiological measurement*, vol. 35, no. 1, p. R1, 2014.

[6] S. Scholle and G. Zwacka, "Arousals and obstructive sleep apnea syndrome in children," *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, vol. 112, no. 6, pp. 984–991, 2001.

[7] C. Guilleminault, J. H. Lee, and A. Chan, "Pediatric obstructive sleep apnea syndrome." *Archives of pediatrics & adolescent medicine*, vol. 159, pp. 775–785, 2005.

[8] H. Colten, B. B. M. Altevogt, and H. Colten., *Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem*, 2006.

[9] C. L. Marcus, J. L. Carroll, C. B. Koerner, A. Hamer, J. Lutz, and G. M. Loughlin, "Determinants of growth in children with the obstructive sleep apnea syndrome." *The Journal of pediatrics*, vol. 125, no. 4, pp. 556–62, Oct. 1994.

[10] S. L. Blunden and D. W. Beebe, "The contribution of intermittent hypoxia, sleep debt and sleep disruption to daytime performance deficits in children: consideration of respiratory and non-respiratory sleep disorders." *Sleep medicine reviews*, vol. 10, no. 2, pp. 109–18, Apr. 2006.

[11] T. Young, P. E. Peppard, and D. J. Gottlieb, "Epidemiology of obstructive sleep apnea: A population health perspective," *American Journal of Respiratory and Critical Care Medicine*, vol. 165, pp. 1217–1239, 2002.

[12] J. R. Smith, "The Electroencephalogram During Normal Infancy and Childhood: III. Preliminary Observations on the Pattern Sequence During Sleep," *The Pedagogical Seminary and Journal of Genetic Psychology*, vol. 53, no. May 2015, pp. 471–482, 1938.

[13] J. D. Geyer, S. Talathi, and P. R. Carney, "Introduction to sleep and polysomnography," *Reading EEGs: a practical approach, Lippincott Williams & Wilkins, Philadelphia*, 2009.

[14] G. D. Church, "The role of polysomnography in diagnosing and treating obstructive sleep apnea in pediatric patients," *Current problems in pediatric and adolescent health care*, vol. 42, no. 1, pp. 2–25, 2012.

[15] A. Sadeh, "The role and validity of actigraphy in sleep medicine: an update." *Sleep medicine reviews*, vol. 15, no. 4, pp. 259–67, Aug. 2011.

[16] C. Crespo, M. Aboy, J. R. Fernández, and A. Mojón, "Automatic identification of activityrest periods based on actigraphy," *Medical and biological engineering and computing*, vol. 50, no. 4, pp. 329–340, 2012.

[17] A. Domingues, T. Paiva, and J. M. Sanches, "An actigraphy heterogeneous mixture model for sleep assessment," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*. IEEE, 2012, pp. 2275–2278.

[18] ——, "Statistical Characterization of Actigraphy Data for Sleep/Wake Assessment," *International Journal of Bioelectromagnetism*, vol. 15, no. 1, pp. 13–19, 2013.

[19] T. B. A. Mason II and A. I. Pack, "Pediatric Parasomnias," *Sleep*, vol. 30, no. 2, pp. 141–151, 2007.

[20] J. Parkka, M. Ermes, P. Korpipaa, J. Mantyjarvi, J. Peltola, and I. Korhonen, "Activity Classification Using Realistic Data From Wearable Sensors," *IEEE Transactions on Information Technology in Biomedicine*, vol. 10, no. 1, pp. 119–128, 2006.

[21] M. Ermes, J. Parkka, J. Mantyjarvi, I. Korhonen, and J. Pärkka, "Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions." *IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 12, no. 1, pp. 20–6, Jan. 2008.

[22] M. S. Blumberg, H. G. Marques, and F. Iida, "Twitching in sensorimotor development from sleeping rats to robots," *Current Biology*, vol. 23, no. 12, pp. R532—-R537, 2013.

[23] J. M. Siegel, "Why we sleep," *Scientific American*, vol. 289, no. 5, pp. 92–97, 2003.

[24] T. Wiedemann and K. Dowden, *Sleep, Volume 8 of Nottingham classical literature studies*. Levante Editori - Bari, 2003.

[25] E. F. Pace-Schott and J. A. Hobson, "The neurobiology of sleep: genetics, cellular physiology and subcortical networks," *Nature Reviews Neuroscience*, vol. 3, no. 8, pp. 591–605, 2002.

[26] E. Aserinsky and N. Kleitman, "Regularly occurring periods of eye motility, and concomitant phenomena, during sleep," *Science*, vol. 118, no. 3062, pp. 273–274, 1953.

[27] R. G. Astill, K. B. der Heijden, M. H. Van IJzendoorn, and E. J. W. Van Someren, "Sleep, cognition, and behavioral problems in school-age children: A century of research meta-analyzed." *Psychological bulletin*, vol. 138, no. 6, p. 1109, 2012.

[28] R. Stickgold, "Sleep-dependent memory consolidation," *Nature*, vol. 437, no. 7063, pp. 1272–1278, 2005.

[29] S. Gais, B. Lucas, and J. Born, "Sleep after learning aids memory recall," *Learning & Memory*, vol. 13, no. 3, pp. 259–262, 2006.

[30] W. H. Moorcroft and P. Belcher, *Understanding sleep and dreaming*. Springer, 2003.

[31] I. Iglowstein, O. G. Jenni, L. Molinari, and R. H. Largo, "Sleep duration from infancy to adolescence: reference values and generational trends," *Pediatrics*, vol. 111, no. 2, pp. 302–307, 2003.

[32] T. a. Lahti, S. Leppämäki, J. Lönnqvist, and T. Partonen, "Transition to daylight saving time reduces sleep duration plus sleep efficiency of the deprived sleep," *Neuroscience Letters*, vol. 406, pp. 174–177, 2006.

[33] T. Kantermann, M. Juda, M. Merrow, and T. Roenneberg, "The Human Circadian Clock's Seasonal Adjustment Is Disrupted by Daylight Saving Time," *Current Biology*, vol. 17, pp. 1996–2000, 2007.

[34] J. Arendt, "Biological Rhythms During Residence in Polar Regions," *Chronobiology International*, vol. 29, no. 4, pp. 379–394, 2012.

[35] M. A. Carskadon and W. C. Dement, "Normal human sleep: an overview," *Principles and practice of sleep medicine*, vol. 4, pp. 13–23, 2000.

[36] M. A. Carskadon and A. Rechtschaffen, "Monitoring and staging human sleep," *Principles and practice of sleep medicine*, vol. 3, pp. 1197–1215, 2000.

[37] B. Frauscher, V. Gschliesser, E. Brandauer, H. Ulmer, C. M. Peralta, J. Müller, W. Poewe, and B. Högl, "Video analysis of motor events in REM sleep behavior disorder," *Movement disorders*, vol. 22, no. 10, pp. 1464–1470, 2007.

[38] C. Iber, A. A. of Sleep Medicine, and Others, *The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications*. American Academy of Sleep Medicine, 2007.

[39] L. M. Hoey, P. Fulbrook, and J. A. Douglas, "Sleep assessment of hospitalised patients: A literature review," *International journal of nursing studies*, 2014.

[40] A. Rechtschaffen and A. Kales, "A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects," 1968.

[41] G. Tononi, "Sleep and dreaming," *The neurology of consciousness: Cognitive neuroscience and neuropathology*, pp. 89–107, 2009.

[42] R. B. Berry, R. Brooks, C. E. Gamaldo, S. M. Harding, C. L. Marcus, and B. V. Vaughn, "The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications, version 2.0," *Darien, Illinois: American Academy of Sleep Medicine*, 2012.

[43] ASTA/ASA, "ASTA/ASA Commentary on AASM Manual for the Scoring of Sleep and Associated Events," Tech. Rep., 2010.

[44] ——, "ASTA/ASA Addendum to AASM Guidelines for recording and scoring of paediatric sleep," Tech. Rep., 2011.

[45] R. E. Dahl and D. S. Lewin, "Pathways to adolescent health sleep regulation and behavior," *Journal of Adolescent Health*, vol. 31, no. 6, pp. 175–184, 2002.

[46] K. L. Knutson, K. Spiegel, P. Penev, and E. Van Cauter, "The metabolic consequences of sleep deprivation," *Sleep Medicine Reviews*, vol. 11, pp. 163–178, 2007.

[47] A. Sadeh, R. Gruber, and A. Raviv, "The effects of sleep restriction and extension on school-age children: what a difference an hour makes." *Child development*, vol. 74, no. 2, pp. 444–455, 2003.

[48] A. H. Messner and R. Pelayo, "Pediatric sleep-related breathing disorders," *American journal of otolaryngology*, vol. 21, no. 2, pp. 98–107, 2000.

[49] D. J. Eckert and M. K. Younes, "Arousal from sleep: implications for obstructive sleep apnea pathogenesis and treatment." *Journal of applied physiology (Bethesda, Md. : 1985)*, vol. 116, no. 3, pp. 302–13, Feb. 2014.

[50] M. H. Bonnet, K. Doghramji, T. Roehrs, E. J. Stepanski, S. H. Sheldon, A. S. Walters, M. Wise, and A. L. Chesson Jr, "The scoring of arousal in sleep: reliability, validity, and alternatives," *J Clin Sleep Med*, vol. 3, no. 2, pp. 133–145, 2007.

[51] A. S. D. A. ASDA, "EEG arousals: scoring rules and examples: a preliminary report from the Sleep Disorders Atlas Task Force of the American Sleep Disorders Association," *Sleep*, vol. 15, no. 2, pp. 173–184, 1992.

[52] D. Y. T. Goh, P. Galster, and C. L. Marcus, "Sleep architecture and respiratory disturbances in children with obstructive sleep apnea," *American journal of respiratory and critical care medicine*, vol. 162, no. 2 Pt 1, pp. 682–686, 2000.

[53] A. Sadeh, R. Gruber, and A. Raviv, "Sleep, Neurobehavioral Functioning, and Behavior Problems in School-Age Children," *Child Development*, vol. 73, no. 2, pp. 405–417, 2002.

[54] G. Jean-Louis, D. F. Kripke, R. J. Cole, J. D. Assmus, and R. D. Langer, "Sleep detection with an accelerometer actigraph: comparisons with polysomnography." *Physiology & behavior*, vol. 72, no. 1-2, pp. 21–8, Jan. 2001.

[55] G. Jean-Louis, D. F. Kripke, W. J. Mason, J. A. Elliott, and S. D. Youngstedt, "Sleep estimation from wrist movement quantified by different actigraphic modalities," *Journal of neuroscience methods*, vol. 105, no. 2, pp. 185–191, 2001.

[56] R. T. Brouillette, S. K. Fernbach, and C. E. Hunt, "Obstructive sleep apnea in infants and children," *The Journal of Pediatrics*, vol. 100, no. 1, pp. 31–40, 1982.

[57] C. H. Jones, J. a. Owens, and B. Pham, "Can a brief educational intervention improve parents' knowledge of healthy children's sleep? A pilot-test," *Health Education Journal*, 2012.

[58] C. L. Marcus and R. W. Emerson, "State of the art: Sleep-disordered breathing in children," *American journal of respiratory and critical care medicine*, vol. 164, no. 11, pp. 16–30, 2001.

[59] R. Arens and C. L. Marcus, "Pathophysiology of upper airway obstruction: a developmental perspective." *Sleep*, vol. 27, no. 5, pp. 997–1019, Aug. 2004.

[60] J. C. Lumeng and R. D. Chervin, "Epidemiology of pediatric obstructive sleep apnea," *Proceedings of the American Thoracic Society*, vol. 5, no. 2, p. 242, 2008.

[61] C. L. Marcus, S. A. McColley, J. L. Carroll, G. M. Loughlin, P. L. Smith, and A. R. Schwartz, "Upper airway collapsibility in children with obstructive sleep apnea syndrome," *Journal of Applied Physiology*, vol. 77, no. 2, pp. 918–924, 1994.

[62] E. Dayyat, L. Kheirandish-Gozal, and D. Gozal, "Childhood obstructive sleep apnea: one or two distinct disease entities?" *Sleep medicine clinics*, vol. 2, no. 3, pp. 433–444, 2007.

[63] S. Powell, H. Kubba, C. O'Brien, and M. Tremlett, "Paediatric obstructive sleep apnoea," *Clinical Otolaryngology*, vol. 35, no. 5, p. 418, 2010.

[64] F. H. Services, "When Your Child Has Obstructive Sleep Apnea (OSA)," May 2014. [Online]. Available: http://www.fairviewebenezer.org/HealthLibrary/Article/88985

[65] S. E. Beck and C. L. Marcus, "Pediatric polysomnography," *Sleep medicine clinics*, vol. 4, no. 3, pp. 393–406, 2009.

[66] R. Lawton, "Pediatric Polysomnogram," 2006.

[67] C. E. Woods, K. J. Usher, H. Jersmann, and G. P. Maguire, "Sleep disordered breathing and polysomnography in australia: Trends in provision from 2005 to 2012 and the impact of home-based diagnosis," *Journal of Clinical Sleep Medicine*, vol. 10, no. 7, pp. 767–772, 2014.

[68] S. H. Foundation and D. A. Economics, "Re-awakening Australia: The economic cost of sleep Sleep Health Foundation October 2011," no. October, p. 14, 2011.

[69] R. B. Mitchell, K. D. Pereira, and N. R. Friedman, "Sleep-Disordered Breathing in Children: Survey of Current Practice," *The Laryngoscope*, vol. 116, no. 6, pp. 956–958, 2006.

[70] P. S. Roland, R. M. Rosenfeld, L. J. Brooks, N. R. Friedman, J. Jones, T. W. Kim, S. Kuhar, R. B. Mitchell, M. D. Seidman, S. H. Sheldon, S. Jones, and P. Robertson, "Clinical Practice Guideline: Polysomnography for Sleep-Disordered Breathing Prior to Tonsillectomy in Children," *Otolaryngology – Head and Neck Surgery*, vol. 145, no. 1 suppl, pp. S1–S15, 2011.

[71] K. E. Bloch, "Polysomnography: a systematic review," *Technology and Health Care*, vol. 5, no. 4, pp. 285–305, 1997.

[72] C. Iber, S. Redline, A. M. K. Gilpin, S. F. Quan, L. Zhang, D. J. Gottlieb, D. Rapoport, H. E. Resnick, M. Sanders, and P. Smith, "Polysomnography performed in the unattended home versus the attended laboratory setting-Sleep Heart Health Study methodology," *Sleep*, vol. 27, no. 3, pp. 536–540, 2004.

[73] S. V. Jacob, A. Morielli, M. A. Mograss, F. M. Ducharme, M. D. Schloss, and R. T. Brouillette, "Home testing for pediatric obstructive sleep apnea syndrome secondary to adenotonsillar hypertrophy," *Pediatric pulmonology*, vol. 20, no. 4, pp. 241–252, 1995.

[74] J. G. Adair, "The Hawthorne effect: A reconsideration of the methodological artifact." *Journal of applied psychology*, vol. 69, no. 2, p. 334, 1984.

[75] S. Scholle, H. Scholle, A. Kemper, S. Glaser, B. Rieger, G. Kemper, G. Zwacka, and Others, "First night effect in children and adolescents undergoing polysomnography for sleep-disordered breathing," *Clinical neurophysiology*, vol. 114, no. 11, pp. 2138–2145, 2003.

[76] L. K. Brown, "Quantum physics and polysomnography: can we prevent the act of measuring sleep from changing sleep?" *Journal of clinical sleep medicine: JCSM: official publication of the American Academy of Sleep Medicine*, vol. 1, no. 2, pp. 133–135, 2005.

[77] Operation Hernia, "Pulse oximetry," 2015. [Online]. Available: http://www.operationhernia.org.uk/

[78] Wikimedia, "Wrist oximeter," 2015. [Online]. Available: http://upload.wikimedia.org/wikipedia/commons/7/7d/Wrist-oximeter.jpg

[79] N. Netzer, A. H. Eliasson, C. Netzer, and D. A. Kristo, "Overnight Pulse Oximetry for Sleep-Disordered Breathing in Adults*A Review," *CHEST Journal*, vol. 120, no. 2, pp. 625–633, 2001.

[80] 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, 2000.

[81] A. J. Williams, G. Yu, S. Santiago, and M. Stein, "Screening for sleep apnea using pulse oximetry and a clinical score," *CHEST Journal*, vol. 100, no. 3, pp. 631–635, 1991.

[82] A. Jubran, "Pulse oximetry," in *Applied Physiology in Intensive Care Medicine*, G. Hedenstierna, J. Mancebo, L. Brochard, and M. R. Pinsky, Eds. Springer Berlin Heidelberg, 2009, ch. 11, pp. 45–48.

[83] S. Ancoli-Israel, R. Cole, C. Alessi, M. Chambers, W. Moorcroft, and C. P. Pollak, "The role of actigraphy in the study of sleep and circadian rhythms." *Sleep*, vol. 26, no. 3, pp. 342–92, May 2003.

[84] L. J. Meltzer, H. E. Montgomery-Downs, S. P. Insana, and C. M. Walsh, "Use of actigraphy for assessment in pediatric sleep research," *Sleep Medicine Reviews*, 2012.

[85] BMedical, "Actiwatch 2," 2015. [Online]. Available: http://www.bmedical.com.au/shop/actiwatch-2-minimitter-philips.html

[86] A. Sadeh, J. Alster, D. Urbach, and P. Lavie, "Actigraphically based automatic bedtime sleep-wake scoring: validity and clinical applications," *Journal of Ambulatory Monitoring*, vol. 2, no. 3, pp. 209–216, 1989.

[87] A. Sadeh, P. Lavie, A. Scher, and R. Epstein, "Actigraphic and Control Infants and Young Children : A New Method for Pediatric Assessment of Sleep- Wake Patterns PhD ;," vol. 87, no. 4, pp. 494–499, 1991.

[88] A. Sadeh, "Assessment of intervention for infant night waking: parental reports and activity-based home monitoring," *Journal of consulting and clinical psychology*, vol. 62, no. 1, pp. 63–68, 1994.

[89] A. Sadeh, C. Acebo, R. Seifer, S. Aytur, and M. A. Carskadon, "Activity-based assessment of sleep-wake patterns during the 1st year of life," *Infant Behavior and Development*, vol. 18, no. 3, pp. 329–337, Sep. 1995.

[90] A. Sadeh, "Evaluating Night Wakings 1n SleepDisturbed Infants: A Methodological Study of Parental Reports and Actigraphy," *Sleep*, vol. 19, no. 10, pp. 757–762, 1996.

[91] B. Gnidovec, D. Neubauer, and J. Zidar, "Actigraphic assessment of sleep–wake rhythm during the first 6 months of life," *Clinical Neurophysiology*, vol. 113, no. 11, pp. 1815–1821, 2002.

[92] E. Sazonov, N. Sazonova, S. Schuckers, M. Neuman, C. S. Group, and Others, "Activity-based sleep–wake identification in infants," *Physiological measurement*, vol. 25, no. 5, p. 1291, 2004.

[93] C. Acebo, A. Sadeh, R. Seifer, O. Tzischinsky, A. Hafer, and M. A. Carskadon, "Sleep-/wake patterns derived from activity monitoring and maternal report for healthy 1- to 5-year-old children," *Sleep*, vol. 28, no. 12, pp. 1568–1577, 2005.

[94] K. So, P. Buckley, T. M. Adamson, and R. S. C. Horne, "Actigraphy correctly predicts sleep behavior in infants who are younger than six months, when compared with polysomnography." *Pediatric research*, vol. 58, no. 4, pp. 761–5, Oct. 2005.

[95] M. Hyde, D. M. O'DRISCOLL, S. Binette, C. Galang, S. K. Tan, N. Verginis, M. J. Davey, and R. S. C. Horne, "Validation of actigraphy for determining sleep and wake in children with sleep disordered breathing," *Journal of sleep research*, vol. 16, no. 2, pp. 213–216, 2007.

[96] N. L. Johnson, H. L. Kirchner, C. L. Rosen, A. Storfer-lsser, L. N. Cartar, S. Ancoli-Israel, J. L. Emancipator, A. M. Kibler, and S. Redline, "Sleep estimation using wrist

actigraphy in adolescents with and without sleep disordered breathing: a comparison of three data modes," *Sleep*, vol. 30, no. 7, p. 899, 2007.

[97] H. Werner, L. Molinari, C. Guyer, and O. G. Jenni, "Agreement rates between actigraphy, diary, and questionnaire for children's sleep patterns," *Archives of pediatrics & adolescent medicine*, vol. 162, no. 4, pp. 350–358, 2008.

[98] S. L. Sitnick, B. L. Goodlin-Jones, and T. F. Anders, "The use of actigraphy to study sleep disorders in preschoolers: some concerns about detection of nighttime awakenings," *Sleep*, vol. 31, no. 3, p. 395, 2008.

[99] M. Sung, T. M. Adamson, and R. S. C. Horne, "Validation of actigraphy for determining sleep and wake in preterm infants," *Acta Paediatrica*, vol. 98, no. 1, pp. 52–57, 2009.

[100] D. M. O'Driscoll, A. M. Foster, M. J. Davey, G. M. Nixon, and R. S. C. Horne, "Can actigraphy measure sleep fragmentation in children?" *Archives of disease in childhood*, vol. 95, no. 12, pp. 1031–1033, 2010.

[101] S. Holley, C. M. Hill, and J. Stevenson, "A comparison of actigraphy and parental report of sleep habits in typically developing children aged 6 to 11 years," *Behavioral sleep medicine*, vol. 8, no. 1, pp. 16–27, 2010.

[102] S. P. Insana, D. Gozal, and H. E. Montgomery-Downs, "Invalidity of one actigraphy brand for identifying sleep and wake among infants," *Sleep medicine*, vol. 11, no. 2, pp. 191–196, 2010.

[103] A. R. Weiss, N. L. Johnson, N. A. Berger, and S. Redline, "Validity of activity-based devices to estimate sleep," *Journal of clinical sleep medicine: JCSM: official publication of the American Academy of Sleep Medicine*, vol. 6, no. 4, p. 336, 2010.

[104] A. M. Gregory, J. C. Cousins, E. E. Forbes, L. Trubnick, N. D. Ryan, D. A. Axelson, B. Birmaher, A. Sadeh, and R. E. Dahl, "Sleep items in the child behavior checklist: a comparison with sleep diaries, actigraphy, and polysomnography," *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 50, no. 5, pp. 499–507, 2011.

[105] K. Spruyt, D. Gozal, E. Dayyat, A. Roman, and D. L. Molfese, "Sleep assessments in healthy school-aged children using actigraphy: concordance with polysomnography," *Journal of sleep research*, vol. 20, no. 1pt2, pp. 223–232, 2011.

[106] E. A. Dayyat, K. Spruyt, D. L. Molfese, and D. Gozal, "Sleep estimates in children: parental versus actigraphic assessments," *Nature and science of sleep*, vol. 3, p. 115, 2011.

[107] M. A. Short, M. Gradisar, L. C. Lack, H. Wright, and M. A. Carskadon, "The discrepancy between actigraphic and sleep diary measures of sleep in adolescents," *Sleep medicine*, vol. 13, no. 4, pp. 378–384, 2012.

[108] T. M. Ward, M. Lentz, G. M. Kieckhefer, and C. A. Landis, "Polysomnography and actigraphy concordance in juvenile idiopathic arthritis, asthma and healthy children," *Journal of sleep research*, vol. 21, no. 1, pp. 113–121, 2012.

[109] K. Boyne, D. D. Sherry, P. R. Gallagher, M. Olsen, and L. J. Brooks, "Accuracy of computer algorithms and the human eye in scoring actigraphy," *Sleep and Breathing*, vol. 17, no. 1, pp. 411–417, 2013.

[110] J. A. Owens, A. Spirito, and M. McGuinn, "The Children's Sleep Habits Questionnaire (CSHQ): psychometric properties of a survey instrument for school-aged children," *Sleep*, vol. 23, no. 8, pp. 1043–1052, 2000.

[111] M.-H. Bao, *Micro mechanical transducers: pressure sensors, accelerometers and gyroscopes*. Elsevier, 2000, vol. 8.

[112] T. M. E. Nijsen, R. M. Aarts, P. J. M. Cluitmans, and P. a. M. Griep, "Time-frequency analysis of accelerometry data for detection of myoclonic seizures." *IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 14, no. 5, pp. 1197–203, Sep. 2010.

[113] J. Y. Hwang, J. M. Kang, Y. W. Jang, and H. C. Kim, "Development of novel algorithm and real-time monitoring ambulatory system using Bluetooth module for fall detection in the elderly," in *Engineering in Medicine and Biology Society, 2004. IEMBS '04. 26th Annual International Conference of the IEEE*, vol. 1, 2004, pp. 2204–2207.

[114] J. Chen, K. Karric, D. Chang, J. Luk, and R. Bajcsy, "Wearable Sensors for Reliable Fall Detection," in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, 2005, pp. 3551–3554.

[115] S. J. Preece, J. Y. Goulermas, L. P. J. Kenney, D. Howard, K. Meijer, and R. Crompton, "Activity identification using body-mounted sensorsa review of classification techniques," *Physiological Measurement*, vol. 30, no. 4, p. R1, 2009.

[116] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," *Pervasive Computing*, vol. 3001, pp. 1–17, 2004.

[117] M. Sekine, T. Tamura, M. Akay, T. Fujimoto, T. Togawa, and Y. Fukui, "Discrimination of walking patterns using wavelet-based fractal analysis," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 10, no. 3, pp. 188–196, 2002.

[118] N. Bidargaddi, A. Sarela, J. Boyle, V. Cheung, M. Karunanithi, L. Klingbei, C. Yelland, and L. Gray, "Wavelet based approach for posture transition estimation using a waist worn accelerometer," in *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, 2007, pp. 1884–1887.

[119] N. H. Lovell, W. Ning, E. Ambikairajah, and B. G. Celler, "Accelerometry Based Classification of Walking Patterns Using Time-frequency Analysis," in *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, 2007, pp. 4899–4902.

[120] D. Gabor, "Theory of Communication," pp. 429–457, 1946.

[121] J. B. Webster, D. Kripke, S. Messin, D. Mullaney, and G. Wyboney, "An activity-based sleep monitor system for ambulatory use," *Sleep: Journal of Sleep Research & Sleep Medicine*, vol. 5, no. 4, pp. 389–399, 1982.

[122] R. J. Cole, D. F. Kripke, W. Gruen, D. J. Mullaney, and J. C. Gillin, "Automatic sleep-/wake identification from wrist activity," *Sleep*, vol. 15, no. 5, pp. 461–469.

[123] N. R. Oakley, "Validation with polysomnography of the Sleepwatch sleep/wake scoring algorithm used by the Actiwatch activity monitoring system." Mini Mitter Co., Inc., Tech. Rep., 1997.

[124] C. A. Kushida, A. Chang, C. Gadkary, C. Guilleminault, O. Carrillo, and W. C. Dement, "Comparison of actigraphic, polysomnographic, and subjective assessment of sleep parameters in sleep-disordered patients." *Sleep Medicine*, vol. 2, no. 5, pp. 389–396, 2001.

[125] J. M. Lachin, "Introduction to sample size determination and power analysis for clinical trials," *Controlled Clinical Trials*, vol. 2, no. 2, pp. 93–113, 1981.

[126] T. J. Cole, K. M. Flegal, D. Nicholls, and A. a. Jackson, "Body mass index cut offs to define thinness in children and adolescents: international survey." *BMJ (Clinical research ed.)*, vol. 335, p. 194, 2007.

[127] T. J. Cole, M. C. Bellizzi, K. M. Flegal, and W. H. Dietz, "and Obesity Worldwide : International Survey," *Bmj*, vol. 320, no. table 1, pp. 1–6, 2000.

[128] TSANZ, "Accreditation of sleep disorders services - Including standards for paediatric laboratories," Tech. Rep., 2009.

[129] A. Oppenheim and J. Lim, "The importance of phase in signals," *Proceedings of the IEEE*, vol. 69, no. 5, pp. 529–541, 1981.

[130] A. P. Bradley, "The use of the area under the ROC curve in the evaluation of machine learning algorithms," *Pattern Recognition*, vol. 30, no. 7, pp. 1145–1159, 1997.

[131] A. R. Feinstein and D. V. Cicchetti, "High agreement but low Kappa: I. the problems of two paradoxes," *Journal of Clinical Epidemiology*, vol. 43, no. 6, pp. 543–549, 1990.

[132] J. Cohen, "Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit." *Psychological bulletin*, vol. 70, no. 4, pp. 213–220, 1968.

[133] R. DerSimonian and N. Laird, "Meta-analysis in clinical trials," *Controlled clinical trials*, vol. 7, no. 3, pp. 177–188, 1986.

[134] M. Egger, G. D. Smith, and A. N. Phillips, "Meta-analysis: principles and procedures." *BMJ: British Medical Journal*, vol. 315, no. 7121, p. 1533, 1997.

[135] F. Galton, "A new instrument for measuring the rate of movement of the various limbs," *The Journal of the Anthropological Institute of Great Britain and Ireland*, vol. 20, pp. 200–204, 1891.

[136] D. Redmond and F. Hegge, "Observations on the design and specification of a wrist-worn human activity monitoring system," *Behavior Research Methods, Instruments, & Computers*, vol. 17, no. 6, pp. 659–669, 1985.

[137] B. M. Maija-Riikka Steenari and E. T. Aronen, "Actigraph placement and sleep estimation in children," *Sleep*, vol. 25, no. 2, p. 235, 2002.

[138] M. F. Hjorth, J.-P. Chaput, C. T. Damsgaard, S.-M. Dalskov, K. F. Michaelsen, I. Tetens, and A. Sjödin, "Measure of sleep and physical activity by a single accelerometer: Can a waist-worn Actigraph adequately measure sleep in children?" *Sleep and Biological Rhythms*, vol. 10, no. 4, pp. 328–335, 2012.

[139] K. W. Adkins, S. E. Goldman, D. Fawkes, K. Surdyka, L. Wang, Y. Song, and B. A. Malow, "A pilot study of shoulder placement for actigraphy in children," *Behavioral Sleep Medicine*, vol. 10, no. 2, pp. 138–147, 2012.

[140] M. C. Souders, T. B. A. Mason, O. Valladares, M. Bucan, S. E. Levy, D. S. Mandell, T. E. Weaver, and J. Pinto-Martin, "Sleep behaviors and sleep quality in children with autism spectrum disorders," *Sleep*, vol. 32, no. 12, p. 1566, 2009.

[141] H. Gjoreski, M. Lustrek, and M. Gams, "Accelerometer placement for posture recognition and fall detection," in *Intelligent Environments (IE), 2011 7th International Conference on*. IEEE, 2011, pp. 47–54.

[142] D. Gravem, M. Singh, C. Chen, J. Rich, J. Vaughan, K. Goldberg, F. Waffarn, P. Chou, D. Cooper, D. Reinkensmeyer, and D. Patterson, "Assessment of Infant Movement With a Compact Wireless Accelerometer System," *Journal of Medical Devices*, vol. 6, no. 2, pp. 21 013–21 017, 2012.

[143] L. Sheng-Fu, Y. Chung-Ping, C. Da-Wei, S. Fu-Zen, L. You-De, L. Yi-Che, and C. Jing-Jhong, "Development of an actigraph system for sleep-wake identification," in *Instrumentation and Measurement Technology Conference (I2MTC), 2011 IEEE*, 2011, pp. 1–6.

[144] M. L. Lamprecht, A. P. Bradley, T. Tran, A. Boynton, and P. I. Terrill, "Multisite accelerometry for sleep and wake classification in children." *Physiological measurement*, vol. 36, no. 1, pp. 133–47, Jan. 2015.

[145] L. E. Dodd and M. S. Pepe, "Partial AUC estimation and regression," *Biometrics*, vol. 59, no. 3, pp. 614–623, 2003.

[146] L. Yu and H. Liu, "Efficient feature selection via analysis of relevance and redundancy," *The Journal of Machine Learning Research*, vol. 5, pp. 1205–1224, 2004.

[147] K. J. Cios, W. Pedrycz, and R. W. Swiniarski, *Data Mining and Knowledge Discovery*. Springer, 1998.

[148] H. Liu and H. Motoda, *Feature selection for knowledge discovery and data mining*. Springer, 1998.

[149] S. Thiemjarus, B. P. L. Lo, K. V. Laerhoven, and G. Z. Yang, "Feature selection for wireless sensor networks," in *the 1st International Workshop on Wearable and Implantable Body Sensor Networks. Imperial, London, UK*, 2004.

[150] P. Mitra, C. A. Murthy, and S. K. Pal, "Unsupervised feature selection using feature similarity," *IEEE transactions on pattern analysis and machine intelligence*, vol. 24, no. 3, pp. 301–312, 2002.

[151] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 8, pp. 1226–1238, 2005.

[152] Y. Saeys, I. n. Inza, and P. Larrañaga, "A review of feature selection techniques in bioinformatics," *Bioinformatics*, vol. 23, no. 19, pp. 2507–2517, 2007.

[153] C. Ding and H. Peng, "Minimum redundancy feature selection from microarray gene expression data," *Journal of bioinformatics and computational biology*, vol. 3, no. 02, pp. 185–205, 2005.

[154] Z. He, J. Zhang, X.-H. Shi, L.-L. Hiu, X. Kong, Y.-D. Cai, and K.-C. Chou, "Predicting Drug-Target Interaction Networks Based on Functional Groups and Biological Features," *PLoS ONE*, vol. 5, no. 3, p. e9603, 2010.

[155] M. Hutter, "Distribution of mutual information," *Advances in Neural Information Processing Systems*, vol. 1, pp. 399–406, 2002.

[156] P. A. Lachenbruch and M. Goldstein, "Discriminant analysis," *Biometrics*, pp. 69–85, 1979.

[157] A. R. Wolfson and H. E. Montgomery-Downs, *The Oxford handbook of infant, child, and adolescent sleep and behavior*, 2013.

[158] H. A. M. Middelkoop, E. M. Dam, D. A. Smildevan Den Doel, and G. Dijk, "45-Hour continuous quintuple-site actimetry: Relations between trunk and limb movements and effects of circadian sleepwake rhythmicity," *Psychophysiology*, vol. 34, no. 2, pp. 199–203, 1997.

[159] G. Giddings, "Normal sleep pattern for children: factors which derange such a pattern (physical factors)," *Journal of the American Medical Association*, vol. 102, no. 7, pp. 525–529, 1934.

[160] C. H. Schenck, S. R. Bundlie, M. G. Ettinger, and M. W. Mahowald, "Chronic Behavioral Disorders of Human REM Sleep : A New Category of Parasomnia," vol. 9, no. December 1985, pp. 293–308, 1986.

[161] H. Yoshimi, K. Sasaguri, K. Tamaki, and S. Sato, "Identification of the occurrence and pattern of masseter muscle activities during sleep using EMG and accelerometer systems," *Head & face medicine*, vol. 5, no. 1, p. 7, 2009.

[162] J. Lockman, R. S. Fisher, and D. M. Olson, "Detection of seizure-like movements using a wrist accelerometer," *Epilepsy & Behavior*, vol. 20, no. 4, pp. 638–641, 2011.

[163] J. Wilde-Frenz and H. Schulz, "Rate and distribution of body movements during sleep in humans," *Perceptual and motor skills*, vol. 56, no. 1, pp. 275–283, 1983.

[164] S. Gori, G. Ficca, F. Giganti, I. D. Nasso, L. Murri, and P. Salzarulo, "Body movements during night sleep in healthy elderly subjects and their relationships with sleep stages," *Brain research bulletin*, vol. 63, no. 5, pp. 393–397, 2004.

[165] F. Giganti, G. Ficca, S. Gori, and P. Salzarulo, "Body movements during night sleep and their relationship with sleep stages are further modified in very old subjects," *Brain research bulletin*, vol. 75, no. 1, pp. 66–69, 2008.

[166] T. Chau, "A review of analytical techniques for gait data. Part 2: neural network and wavelet methods." *Gait & posture*, vol. 13, no. 2, pp. 102–20, Apr. 2001.

[167] M. B. I. Raez, M. S. Hussain, and F. Mohd-Yasin, "Techniques of EMG signal analysis: detection, processing, classification and applications," *Biological procedures online*, vol. 8, no. 1, pp. 11–35, 2006.

[168] T. M. E. Nijsen, A. J. E. M. Janssen, F. Ieee, and R. M. Aarts, "Analysis of a wavelet arising from a model for arm movements during epileptic seizures," *ProRisc*, no. 5, pp. 1–4, 2007.

[169] O. Geman, C. Zamfir, and G. Asachi, "Using wavelet for early detection of pathological tremor," in *Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*, no. Eusipco, 2012, pp. 1723–1727.

[170] S. T. Aaronson, S. Rashed, M. P. Biber, and J. A. Hobson, "Brain state and body position: A time-lapse video study of sleep," *Archives of General Psychiatry*, vol. 39, no. 3, pp. 330–335, 1982.

[171] S. J. Preece, J. Y. Goulermas, L. P. J. Kenney, and D. Howard, "A comparison of feature extraction methods for the classification of dynamic activities from accelerometer data." *IEEE transactions on bio-medical engineering*, vol. 56, no. 3, pp. 871–9, Mar. 2009.

[172] M. L. Lamprecht, P. I. Terrill, C. L. Parsley, and A. P. Bradley, "Characterization of Movements during Restless Sleep in Children : A Pilot Study," *36th Annual International Conference of the IEEE Engineering in Medicince and Biology Society*, pp. 274–277, 2014.

[173] A. P. Bradley, "Shift-invariance in the discrete wavelet transform," *Proceedings of VIIth Digital Image Computing: Techniques and Applications. Sydney*, 2003.

[174] M. Shensa, "The discrete wavelet transform: wedding the a trous and Mallat algorithms," *Signal Processing, IEEE Transactions on*, vol. 40, no. 10, pp. 2464–2482, 1992.

[175] M. González-Audícana, X. Otazu, O. Fors, and A. Seco, "Comparison between Mallat's and the à trous discrete wavelet transform based algorithms for the fusion of multispectral and panchromatic images," *International Journal of Remote Sensing*, vol. 26, no. 3, pp. 595–614, 2005.

[176] G. Hansson, P. Asterland, N.-G. Holmer, and S. Skerfving, "Validity and reliability of triaxial accelerometers for inclinometry in posture analysis," *Medical and Biological Engineering and Computing*, vol. 39, no. 4, pp. 405–413, 2001.

[177] C. V. C. Bouten, A. A. H. J. Sauren, M. Verduin, and J. D. Janssen, "Effects of placement and orientation of body-fixed accelerometers on the assessment of energy expenditure during walking," *Medical and Biological Engineering and Computing*, vol. 35, no. 1, pp. 50–56, 1997.

[178] S. W. Mitchell, "Some Disorders of Sleep," *The American Journal of the Medical Sciences*, vol. 100, no. 2, pp. 109–127, 1890.

[179] I. A. Kelmanson, "Clinical and physiological aspects of the arousal reaction in young children," *Human Physiology*, vol. 39, no. 6, pp. 625–634, 2013.

[180] M. A. Mograss, F. M. Ducharme, and R. T. Brouillette, "Movement/arousals. Description, classification, and relationship to sleep apnea in children," *American journal of respiratory and critical care medicine*, vol. 150, no. 6, pp. 1690–1696, 1994.

[181] M. J. Drinnan, A. Murray, J. E. S. White, A. J. Smithson, G. J. Gibson, and C. J. Griffiths, "Evaluation of activity-based techniques to identify transient arousal in respiratory sleep disorders," *Journal of sleep research*, vol. 5, no. 3, pp. 173–180, 1996.

[182] M. L. Lamprecht, A. P. Bradley, G. Williams, and P. I. Terrill, "Temporal associations between arousal and body / limb movement in children with suspected obstructed sleep apnoea," *Physiological Measurement*, vol. 37, pp. 115–127, 2016.

[183] M. M. Grigg-Damberger, "The AASM scoring manual four years later," *Journal of Clinical Sleep Medicine*, vol. 8, no. 3, pp. 323–332, 2012.

[184] L. M. Walter, G. M. Nixon, M. J. Davey, D. M. O'Driscoll, J. Trinder, and R. S. C. Horne, "Sleep disturbance in pre-school children with obstructive sleep apnoea syndrome," *Sleep Medicine*, vol. 12, no. 9, pp. 880–886, 2011.

[185] T. C. W. Landgrebe, P. Paclik, R. P. W. Duin, and A. P. Bradley, "Precision-Recall Operating Characteristic (P-ROC) curves in imprecise environments," *Proceedings - International Conference on Pattern Recognition*, vol. 4, no. c, pp. 123–127, 2006.

[186] J. C. T. Pepperell, R. J. O. Davies, and J. R. Stradling, "Sleep studies for sleep apnoea." *Physiological measurement*, vol. 23, pp. R39–R74, 2002.

[187] A. Sadeh, K. M. Sharkey, and M. a. Carskadon, "Activity-based sleep-wake identification: an empirical test of methodological issues." *Sleep*, vol. 17, pp. 201–207, 1994.

Glossary

A

AHI Apnoea hypopnea index (AHI) is a representation of the number of apnoes or hypopneas per hour of sleep.

ANOVA Analysis of variance (ANOVA) is a statistical model used to analyse differences between group means. ANOVA provides a statistical test that determines whether the means of groups of data differ. The ANOVA is a generalisation of the t-test to several variables.

Apnoea Apnoea is where there is a cessation of air intake caused by an obstruction in the airways or weak respiratory muscles. This cessation of air intake causes reduced oxygen saturation in the blood.

E

Electrocardiography Electrocardiography (ECG) measures electrical signal fluctuations on the surface of the skin caused by heart activity. Electrodes are placed on the surface of the skin across the thorax and measure any voltage fluctuations caused by heart activity over a period of time.

Electroencephalography Electroencephalography (EEG) records the electrical voltage potential fluctuations along the scalp caused by brain activity. Multiple electrodes are placed on the skin surface across the scalp and measure the electrical signals caused by neural activity in the brain. These signals contain the frequency information of neural activity during sleep that can be used to determine the different sleep stages in sleep analysis.

Electromyography Electromyography (EMG) measures electrical voltage potential generated by muscle cells during muscular activity. Electrodes are placed on the surface of the skin and measures the electrical potential that is generated when the muscle under the electrode is activated.

Electrooculography Electrooculography (EOG) measures eye movements by recording the electrical voltage potential between the front and rear of the eye. Pairs of electrodes are placed such that the eye is between them (for example, at the top and bottom or left and right of the eye). These electrodes then measure the potential difference between the electrodes. If the eye moves in one direction it will create a positive potential. Conversely, if the eye moves in the other direction it will cause a negative potential. This difference is used to determine the position of the eye.

H

HSD Tukey's honest significant difference (HSD) test is used in conjunction with an ANOVA to determine the means that are significantly different and their statistical significance value.

Hypopnea Hypopnea is where there is a reduction in air intake caused by an obstruction in the airways or weak respiratory muscles. This reduction in air intake causes reduced oxygen saturation in the blood.

Hypoxia Hypoxia refers to an inadequate supply of oxygen to a region of the body.

I

Indicator function The indicator function takes the value 1 when an event $\mathbb{E}$ occurs, and 0 when it does not. It is mathematically defined as:

$$\mathbb{1}(\omega) = \begin{cases} 1, & if \omega \in \mathbb{E} \\ 0, & if \omega \notin \mathbb{E} \end{cases} \tag{G.1}$$

K

Kappa Cohen's Kappa is a measure of the agreement that adjusts for chance. It is defined mathematically as:

$$\kappa = \frac{p_o - p_e}{1 - p_e} \tag{G.2}$$

where p_o is the relative observed agreement, and p_e is the hypothetical probability of chance agreement [132].

M

Myoclonic twitches Myoclonic twitches are very small body twitches (in the order of less than a second in duration) that occur during REM sleep.

N

Non-stationary signal A non-stationary signal has time-varying variance and mean parameters; the joint probability distribution changes when shifted temporally.

P

Polysomnography Polysomnography (PSG) is the current gold standard for diagnosing sleeping disorders. Polysomnography measures various biophysiological signals across the night, including brain activity (see electroencephalography), eye movements (see electrooculography), heart rhythm (see electrocardiography) and muscular activity (see electromyography). These signals require many sensors to be placed on the body and as such, polysomnography is generally intrusive and requires a fully trained nurse in the sleep laboratory to administer.

Precision Precision represents the percentage of positive predictions that are actually correctly predicted,

$$Specificity = \frac{TP}{TP + FP}. \tag{G.3}$$

where TP refers to the number of true-positives, and FP refers to the number of false-positives.

Q

Quiet rest Quiet rest refers to periods of wake where there is no associated movement. This will often occur for patients with sleeping disorders such as insomnia or those causing fragmented sleep. This will also occur at any time during the night when the patient is immobile whilst awake.

R

R&K Scoring Guidelines The R&K scoring guidelines were developed by Allan Rechtschaffen and Anthony Kales in 1986 to aid in scoring sleep stages.

Recall Recall represents number of positives that were correctly classified,

$$Recall = \frac{TP}{TP + FN}. \tag{G.4}$$

where TP refers to the number of true-positives, and FN refers to the number of false-negatives.

RERA Respiratory effort-related arousal (RERA) is defined as an event that cannot be classified as a hypopnea that lasts greater than 10 seconds, is accompanied by a progressively negative oesophageal pressure (caused by increased inspiratory effort) and is terminated by an arousal.

Restless Sleep Restless sleep refers to periods of sleep where there is associated movement. This will often occur for patients with sleeping disorders such as periodic leg movement or REM behaviour disorder. This will also occur at any time during the night when a patient moves whilst asleep, for example, to scratch their nose or change sleeping positions.

S

Sensitivity Sensitivity represents the ability to accurately detect sleep, and is defined as the percentage of sleep epochs or samples correctly scored as 'sleep',

$$Sensitivity = \frac{TP}{TP + FN}. \tag{G.5}$$

where TP refers to the number of true-positives, and FN refers to the number of false-negatives.

Signal to Noise Ratio Signal to noise ratio (SNR) compares the level of a desired signal strength to that of the background noise. It is defined as the ratio of the desired signal power to the noise power.

Sleep apnoea Sleep apnoea is a sleep-related breathing disorder where breathing is partially or completely restricted by either an obstruction in the airways or unregulated respiratory muscle control. An obstruction in the airways is often due to enlarged tonsils or adenoids in children. The restriction of breathing control or intake causes a reduction in oxygen levels in the blood. This leads to poor respiratory function and/or fragmented sleep.

Sleep-disordered breathing Sleep-disordered breathing refers to breathing abnormalities during sleep, such as those related to snoring, sleep apnoea-hypopnea syndrome, etc.

Specificity Specificity represents the ability to accurately detect wake, and is defined as the percentage of wake epochs or samples correctly scored as 'wake',

$$Specificity = \frac{TN}{TN + FP}. \tag{G.6}$$

where TN refers to the number of true-negatives, and FP refers to the number of false-positives.

Spectrogram Visual representation of the frequency spectrum of a signal. The horizontal axis in a spectrogram represents the time scale, and the vertical axis represents the frequency scale.

T

Transient Arousal Transient arousals describe the change in sleep stage (not necessarily to wake) from an internal or external stimulus.

Index