Flinders UNIVERSITY
ADELAIDE · SOUTH AUSTRALIA

The Impact of Sleep Fragmentation on Health

Bastien Lechat

A thesis submitted to Flinders University
for the degree of
Doctor of Philosophy

College of Science and Engineering
Flinders University, Adelaide

August 6, 2021

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

AHI	Apnoea-hypopnoea index
BMI	Body Mass Index
CBTi	Cognitive behavioural therapy for insomnia
CFS	Cleveland family study
COMISA	Co-morbid insomnia and obstructive sleep apnoea
CPAP	Continuous positive airway pressure
CVD	Cardiovascular disease
DIMS	Difficulties initiating and maintaining sleep
DNN	Deep neural network
DWT	Discrete wavelet transform
ECG	Electrocardiography
EEG	Electroencephalography
EMG	Electromyography
EOG	Electrooculography
ESS	Epworth sleepiness scale
HR	Hazard ratio
KC	K-complex
NSRR	National Sleep Research Resources
NREM	Non-rapid eye movement
MASS	Montreal Archive of Sleep Study
OSA	Obstructive sleep apnoea
OR	Odds ratio
ORP	Odds ratio product
PPV	Positive predictive value
PSQI	Pittsburgh sleep quality index
REM	Rapid eye movement
SHHS	Sleep Heart Health Study
SPL	Sound pressure level
TPR	True positive rate
WNSS	Weinstein noise-sensitivity scale
WFN	Wind farm noise

SUMMARY

Inadequate sleep is highly prevalent in the population and leads to an increased risk of a broad range of adverse health impacts. These include traffic and other accidents, mental health problems including depression, psychiatric disorders and suicidal tendencies and cardio-metabolic diseases such as diabetes, hypertension, heart disease and stroke. Assessments of sleep fragmentation and sleep quality are important to evaluate and mitigate risks associated with poor sleep. However, sleep fragmentation has traditionally been defined according to conventional sleep scoring in 30-sec epochs and obvious electroencephalographic (EEG) changes associated with arousals and awakenings, rather than more systematic and potentially more sensitive physiologically guided measurements derived from modern signal processing methods. Current markers of sleep fragmentation repeatedly fail to predict important clinical outcomes, such as sleepiness or cardiovascular events. Therefore, the aim of the work presented in this thesis was to develop new markers of sleep fragmentation based on key features of EEG changes during sleep. These biomarkers were subsequently tested for clinical utility in several population groups relevant to sleep fragmentation, including a sample of individuals exposed to experimental environmental noise manipulations and several large population samples including participants with sleep disorders.

Phasic sleep fragmentation due to experimental environmental noise was quantified using K-complexes, a subtle EEG marker of sensory processing during sleep. K-complexes were automatically detected and scored using a deep learning algorithm that was developed as part of this thesis. The effect of different types of environmental noise (traffic noise and wind farm noise) on sleep fragmentation was assessed in a pilot-study of 21 individuals exposed to a range of noises at different sound pressure levels throughout sleep. K-complexes were a more sensitive sensory disturbance marker of noise exposure during sleep than traditional metrics, such as arousals and awakenings. Statistically significant K-complex responses were observed at sound pressure levels as low as 33 dBA (75% more likely than control) and K-complex response probability further increased with sound pressure level. In contrast, arousals and awakenings were only detectable with noise exposures above 39 dBA. Overall, K-complexes were two times more likely to occur in response to noise than EEG arousals or awakenings, clearly indicating their superior sensitivity to noise exposure compared to traditional arousal scoring.

In a separate study and analysis, deep sleep fragmentation was assessed using a technique conceived during this thesis work, which combines power spectral analysis

of the delta-frequency band (0.5 Hz to 4.5 Hz) with a measure of signal complexity via spectral entropy. The association between deep sleep fragmentation assessed with this new entropy metric and all-cause mortality was studied in the Sleep Heart Health Study (SHHS), a large US-based cohort (N = 5804). Delta sleep fragmentation was associated with a ~30% increased risk of all-cause mortality compared to no sleep fragmentation. This association was similar to a reduction in total sleep time from 6.5h to 4.25h. Conventional measures of sleep quality, including wake after sleep onset and arousal index were not predictive of all-cause mortality.

Hyperarousal – a pathophysiological trait sometimes observed in patients with insomnia, was quantified using the odds ratio product (ORP), a novel marker of sleep alertness. Association between the ORP during wake (hypothesised to reflect hyperarousal) and sleepiness/poor sleep quality was assessed in two large cohort studies (HypnoLaus N = 2162; MAILES N = 754). Hyperarousal was associated with around a 30% increased risk of self-reported poor sleep quality (Pittsburgh Sleep Quality Index score >5) in both HypnoLaus (28%) and MAILES (36%), but an approximately 20% decrease in excessive daytime sleepiness (Epworth sleepiness scale score >10) in the combined dataset. In contrast, no associations were detected using any traditional polysomnography markers.

The additive effect of multiple sleep disorders (co-occurrence of insomnia and obstructive sleep apnoea (COMISA)) on all-cause mortality and sleep fragmentation was studied in the SHHS cohort (N = 5804). COMISA was associated with greater sleep fragmentation and COMISA patients were at higher risk of all-cause mortality (30%) and cardiovascular events (30%). Insomnia-alone and obstructive sleep apnoea (OSA)-alone were not associated with all-cause mortality risk or cardiovascular event risk.

The work presented in this thesis suggests that metrics designed to encapsulate core physiological and pathophysiological processes of sleep, sleep fragmentation and sleep disorders provide more informative markers that may be important predictors of adverse health outcomes. Specifically, disrupted deep sleep and an increased state of hyperarousal were two pathways identified as potentially contributing to all-cause mortality, sleepiness and poor sleep quality. K-complexes were also established to be a more sensitive marker of sensory processing during sleep to environmental noise disturbances than conventional metrics. Together, these findings make an important contribution to understanding the impact of sleep fragmentation on health and provide multiple EEG biomarkers with major potential to substantially improve clinical sleep medicine.

DECLARATION

I certify that this thesis does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text.

Signed Bastien Lechat

Date 06/08/2021

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest thanks to my supervisors, Kristy, Branko and Peter. Your mentoring, belief in my work and unparalleled support has made my degree most enjoyable and insightful. Kristy, thank you for giving me the opportunity of joining the lab in the first place – your dedication to help and guidance to every of your students is most inspiring. I would also like to acknowledge Flinders Foundation for the support provided through the Nick Antic Sleep Research PhD Scholarship.

This work wouldn't have been possible with the support, both academically and emotionally, of colleagues and friends at the Adelaide Institute for Sleep Health. Amid a pandemic, you made me feel at home. I would also like to thank all those who participated in, or were involved in, the studies presented in this thesis. To my friends, in Adelaide and back home, thank you for keeping me sane and for your emotional support.

No words can express how thankful I am to my family. Maman, Papa, Noémie and Thomas, your love, support, encouragement and understanding gave me the strength to finish this thesis. I cannot wait to see you all again.

Finally, my deepest thanks must go to Cassie. Thank you for your love.

PUBLICATIONS

During this PhD candidature, I published and submitted for publication the following work:

Published:

1. **Lechat B**, Hansen K, Catcheside P, Zajamsek B. Beyond K-complex binary scoring during sleep: probabilistic classification using deep learning. *Sleep.* 2020;43(10).
2. Scott H, **Lechat B**, Lovato N, Lack L. Correspondence between physiological and behavioural responses to vibratory stimuli during the sleep onset period: A quantitative electroencephalography analysis. *J Sleep Res.* 2020: e13232.
3. **Lechat B**, Hansen K, Micic G, Decup F, Dunbar C, Liebich T, Catcheside P, Zajamsek B, K-complexes are a sensitive marker of noise–related sensory processing during sleep: A pilot study. *Sleep* 2021
4. Sweetman A, **Lechat B**, Catcheside P, Smith S, Antic NA, Chai-Coetzer C, Douglas J, O'Grady A, Dunn N, Robinson Jan, Paul1 D, McEvoy RD, Lack L. Polysomnographic predictors of treatment response to cognitive behavioural therapy for insomnia in participants with co-morbid insomnia and sleep apnoea: Secondary analysis of a randomized controlled trial *Front. in Psychology.* 2021;12

Under-review:

5. **Lechat B**, Hansen K, Melaku Y.A, Vakulin A, Micic G, Adams R.J, Catcheside P, Zajamsek B, Sleep delta wave activity fragmentation and mortality: re-analysis of the sleep heart health study [Annals ATS]
6. **Lechat B**, Appleton S, Melaku Y.A, Hansen K, McEvoy R.D, Adams R.J, Catcheside P, Lack L, Eckert D.J, Sweetman A. Co-morbid insomnia and obstructive sleep apnoea is associated with all-cause mortality. [European Repsiratory Journal]
7. **Lechat B**, Hirotsu C, Appleton S,Younes M, Adams R.J, Vakulin A, Hansen K, Zajamsek B, Wittert G, Catcheside P, Heinzer R, Eckert D.J, A novel EEG marker predicts perceived sleepiness and poor sleep quality [Sleep]
8. Nguyen DP, Hansen K, **Lechat B**, Catcheside P, Zajamsek B A machine learning approach for detecting wind farm noise amplitude modulation [Applied Acoustics]
9. Hansen K, Nguyen P, Micic G, **Lechat B**, Catcheside P, Zajamsek B, Wind farm noise relationship with annoyance: a longitudinal study [The Journal of the Acoustical Society of America]

CHAPTER 1. LITERATURE REVIEW

The negative consequences of severe sleep restriction are well known. Ultimately, continuous sleep restriction in rats causes weight loss, skin and gastro-intestinal lesions, sepsis and ultimately death after 2-3 weeks (Rechtschaffen et al., 2002) strongly suggesting that sleep is essential for good health. In humans, a lack of sleep has been associated with cardio-metabolic diseases such as diabetes and cardiovascular disease (Gottlieb et al., 2006, Patel et al., 2009). Sleep restriction is also associated with neuro-cognitive impairment, sleepiness and vigilance deficits that likely contribute to an increased risk of motor-vehicle accidents (Lyznicki et al., 1998, Belenky et al., 2003). Sleep deprivation is associated with social withdrawal, psychiatric disorders and an increase in suicide ideation (Bernert et al., 2015, Ben Simon and Walker, 2018, Freeman et al., 2020). Given the importance of sleep for good health, tools to assess sleep fragmentation and sleep quality are an important step towards improvement of sleep medicine.

Poor sleep due to sleep disorders, sleep restriction and/or poor sleep hygiene is a growing public health concern. Sleep disorders are diagnosed using gold-standard polysomnography methods to assess sleep which requires electroencephalography (EEG), electrooculography (EOG) and electromyography (EMG) to classify wake, non-REM and rapid-eye-movement (REM) sleep against international manual sleep scoring criteria (Berry et al., 2012). Despite exponential advances in modern computing, sleep medicine remains based on manual scoring methods originating from the 1960s, when chart recorders necessitated manual sleep study scoring literally page-by-30-sec-page (Rechtschaffen and Kales, 1968). Computerised systems have replaced paper-based sleep recording, but sleep scoring methods remain largely manual and unchanged and thus labour intensive (30-90 min per study). Manual scoring is therefore costly and captures only gross visually discernible EEG features with much poorer time and frequency resolution than is available within sleep recordings. Manual scoring also shows considerable intra- and inter-scorer variability.

The potential for big data, automated methods and artificial intelligence to improve sleep medicine practices is thus well accepted (Redline et al., 2013, Goldstein et al., 2020) and several promising new sleep metrics have recently

been advanced (Younes, 2017, Stephansen et al., 2018). However, these remain early-stage advances with limited evidence to support their utility. The first part of this review therefore describes normal sleep structure and conventional measures of sleep fragmentation (and their limitations). The effect of sleep disorders and external factors on sleep structure is also described. The second part of this review is focused on novel ways of measuring sleep quality and fragmentation. At the end of this chapter, the aims and outline of this thesis are presented.

1.1 Sleep architecture

1.1.1 Normal sleep structure

a. EEG activity during sleep

Sleep is characterised according to classic methods first described in detail in 1968 (Rechtschaffen and Kales, 1968). This method recognises that behavioural responsiveness changes occur simultaneously with characteristic changes in EEG features and classifies sleep and wake into 5 stages based on 30 sec EEG epochs. In the wake stage, brain electrical activity is characterized by low amplitude (10 to 30 µV) and relatively fast and mixed (16 to 25 Hz) frequencies. During quiet relaxed wake, particularly with eyes closed, EEG frequencies begin to reduce and often show strong 10 Hz (alpha waves) activity. The transition from the wake state to stage 1 sleep is characterized by a further decrease in frequency to theta waves (3 - 7 Hz). Classically, sleep was then divided into 4 stages of non-REM sleep and rapid eye-movement (REM) sleep. However, more recent American Academy of Sleep Medicine classification methods classify 3 stages of non-REM (N1-N3) by combining classic stage 3 and 4 into N3 (Iber et al., 2007). The main features of this classification schema are shown in Figure 1-1:

- N1 typically occupies around 5% of total normal sleep and is a light transitional stage of sleep, during which the sleeper is easily awoken and may still respond to verbal cues.
- The N2 stage usually comprises around 50% of total normal sleep, and shows characteristic brief bursts of EEG activity including sleep spindles (0.5 to 1.5 sec bursts of 8 to 12 Hz activity) and KCs (large amplitude -

75 µV - positive wave immediately followed by a negative wave, lasting at least 0.5 sec and at a frequency between <1 Hz).

- N3 or slow wave sleep occupies around 20 to 25% of total normal sleep and consists of large amplitude waves at low frequency thought to reflect synchronous low-level activity of large populations of neurons and to be a key marker of sleep homeostatic mechanisms. The sleeper is most difficult to awaken from N3, so N3 is often referred to as deep sleep (Bersagliere and Achermann, 2010).
- REM occupies around 15 to 20% of total normal sleep, and displays wake-like EEG activity, but profoundly reduced muscle activity and characteristic rapid eye movements (Iber et al., 2007) associated with dream activity.

Sleep is typically composed of multiple cycles lasting between 40 to 90 min and repeating 4 to 6 times over a full night of sleep, as shown in Figure 1-1. Multiple short-time scale features can be observed within the typical sleep stages. For example, Figure 1-1 shows an example of a manually scored wake stage which contains both fast/low voltage EEG, indicative of higher alertness (von Stein and Sarnthein, 2000, Kaminski et al., 2012), and alpha waves, indicative of drowsiness/cortical inhibition (Snyder and Foxe, 2010). Similar findings can be observed for the other sleep stages. Manual scoring of sleep in 30-sec epochs therefore ignores potentially informative short-time scale EEG features.

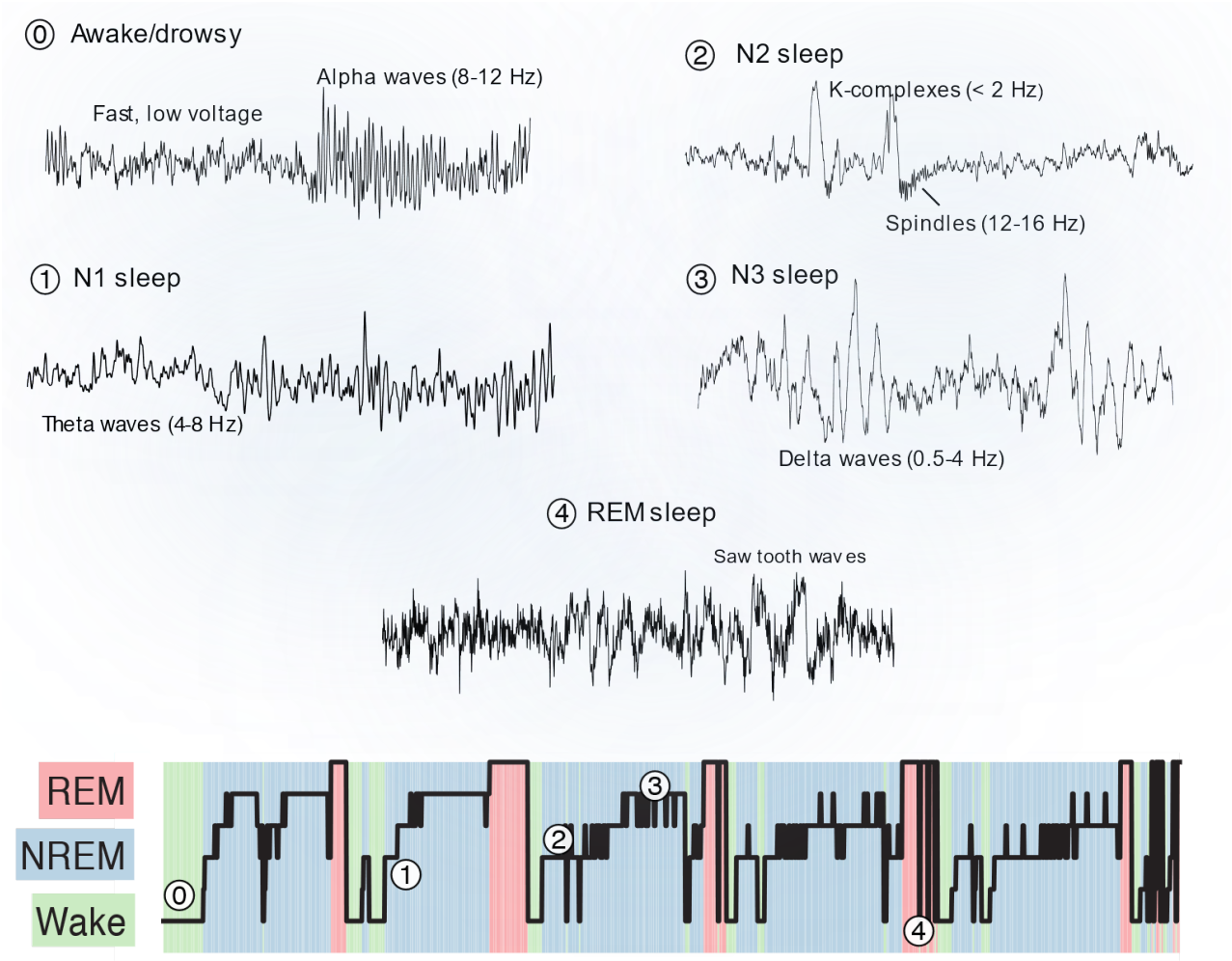

Figure 1-1: EEG activity during sleep.

b. **Slow wave sleep**

Slow wave sleep in humans is defined by the presence of high voltage (> 75 µV) synchronized EEG waveforms. The EEG power within the 0.5 to 4.5 Hz frequency range is typically referred to as slow wave activity and encompasses delta oscillations (1 to 4.5 Hz) and slow oscillations (< 1 Hz). Slow wave activity is typically highest during the first 1-2 sleep cycles and subsequently decreases with the time spent asleep (Figure 1-2). Slow waves seem to be mainly produced locally, meaning that some brain regions can be active in producing slow waves while others brain regions are silent and/or are producing oscillations at different frequencies (Massimini et al., 2004, Nir et al., 2011). Slow waves are produced more and more locally towards the end of sleep periods, and potentially explain the lower amplitude of slow wave activity (Nir et al., 2011) towards the end of the night. The decrease in amplitude might also

be partly explained by circadian phase (Lazar et al., 2015) or decreased sleep pressure and cortical synaptic strength across the night (Esser et al., 2007, Riedner et al., 2007)

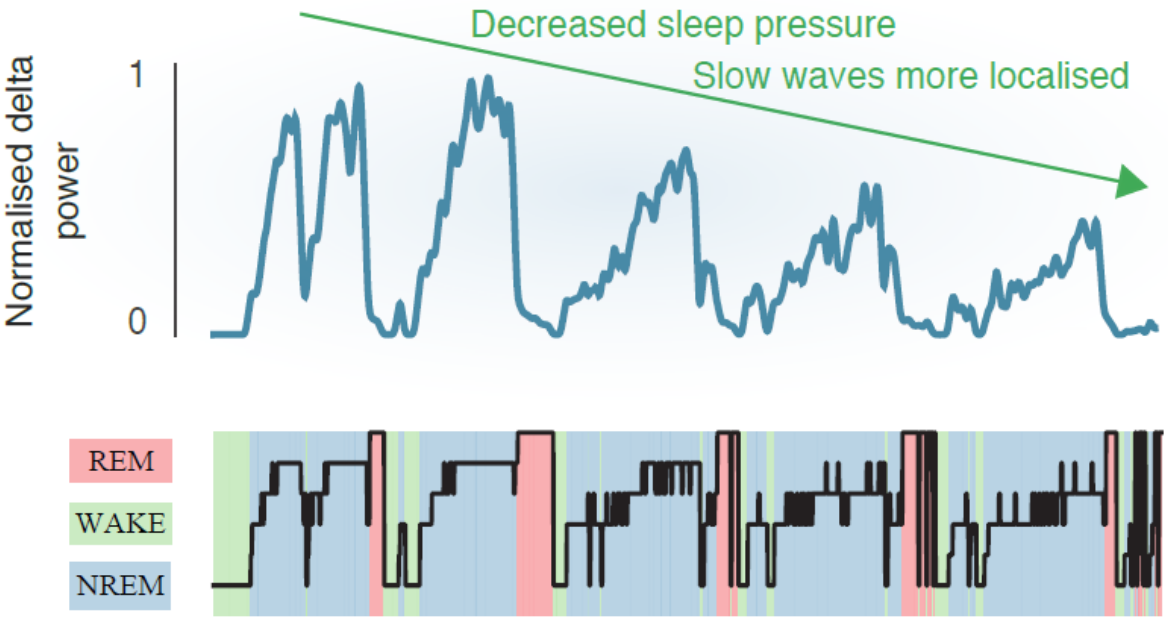

Figure 1-2: Slow wave activity across the night.

Slow oscillations occur when cortical neurons become bistable and undergo a slow oscillation (<1 Hz) in membrane potential (Steriade et al., 1993, Steriade et al., 2001). This consists of a depolarized-up state, when neurons show sustained firing, and a hyperpolarized down state, characterized by neuronal silence, which corresponds to the negative downstroke of EEG slow waves. Slow wave sleep is a good candidate as a physiological marker of sleep homeostasis as it is considered to be a key component of synaptic homeostasis. According to this theory, wakefulness related high levels of neuronal activity potentiate the strength of synaptic connections in the cortex for which slow wave activity during sleep plays a key role in synaptic downscaling (i.e. restore synapses to their baseline strength) needed to re-optimise and re-organise finite synaptic resources underpinning normal brain function and memory and learning (de Vivo et al., 2019); see Tononi and Cirelli (2006) and Tononi and Cirelli (2014) for a in depth review of this hypothesis. The alternating and dissipating pattern of slow wave activity prior to progressively lengthening periods of REM over the course of the night suggests that both slow wave activity and REM

sleep play key roles in synaptic re-organisation during sleep, and that slow wave activity may be a fundamental pre-requisite for the overall process.
The evidence linking impaired slow wave sleep and slow wave activity with adverse health outcomes is overwhelming. Slow oscillations have been hypothesized to be involved in memory consolidation during sleep (Stickgold, 2005, Marshall et al., 2006, Rasch et al., 2007, Maingret et al., 2016) and subsequently, reduced delta power across the night may be associated with cognitive impairment (Taillard et al., 2019). Some evidence supports that slow wave sleep is also involved in systemic metabolic regulation and tissue growth and repair. Indeed, supressed slow wave sleep may adversely affect glucose homeostasis (Tasali et al., 2008) and has been implicated to be involved in the development of Alzheimer's disease (Ju et al., 2017). Finally, slow wave sleep activity has been coupled to cerebrospinal fluid flow, which is associated with clearance of metabolic waste products from the brain (Ju et al., 2017, Fultz et al., 2019). Both experimental and epidemiological studies suggest that slow wave sleep might also be involved in cardiovascular system regulation (Javaheri and Redline, 2012, Silvani and Dampney, 2013, Brindle et al., 2018, Javaheri et al., 2018). Multiple associations between slow wave sleep and a wide range of cardio-metabolic outcomes support the concept that slow wave sleep disruption may contribute to adverse health outcomes. Despite the overwhelming evidence regarding slow wave sleep and good health and its use in sleep research/clinics, slow wave sleep is not used as a criterion of sleep disorder severity according to the international classification of sleep disorders (AASM, 2014, Leger et al., 2018).

c. **K-complexes**

KCs, as shown in Figure 1-3, are bi- or tri- phasic events with components generally accepted as N350, N550 and P900 (where the N and P represent negative and positive peaks, respectively, and the numerical value represents the approximate timing of the peak relative to stimulus onset in ms). KCs are most easily distinguishable and a characteristic feature in N2 sleep but can also occur in other NREM stages of sleep, as a spontaneously occurring event or as an evoked response to a sensory stimulus (Bastien and Campbell, 1992, Bastien and Campbell, 1994). The first studies of evoked KCs showed that

they could be elicited by all types of sensory stimuli but occur more frequently to auditory stimulation (Colrain, 2005).

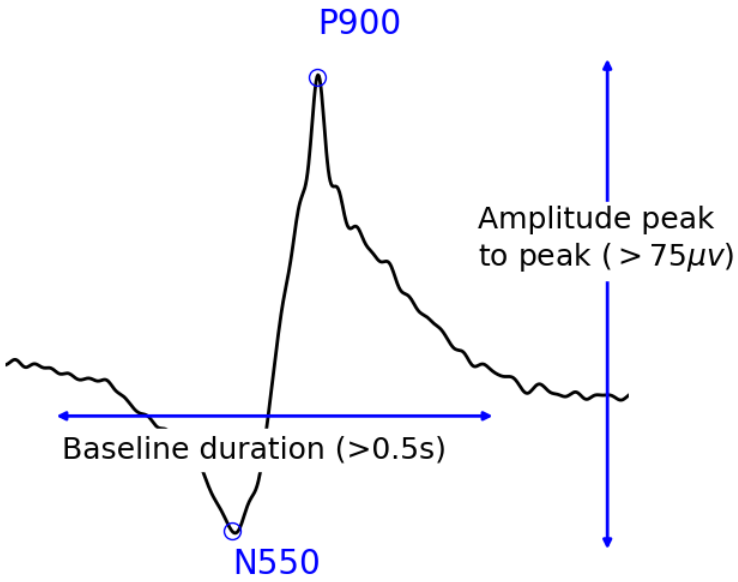

Figure 1-3: A bi-phasic (N550/P900) K-complex.

The physiological significance of KCs has been the subject of a significant debate (Halász, 2005) in which there are 2 main hypotheses. The first hypothesis is that KCs initiate activation of higher cortical centres in the brain (cortical arousal) and augment cardio-respiratory activity. The relationship between KCs and activation of the autonomic nervous system (increase in heart rate and in blood pressure) are well known (Church et al., 1978), and therefore some studies consider KCs as a form of arousal (Monstad and Guilleminault, 1999). The second hypothesis is that KCs and slow waves are similar phenomena related to sleep homeostasis for which KCs may help to preserve sleep in response to potentially sleep-disruptive external stimuli. This was first hypothesized by Amzica and Steriade (2002) following the observation that KCs and slow wave oscillations share the same neuronal sources and fluctuations of membrane potentials (Amzica and Steriade, 1997b, Cash et al., 2009). The likelihood of evoking KCs in response to an acoustic stimulus ranges from 10 to 90% (Bastien and Campbell, 1992, Bastien and Campbell, 1994, Colrain et

al., 1999, Forget et al., 2011). Thus, KCs are clearly a potentially useful marker of noise-induced sensory disturbances during sleep.

While KCs are usually studied in N2 sleep, they are also present in N3 and follow a similar morphological pattern as in N2 (Mak-McCully et al., 2015). Slow waves are formed by a succession of up– and down– states (Steriade et al., 1993) reflecting synchronous firing of cortical neurons and periods of neuronal silence, respectively. Furthermore, KCs can be seen as an up–state–deprived slow oscillation but with the same down–state component (Cash et al., 2009), which makes them difficult to differentiate from EEG slow waves in N3. A significant reason for the slow progression in research surrounding KCs is the current reliance on manual scoring, which is time-consuming given that the KC occurrence rate is around 1 to 2 per minute (Colrain, 2005). Therefore, large-scale studies of KCs are currently impractical and automated methods are needed to upscale research on KCs.

d. **Conventional definition of sleep quality/fragmentation**

Sleep quality and fragmentation are loosely defined terms (Krystal and Edinger, 2008), usually encompassing a broad range of subjective measurements, such as the Pittsburgh sleep quality index (PSQI) (Buysse et al., 1989) and/or sleep diaries, and objective measures of sleep, using gold-standard polysomnography. Among these objective indices are measures such as sleep onset latency, total sleep time, wake time after sleep onset, sleep efficiency, and the number of awakenings or arousals ($>$ 3 sec of fast, wake-like EEG during sleep). Thus in 2017, the National Sleep Foundation formed a committee, including multiple stakeholders and experts, to define sleep quality (Ohayon et al., 2017). The committee recommended sleep onset latency, number of awakenings with duration greater than 5 minutes, wake after sleep onset and sleep efficiency as appropriate measures of sleep quality. More fine-grained metrics of sleep structure, such as slow wave power, were not considered by the committee due to their "limited insight into home setting of the general population" (Ohayon et al., 2017). However, these currently recommended measures are not predictive of important clinical outcomes, such as all-cause mortality (Punjabi et al., 2009, Kendzerska et al., 2014a) or sleepiness (Rosenthal and Dolan, 2008, Adams et al., 2016). This suggests that while these measurements might have

clinical utility, biomarkers which better define sleep quality are needed for better prediction of important clinical outcomes relevant to sleep (Redline et al., 2013).

1.1.2 Sleep fragmentation and sleep disorders

The international classification of sleep disorders consists of seven major categories of sleep disorders as follows: insomnia, sleep-related breathing (OSA, central sleep apnoea, hypoventilation), central disorders of hypersomnolence (such as narcolepsy), circadian rhythm sleep-wake disorders (such as non-24-h sleep-wake rhythm disorder), sleep-related movement disorders (such as restless legs syndrome), parasomnias and other sleep disorders (AASM, 2014). This thesis focuses on the development and clinical testing of biomarkers that may help to quantify sleep disturbance and better inform pathophysiological mechanisms of insomnia and OSA. Given that the prevalence and consequences of these disorders have already been studied in large cohort studies, such as the sleep heart health study (Quan et al., 1997), comparisons between novel metrics and traditional markers of sleep quality are very useful for exploring the potential utility of novel markers of sleep disruption (see section 1.1.1d).

a. Obstructive sleep apnoea (OSA)

OSA is the most common pathological respiratory disorder in sleep. It is estimated that 936 million (95% CI 903–970) and 425 million (399–450) adults aged 30–69 years globally have mild to severe and moderate to severe OSA, respectively (Benjafield et al., 2019). Socio-economic consequences of this are high. The cost of undiagnosed OSA in the US in 2015 totalled nearly $150 billions for reasons such as absenteeism and loss of productivity ($87 billions), but also increased risk of cardio-metabolic a psychiatric conditions ($30 billions) and motor accidents ($26 billions) (Watson, 2016).

The upper airway frequently collapses in patients with sleep apnoea either partly or completely to cause hypopneas (a significant reduction in ventilation) and apnoea (complete cessation of ventilation), respectively. OSA diagnosis and severity are assessed based on the total number of apnoea and hypopnoea events per hour of sleep, called the apnoea hypopnoea index (AHI). The AHI definition and particularly hypopnea scoring criteria have undergone multiple changes in 1999, 2007, 2012, with little regard for impacts of scoring rule

changes on clinical diagnostic cut-offs. This has led to remarkably large differences in estimates of prevalence and consequences of OSA in the general population. For example, in the HypnoLaus cohort, the prevalence of OSA using the originally recommended AHI cut-off > 5 according to the 2007 and 2012 criteria was 35.5% and 60.8%, respectively (Hirotsu et al., 2019). More problematic is that there was a two-fold increase in the threshold associated with cardio-metabolic conditions for the 2007 compared to the 2012 definition.
Severe OSA has been associated with a higher likelihood of developing cancer (Marshall et al., 2014), cardiovascular disease (Redline et al., 2010, Kendzerska et al., 2020) increased motor vehicle accidents (Teran-Santos et al., 1999), decreased cognitive function (Redline et al., 1997, Beebe and Gozal, 2002), reduced quality of life (Appleton et al., 2015), increased depression (Lang et al., 2017a) and ultimately all-cause mortality (Punjabi et al., 2009, Kendzerska et al., 2014a). Associations between adverse health outcomes and OSA have been generally based on variables related to disturbed ventilation and hypoxia, such as (from the most common to the least common) AHI, oxygen desaturation index, and the percentage of sleep time spent with arterial oxygen saturation less than 90%. Furthermore, sleep fragmentation, usually quantified using the arousal index, has largely been discounted as a potential pathophysiological pathway contributing to adverse health outcomes, since several studies have consistently shown no association between arousal index and adverse outcomes (Shahar et al., 2001, Punjabi et al., 2009, Kendzerska et al., 2014a). However, AHI is typically dominated by hypopnoea events for which most scoring criteria require both a period of reduced ventilation and an oxygen desaturation or an arousal (or both). Thus, AHI is a composite outcome for which the potential contribution of arousal to adverse outcomes remains unclear. More problematic is that reducing the impact of hypoxia using continuous positive airway pressure to reduce AHI in patients with pre-existing cardiovascular disease had no effect on the incidence of secondary cardiovascular events (McEvoy et al., 2016). The absence of association between adverse health outcomes and sleep fragmentation in these earlier studies could well be confounded through the use of the arousal index as a marker of sleep fragmentation rather than a true absence of an association. Indeed, recent studies have shown associations between other components of sleep, such as reduced slow wave (Javaheri et al.,

2018) and REM sleep (Leary et al., 2020), and hypertension and all-cause mortality, respectively.

b. **Insomnia**

Insomnia symptoms include difficulties initiating sleep, maintaining sleep or waking too early and difficulties returning to sleep (Kupfer and Reynolds, 1997, Morin and Benca, 2012), from which chronic insomnia is typically defined on the basis of sleep onset latency > 30 minutes and/or wake after sleep onset > 30 minutes at least three times a week for the past 3 months. A more conservative definition of insomnia includes daytime impairment, such as fatigue or feeling unrested. Depending on the definition, insomnia symptoms and/or disorder impact between ~6 to 40% of the general population (Ohayon and Reynolds, 2009, Sweetman et al., 2019, Zhang et al., 2019b). The prevalence of insomnia is also higher in populations reporting other sleep disorders (Morin and Benca, 2012), especially OSA (Sweetman et al., 2019).

The most common accepted questionnaire to diagnose insomnia is the insomnia severity index (Morin et al., 2011). Risk factors for insomnia include low socio-economics status, physical pain, anxiety and stress, shift-work, depression and psychiatric disorders (LeBlanc et al., 2009, Morin and Benca, 2012), although a bi-directional relationship might exist between some of the psychiatric disorders and insomnia or poor sleep (Wulff et al., 2010, Freeman et al., 2020).

Definitive mechanisms underpinning insomnia have not yet been identified. However, hyperarousal is a popular common pathophysiological trait observed in patients with insomnia (Bonnet and Arand, 2010, Riemann et al., 2010). The concept of hyperarousal is characterised by 24-h increased cognitive/emotional (e.g. ruminations about sleep, anxiety, catastrophizing) (Bonnet and Arand, 1997, Harvey, 2002, Bonnet and Arand, 2010), and physiological arousal with increased autonomic and central nervous system activity (Nofzinger et al., 2004, Li et al., 2015). Evidence of increased autonomic activation in patients with insomnia are based on observations of a higher heart rate in patients with insomnia (Stepanski et al., 1994, Bonnet and Arand, 1998) and/or higher levels of cortisol compared to controls without insomnia (Rodenbeck et al., 2002). Absolute power in some EEG frequencies (beta and gamma), likely reflecting higher cognitive processing activity, have also been

shown to be increased in participants with insomnia (Freedman, 1986, Perlis et al., 2001, Hogan et al., 2020) in support of the hyperarousal hypothesis. However, these studies were relatively small in sample size (N ~ 50) and were generally designed to compare insomnia versus control participants. Thus, evidence of physiological arousal in the general population presenting "insomnia-like" symptoms remains unclear and requires more detailed investigation.

c. Co-morbid insomnia and OSA

Insomnia disorder is more common in patients diagnosed with OSA compared to the general population, with a prevalence rate estimated between 30 and 70% (Sweetman et al., 2017a). Similarly, a recent meta-analysis review estimated that 38% of OSA patients meet diagnostic criterial for insomnia (Zhang et al., 2019b). Given the high prevalence of both insomnia and OSA in populations presenting with either primary complaint, Sweetman et al. (2017a), (2017b) suggest that bi-directional relationships may exist between insomnia and OSA, and subsequently coined the term co-morbid insomnia and OSA (COMISA) to describe their overlap. Potential bi-directional relationships have been further reinforced by a recent randomised controlled trial which showed that treating insomnia using cognitive behavioural therapy for insomnia (CBTi) produced a small but significant reduction in OSA severity in COMISA patients (Sweetman et al., 2020).

COMISA patients are at greater risk of adverse health outcomes, such as depression, anxiety and quality of life compared to patients with either insomnia or OSA alone (Lang et al., 2017b, Tasbakan et al., 2018). However, increased cardiovascular/all-cause mortality risk for patients with co-morbid insomnia and OSA compared to any disorder alone has yet to be studied in any detail.

1.1.3 Phasic sleep fragmentation and environmental factors

a. Environmental noise as a public health concern

A report from the World Health Organization (2011) estimated an annual loss of at least one million healthy life years due to annoyance and sleep disturbance caused by environmental noise in western Europe. Children chronically exposed to traffic noise show poorer reading ability, memory and performance

than children who are not exposed to noise at school (Hygge et al., 2002, Lercher et al., 2003). Adults exposed to night-time wind turbine noise are 14% and 17% more likely to seek sleep medication and antidepressants, respectively (Poulsen et al., 2019). Long term effects of chronic noise exposure may also include cardiovascular disease, myocardial infarction or stroke (Babisch, 2011, Basner et al., 2014) and/or chronic disorders such as increased risk of hypertension, diabetes, ischemic heart diseases and atherosclerosis (Babisch, 2011, Basner et al., 2014, Munzel et al., 2018, Zare Sakhvidi et al., 2018).

Since similar health effects are associated with insufficient or disturbed sleep, sleep disturbance has been hypothesised to be an underlying mechanism explaining the adverse health effects of noise exposure (Basner et al., 2014, Basner and McGuire, 2018). This hypothesis is supported by survey data reporting that in an adult population, self-reported shortened sleep duration is associated with nocturnal traffic noise (Evandt et al., 2017).

Evidence of environmental noise-induced sleep disruption is generally based on transportation noise (such as road, rail and aircraft traffic noise). Wind farm noise (WFN) is an atypical environmental noise as it is dominated by low frequencies that propagate long distances and through buildings more readily than higher frequency noise, and is often amplitude modulated (Nguyen et al., 2019) and thus potentially more annoying and disruptive compared to more common environmental noise sources (Schaffer et al., 2016). Wind farm noise is also at its highest sound pressure level (SPL) during night-time hours (Hansen et al., 2019, Nguyen et al., 2019) when other background-noise is usually lowest. Thus, WFN might be particularly problematic for sleep and fundamentally different to transportation noise which is generally lower during the night.

Very few studies have investigated the impact of WFN on sleep (Jalali et al., 2016, Michaud et al., 2016, Smith et al., 2020). Two of these studies did not find any significant effect of noise on sleep. One was a study of pre- and post-WFN operation sleep disturbances in the field, and while no significant subjective or objective sleep impacts were found between pre- and post-WFN operation, the small sample size of three participants and the short analysis period of two days limits the generalisability of the study findings (Jalali et al., 2016). The two other studies were large epidemiological studies that relied

solely on noise prediction models for noise quantification, and actigraphy based signal for estimation of sleep disturbances (Michaud et al., 2016, Michaud et al., 2021). No significant associations between WFN average SPL and sleep disruption were observed (Michaud et al., 2016, Michaud et al., 2021). Variability of WFN SPL across the night was significantly associated with an increased rate of awakenings (and movement time). While significant, the effect size was relatively small and equivalent to only a 1-min increase of wake time overnight for a 5dBA increase in SPL variability. However, noise modelling was done based on ISO 9613-2, which may underestimate true SPL for WFN (Keith et al., 2018). Furthermore, actigraphy cannot adequately measure sleep parameters that may be predictive of noise-induced sleep disturbances (such as arousals).

Only one research group has investigated the effects of WFN on sleep in a laboratory setting using polysomnography(Ageborg Morsing et al., 2018, Smith et al., 2020). The first study was a pilot study with 6 participants and suggested that WFN could potentially increase the number of awakenings and decrease the duration of N2/N3 sleep (Ageborg Morsing et al., 2018). These findings were partially reproduced in a later from the same group study suggested that WFN delayed REM sleep, reduced REM sleep duration and disrupted sleep quality over the course of the night (Smith et al., 2020). The occurrence of noise-induced phasic events (such as arousals and awakenings) was not significantly worse during WFN exposure. Therefore, there is a clear need for studies more specifically designed to investigate potential dose-response relationships between wind farm noise exposure and markers of sleep disturbance (Micic et al., 2018).

b. **Noise-induced sleep disturbance**

Deriving exposure-response curves between sleep disturbance and SPL is of major importance for informing public policy decision making (Basner and McGuire, 2018). Several studies have derived noise exposure responses curves for various types of traffic noise compared to awakenings, arousals and brief autonomic arousal events derived from cardiovascular response markers (Jakovljevic et al., 2006, Basner et al., 2008, Griefahn et al., 2008). Dose-response relationships between awakening rates and daytime function are also relatively well established, and show that an increase in evoked awakenings

results in impaired cognitive function, mood, and alertness, even if the total sleep time remains unchanged compared to a no-noise control night (Bonnet, 1985, Bonnet, 1987, Martin et al., 1997).

Exposure-response functions between environmental noise and sleep disturbance usually show low levels of noise-related sleep disturbance. For example, between 5 to 10% of noise occurrences between 33 dBA and 43 dBA were found to evoke awakenings and arousals (Elmenhorst et al., 2012, Basner and McGuire, 2018). While these findings suggest that responses at lower noise levels may be subtle, relationships between markers of sleep disruption and next-day impacts remain largely unknown. Experimental data showing that overnight noise exposure without any apparent changes in sleep time, arousals or awakenings causes next day sleepiness and mood impairment supports that even subtle noise-related sleep disruption is sufficient to cause negative impacts (Martin et al., 1997). Sub-cortical autonomic responses including heart-rate acceleration (Martin et al., 1997) and peripheral vasoconstriction (Catcheside et al., 2002) are observed with noises that do not necessarily elicit cortical arousals or awakenings and are generally associated with KC co-occurrence (de Zambotti et al., 2016). Thus, the investigation of dose-response relationships between more sensitive markers of sleep disturbance and environmental noise (SPL and types) is an important next step towards understanding the potential long-term effects of chronic noise exposure during sleep.

1.2 Emerging biomarkers in sleep research

1.2.1 Limitations of human scoring

Sleep is traditionally scored based on 30-sec epochs. For example, N3 sleep is characterized by at least 20% of the epoch containing more than 75 µV amplitude delta activity. In its current form, this scoring does not differentiate between '20% delta activity N3 epochs' and '100% delta activity N3 epochs' and therefore substantial changes in delta activity could be masked by current 30-second epoch scoring. More generally, the coarse time-scale of 30-second epochs and practical constraints of manual scoring are very poorly suited to the systematic study of shorter time scale micro EEG events such as arousals, KCs and sleep spindles.

Manual scoring also has large intra- and inter-scorer variability, which remains significantly problematic in sleep medicine despite AASM scoring criteria updates attempting to reduce scoring variability. Magalang et al. (2013) studied the inter-scorer agreement in sleep stage classification between 9 scorers from international sleep centres. Wake, N3 and REM stages showed relatively good agreement with kappa statistics (mean ± SD) of 0.78 ± 0.01, 0.67 ± 0.02 and 0.78 ± 0.01 respectively. However, N1 (0.31 ± 0.01) and N2 (0.60 ± 0.01) showed lower agreement. Furthermore, the mean intra-class correlation coefficient of the arousal index was relatively low (0.68) with high within-scorer variance (± 0.15). Other studies report similar findings (Danker-hopfe et al., 2009, Rosenberg and Van Hout, 2013).

High inter- and intra-scorer variability inevitably leads to approximation in markers of sleep disorders, such as AHI measures and OSA diagnosis (Thomas et al., 2020). In this regard, the use of automated methods, instead of manual scoring, has been recognized as an important step needed towards improvement of sleep medicine (Redline et al., 2013, Younes, 2017).

1.2.2 Deep learning and automated sleep scoring

Deep learning has gained widespread interest in the last decade and has been shown to perform well in computer vision, natural language processing and medical applications (Hinton, 2018). Sleep medicine has also been influenced by the rise of deep learning, and there are now several deep learning approaches to sleep stage classification that can achieve human-level performance (Tsinalis et al., 2016, Supratak et al., 2017, Chambon et al., 2018, Phan et al., 2019). The inputs to those networks are usually raw EEG data, sometimes with EMG or EOG data. These methods are based on convolutional neural networks (CNN). In addition, recurrent neural networks (RNN) based on long short-term memory (LSTM) layers are sometimes used and this algorithm can be trained to learn sleep stage transitions.

The results obtained by these algorithms are similar to inter-scorer agreement (around 80% for all sleep stages, except N1) obtained with human scoring (Rosenberg and Van Hout, 2013, Ruehland et al., 2015) and interestingly both human-algorithm and human-human agreement remains low for sleep N1 (around 50%). All authors discussed potential improvements of their algorithm using training data scored by multiple scorers. While this might be a good

solution to improve algorithm performance against human scoring, these algorithms remain constrained by traditional 30-second epoch scoring against arbitrary EEG classifications founded on traditional sleep pattern recognition from the 1960's. As explained in section 1.1.1a, this approach continues to ignore most of the finer-grained quantifiable features within the EEG highly likely to be more strongly predictive of clinical outcomes than largely arbitrary rule-based coarser time-scale metrics.

There have been several attempts to automate the scoring of more subtle EEG elements using deep learning, such as the identification of KCs/spindles (Chambon et al., 2019) and/or the cyclic alternating pattern (Hartmann and Baumert, 2019). Deep learning has also been applied to other polysomnography signals, such as for the identification of apnoeic events (see Mostafa et al. (2019) for a review). However, the use of deep learning for scoring fine-grained elements in sleep is only emerging and there remains considerable room for improvement in both accuracy and interpretability of the algorithms. Most of the algorithms described above use threshold function to binarize outcomes (0 or 1) and therefore do not quantify uncertainty. Uncertainty quantification is a way of assessing the reliability of automated decisions (Begoli et al., 2019) by giving a "confidence score". For example, uncertainty quantification was used to identify "difficult cases" in an algorithm detecting diabetic retinopathy (Leibig et al., 2017), thus helping to avoid misdiagnosis by referring more equivocal cases to medical experts for further assessment.

1.2.3 Sleep microstructure biomarkers

a. K-complexes

Most studies of KCs have been focused towards establishing exposure-response curves between KC occurrence and stimuli characteristics, such as SPL for acoustic stimuli (see Colrain (2005) for a review). In the context of noise, the likelihood of evoking a KC ranges from 10% (Forget et al., 2011) to 80-90% (Bastien and Campbell, 1992, Colrain et al., 1999, Nicholas et al., 2006, Colrain et al., 2010). Multiple demographic, behavioural and clinical factors can affect the rate of evoked KCs such as age (Colrain et al., 2010) sleep pressure (Nicholas et al., 2002), alcoholism (Colrain et al., 2009), neuropathology

(Crowley et al., 2005) and OSA (Afifi et al., 2003, Nguyen et al., 2016). However, most auditory evoked KC studies used pure tones at high SPL (~80dBA) to maximize KC occurrence, which are not representative of real-world environmental noise.

Previous research has mainly focused on KC occurrence in response to artificially synthesised, high-level tones. However, there is a lack of research on KC occurrence in response to more realistic environmental noise exposure levels. It is therefore unknown whether there is a threshold below which KCs are no longer elicited and whether the noise type has an influence over KC occurrence. Furthermore, it is unknown whether KC response characteristics remain fixed or whether they change depending on noise exposure conditions.

b. **Slow oscillations**

Characteristic features of EEG slow oscillations have been mainly studied using power spectrum analysis, which will be discussed in section 1.2.4. However, a few studies have focused on specific time-based aspects of slow waves. For example, the slope of half slow-waves (Bersagliere and Achermann, 2010) (.i.e. the slope between the up-state and the down-state) has been shown to be sensitive to sleep restriction, suggesting that sleep need or "pressure" to sleep has an effect on the shape of slow oscillations. Furthermore, the proximity of other EEG oscillations (such as spindles) to slow oscillations has been a major area of research interest to study the effect of sleep on memory consolidation. These studies suggest that the proximity, and the phase of coupling between slow oscillations and spindles is involved in memory formation (Hahn et al., 2020) and consolidation (Helfrich et al., 2019, Muehlroth et al., 2019).

However, these metrics do not take into account the particular distribution of slow wave oscillations across the night (see section 1.1.1b) and they consider slow oscillations at the beginning and end of the night as the same "entity". Assuming that the distribution of slow oscillations likely reflects multiple physiological processes, then a marker encapsulating the more complex dynamics of slow wave distribution across sleep could provide a particularly useful physiological marker of sleep quality with significant practical and clinical value.

c. Cyclic alternating pattern (CAP)

The cyclic alternating pattern is an additional sleep scoring system beyond traditional AASM sleep scoring which aims to quantify sleep NREM discontinuity by characterizing phases of activation (A phases) and periods of inactivity (B phases) (Terzano et al., 2001). Subgroups of A phases (1 to 3) represent different degrees of cortical activation, with A1 phases mainly composed of slow oscillation activity, and A3 phases containing more high-frequency EEG, supposedly representing a higher degree of cortical activation (Parrino et al., 2012). Automatic methods of CAP scoring have been proposed (Hartmann and Baumert, 2019) and have been applied to study and define NREM instability in large population-based studies (Hartmann et al., 2020b, Hartmann et al., 2020a). However, cyclic alternating pattern scoring has its own issues with inter-scorer agreement with a Cohen's kappa coefficient of around 0.6 (Ferri et al., 2005). Furthermore, as this technique groups micro-elements together (Parrino et al., 2012), cyclic alternating pattern scoring ignores potentially useful information of single events, for example, KCs or spindles alone.

1.2.4 Power spectral analysis of EEG signals

Power spectral analysis has been applied to categorize physiological hyperarousal in people suffering from insomnia through identification of higher absolute power in the beta frequency band (Krystal et al., 2002). People suffering from insomnia may experience smaller amounts of slow wave activity (measured by the absolute power of delta activity) (Cervena et al., 2004, Krystal and Edinger, 2010). However, a recent meta-analysis suggests that only relative delta activity (defined as absolute delta power divided by absolute power from 0.5 Hz to 35 Hz) is associated with insomnia (Zhao et al., 2021).

Power spectral analysis has also been used to assess EEG power pre- to post-treatment of insomnia using CBTi. Collectively, these results suggest that cognitive behavioural therapy for insomnia may increase the amount of absolute power in the delta band (Cervena et al., 2004) during NREM, but also increase the amount of delta in the first NREM cycle (Krystal and Edinger, 2010).

Krystal and Edinger (2010) also studied metrics such as slow wave power only in the first sleep cycle, and the decrease of slow wave power across sleep cycles. These metrics have been used to study the effect of an early evening nap on

sleep EEG power (Werth et al., 1996) in one of the earliest clinical studies on the topic. The same methodology was recently used by the same research group to show a slower decrease of EEG delta power in individuals with insomnia versus controls (Lunsford-Avery et al., 2021).

Similar techniques have also been applied to OSA, where the mean absolute power of a given frequency band (delta, alpha, theta, sigma, beta), usually averaged over NREM and REM sleep, was shown to be predictive of disease severity (Appleton et al., 2019). The same methodology has also been used to predict daytime impact of OSA, such as excessive daytime sleepiness and cognitive functioning (see (D'Rozario et al., 2017a) for a review).

However, there are several important limitations with spectral power analyses employed in previous studies. Firstly, spectral analysis is typically only applied following traditional sleep stage scoring and therefore importantly governs which sleep epochs are analysed. As discussed in section 1.2.1, this inevitably introduces problems with inter-scorer agreement which have the potential to mask more subtle EEG changes than can be detected through manual methods. Averaging over traditionally scored sleep stages (30-sec) is convenient and likely helps to reduce effects of manual scoring, but remains highly likely to mask more subtle sleep and environmental noise dependent changes over both short (<30 sec) and longer-time scales (minutes or hours). Clearly EEG spectral features can and do change both within and between sleep stages, so averaging across NREM and/or REM remains inherently problematic for detecting more subtle changes in the EEG than is possible with conventional manual scoring. Finally, power spectral density analysis has generally been applied to relatively small datasets, and more epidemiological studies (such as the one performed by Djonlagic et al. (2021)) accounting for confounders, are needed to systematically evaluate the potential clinical utility of power spectral analysis in sleep medicine.

1.2.5 **Time-frequency analysis of EEG signals**

Sleep EEG is naturally suited for frequency and time-frequency analysis, since different stages or micro-elements (such as spindles, KCs, slow waves) have specific frequencies (Steriade, 2006, Scammell et al., 2017). Power spectral analysis of EEG may thus offer a more sensitive and objective marker for identifying patient phenotypes in different sleep disorders than current manual

methods. Prerau et al. (2017) revisited an older method, called multi-taper analysis (Park et al., 1987, Babadi and Brown, 2014), to systematically quantify EEG spectral power across the full range of EEG frequencies with much better resolution than manual scoring. This approach is similar to traditional fast Fourier transformation based methods (D'Rozario et al., 2017b) but has superior noise reduction and feature extraction capabilities. This is an important consideration for clinical EEG acquisition where signal quality is highly variable. Figure 1-4 shows an example of multi-taper analysis where a transition from slow wave sleep (1) to N2 sleep (3) with an arousal in the middle (2) is observed. Slow wave sleep is characterised by high absolute power at frequencies less than 4 Hz and very little power at high frequencies, thus making the identification of high frequency (8-16 Hz) arousals straightforward. The transition from arousal to N2 sleep is also very specific, with a reduction in high frequency power, a sparse low frequency burst (likely reflecting KCs), sometimes followed by a burst of energy around 12 to 16 Hz, likely representing spindles.

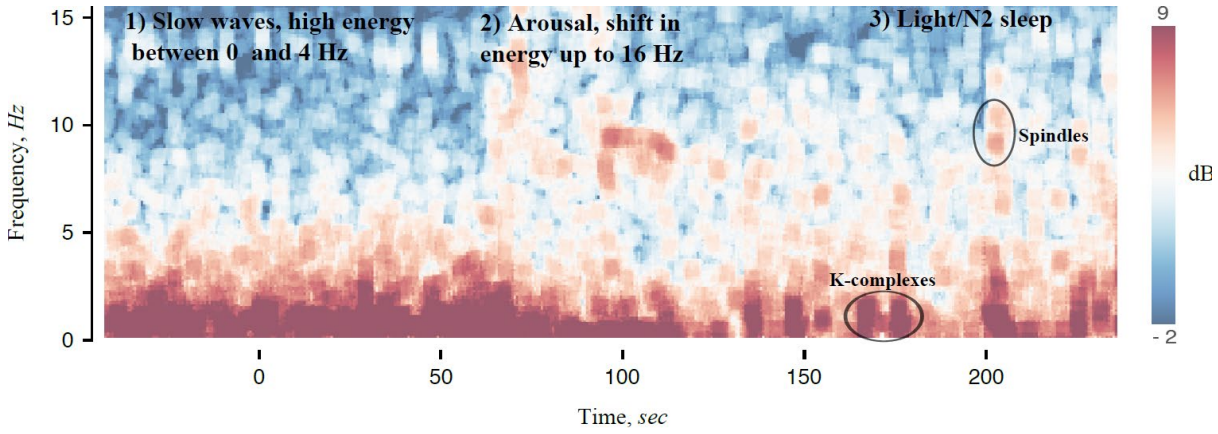

Figure 1-4: Spectrogram of sleep EEG signals using multi-taper based method.

Multi-taper-based methods have not yet been used as a means to quantify sleep processes in a clinical population such as OSA, insomnia or COMISA. Therefore, although likely, it is currently unknown if markers based on multi-taper analysis could have clinical utility.

1.2.6 The odds ratio product (ORP)

The odds ratio product (ORP) is a novel EEG-derived metric that provides a continuous index of sleep depth and alertness (Younes et al., 2015, Younes,

2017). ORP is calculated in 3-sec non-overlapping segments, and is based on the absolute power, calculated using the fast Fourier transform, of 4 frequency bands: 0.33 - 2.33 Hz (slow delta), 2.67 - 6.33 Hz (range 2, includes theta and fast delta), 7.0 - 14.0 Hz (alpha/sigma) and 14.0 - 35.0 Hz (beta). The power in each of the four frequency ranges is calculated and assigned a rank (0–9) depending on pre-selected cut-off values, which were determined based on 58 polysomnography studies included in the original ORP study (Younes et al., 2015). These ranks are then used to determine the probability of being awake in each 3-second segment. This probability, which ranges from 0 to 100%, is then divided by 40 resulting in an ORP ranging from 0 to 2.5, where 0 indicates very deep sleep and 2.5 is widely awake. ORP values correlate well with the visual appearance of EEG across the night, as shown in Figure 1-5. There is an excellent correlation ($r^2 = 0.98$) between average ORP in 30-sec epochs and the arousability index (defined as sum of subsequent arousals and awakenings * 100 / total number of epochs) of the following epochs, suggesting that ORP might be a reliable marker of arousability (Younes et al., 2015, Younes et al., 2020).

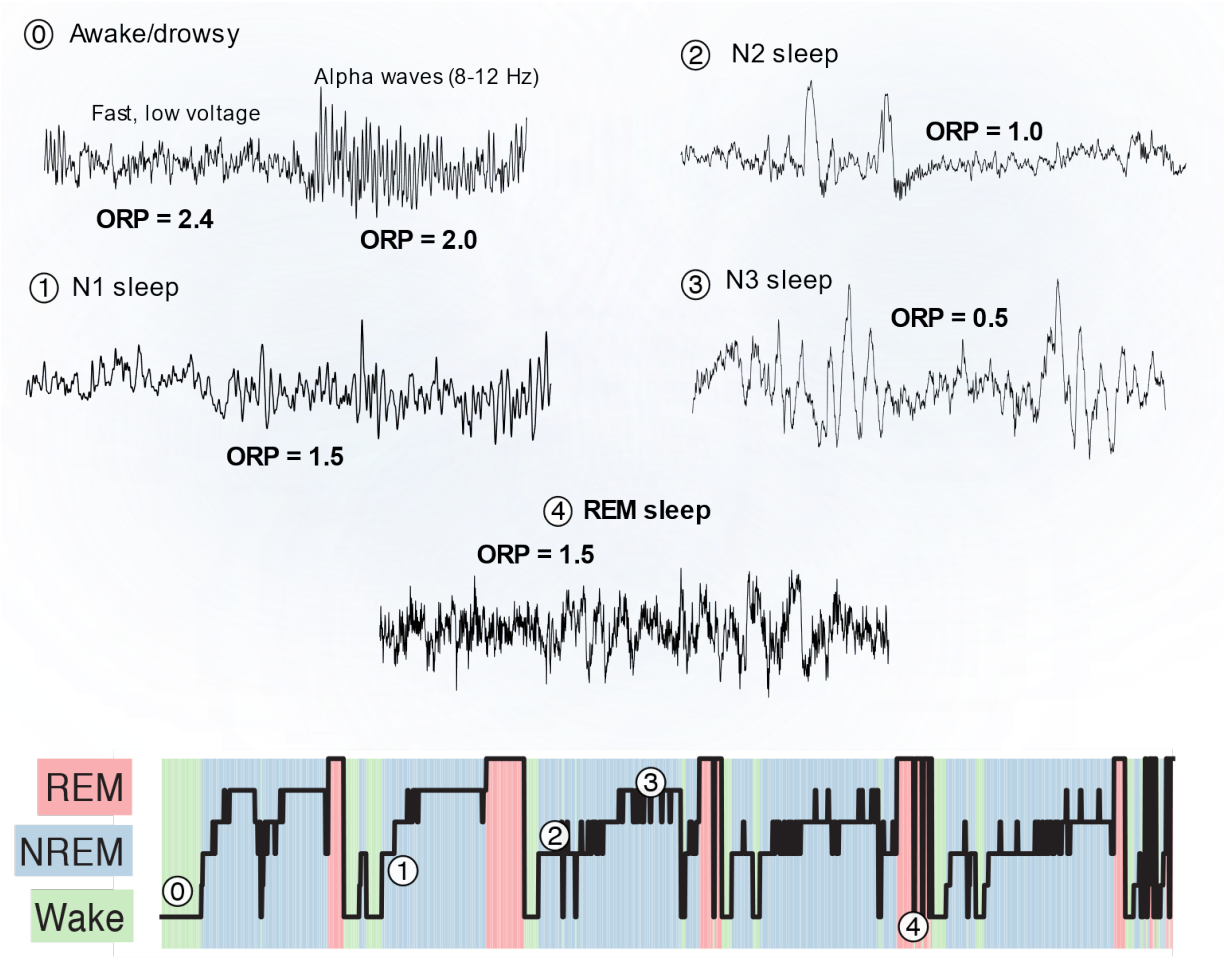

Figure 1-5: The odd ratio product, a marker of sleep depth and alertness.

ORP derived metrics may be useful for a wide range of clinical applications. For example, abnormally low ORP values during the wake period have been observed in mechanically ventilated patients (Dres et al., 2019). These participants with an abnormally low wake ORP were more likely to fail at a spontaneous breathing test, which tests a patient's capacity to breath without the need of a ventilator. Together these findings suggest that ORP may be useful for guiding decisions around ventilator weaning. High post-arousal ORP values have also been associated with lower sleep continuity in patients with sleep disordered breathing suggesting that the ORP may be a non-invasive marker of respiratory-related arousal threshold, which otherwise requires more invasive measurement to determine (Younes and Hanly, 2016). Other potential clinical applications of ORP-derived metrics include the diagnosis of sleep disorders (Younes and Giannouli, 2020). A higher mean ORP may also be a useful marker of decreased sleep depth/quality following environmental noises (Smith

et al., 2019). While promising, the performance of ORP-based metrics for predicting adverse health outcomes requires more studies, including cross-sectional and longitudinal epidemiological studies with large sample sizes necessary to account for potential confounders, and ultimately randomised controlled treatment trials to more definitively demonstrate clinical utility.

1.2.7 Measuring signal complexity via entropy

Surface EEG signals reflect the sum of electrical potentials radiating from neuronal and non-neuronal sources with different electrical potentials, action potential firing rates and distances from the recording electrodes. EEG is highly nonlinear and traditional signal processing techniques, such as EEG spectral-band power analysis, is likely to miss some potentially informative underlying signal features (Bradley and Kantz, 2015). Nonlinear time-series analysis (Bradley and Kantz, 2015, Zou et al., 2019) is a set of tools very commonly used in fields such as weather prediction (Goswami et al., 2018), cardiology (Kumar et al., 2017), finance (Zhou et al., 2013) and epilepsy detection (Kannathal et al., 2005). Multiple entropies have been used to quantify the degree of signal complexity. For example, decreased EEG complexity was measured in patients with Alzheimer's disease (Abasolo et al., 2006). In sleep research, entropy has been generally used as a feature to classify sleep stages (Ma et al., 2018). Two studies have used entropies, calculated over different sleep stages, as a potential biomarker of Parkinson's disease (Chung et al., 2013) and neurodevelopment in newborns (Zhang et al., 2009). However, these studies were limited by relatively small sample sizes and entropy-based metrics have not been examined as potential markers of sleep disorder consequences to date. Given promising results to date with these techniques, nonlinear time series analysis of sleep signals clearly warrants further research.

1.2.8 Signal coupling, network physiology and machine learning-based approaches

While sleep quality metrics are not the specific focus of this thesis, and a detailed review is available elsewhere (Mendonca et al., 2019, Lim et al., 2020), a few particularly influential metrics warrant mention.

a. **Signal coupling and network physiology**

Several research groups have investigated the coupling between multiple physiological signals, such as heart rate in combination with respiratory signals (named cardio-pulmonary coupling) (Thomas et al., 2005, Bartsch et al., 2012, Penzel et al., 2016, Thomas et al., 2018). Coupling-based analyses have also been applied between sleep EEG and heart rate (Brandenberger et al., 2001) to facilitate the study of interactions between the central and autonomic nervous system activity. The theoretical concept of coupling-functions between different physiological systems has been recently generalised under the framework of network physiology, which aims to study relationships between different types of signals (Bashan et al., 2012, Ivanov et al., 2016).

b. **Machine-learning based approaches**

Machine-learning has been used to automatically detect sleep disorders such as narcolepsy (Stephansen et al., 2018), and sleep apnoea subtypes (Mazzotti et al., 2019). These techniques can derive and explore a large number of features from polysomnography signals and use machine-learning to infer sleep disorders (Stephansen et al., 2018). Classification accuracy of these methods partly depends on the features used and their quality and the nature of the outcomes against which their performance is optimised and assessed. Given the data and feature rich nature of sleep EEG there is potential high value in applying machine-learning methods to explore and define more data-driven physiological based biomarkers of sleep than is possible through traditional manual sleep scoring methods. However, while machine learning methods are promising, more work is needed to assess where modern machine learning methods have advantages over traditional approaches (Christodoulou et al., 2019).

1.3 Summary

Sleep is vital for good cardio-metabolic and mental health. The practice of sleep medicine remains heavily dependent on techniques established in the sixties based on manual scoring. Given poor intra- and inter-scorer agreement of manual scoring of EEG events during sleep, it is quite likely that associations between sleep quality and adverse health outcomes may be systemically underestimated.

While there are multiple plausible biological mechanisms linking slow wave sleep and good health, such as synaptic homeostasis and glymphatic system removal of wake-accumulated metabolites, there has been very little epidemiological research studying the impact of fragmented slow wave sleep on health. Quantifying the impact of fragmented slow wave sleep, using a combination of power spectral analysis and measures of signal complexity, is likely to provide informative markers of poor sleep.

Previous evidence supports that KCs are a sensitive marker of sensory processing during sleep. However, systematic manual scoring of KCs is impractical so little is known regarding their potential utility to assess sleep disturbance to low level environmental noise. Given community complaints regarding environmental noise, as well as the potential impact of environmental noise on sleep, testing for associations between KCs and environmental noise is clearly an important step towards understanding potential noise-related sleep disturbance effects.

Insomnia patients sometimes exhibit a pathophysiological trait called hyper-arousal, which is conceptualised as a chronic state of increased cognitive and physiological arousal. Previous studies suggest that greater power in high frequency EEG during sleep is associated with heightened physiological arousal. However, the available evidence is limited and based on a small number of participants. The ORP is likely to be a good marker of hyper-arousal, given that it is a marker of sleep alertness, but independent validation of the ORP as a potentially useful clinical measure has not been done to date.

Finally, emerging research suggests increased morbidity in patients with co-morbid insomnia and OSA compared to insomnia and OSA alone, but this hypothesis remains to be more rigorously tested using large datasets.

1.4 Aims

The central aims of this thesis were as follows:

1. **To determine the exposure-response curve of KC occurrence for environmental noise.**
   An automated KC algorithm was developed and validated and then used to determine the probability of evoked KC occurrence in response to traffic noise and wind farm noise in carefully controlled laboratory experiments. Evoked KC probability was compared to traditional sleep disruption markers of arousals and awakenings. *This work tested the hypothesis that KCs are a more sensitive marker of sensory processing of environmental noise during sleep than traditional markers of arousal from sleep.*
2. **To design novel sleep quality markers predictive of all-cause mortality.**
   A marker of slow wave sleep overnight structure was developed through the lens of delta (0.5 to 4.5 Hz) activity patterns overnight, calculated using a Fourier-based method and spectral entropy. The association between this marker of sleep quality and all-cause mortality was then determined in a US-based large cohort study and subsequently compared to traditional markers of objective sleep quality. *This work tested the hypothesis that the shape of the distribution of delta activity overnight is a stronger predictor of all-cause mortality than current traditional sleep markers.*
3. **To establish the association between sleepiness, poor sleep quality and novel EEG-markers of sleep alertness/depth.**
   The association between high alertness during wake periods, as measured with the ORP, and sleepiness/poor sleep quality were determined and cross-validated in two independent large study cohorts. *This work tested the hypothesis that ORP-based metrics predict self-reported daytime sleepiness and poor sleep quality better than traditional objective sleep markers.*

4. **To determine whether people with co-occurring sleep disorders are at greater risk of morbidity and mortality compare to single disorder.**

   The consequences of co-occurring sleep apnoea and insomnia on all-cause mortality risk and cardiovascular event risk were examined and compared to either disorder alone. *This work tested the hypothesis that patients with co-occurring disorders are at greater risk of morbidity and mortality than patients with either disorder alone.*

1.5 Thesis outline

To address the aims outlined above, the thesis is structured as follows:

Chapter 1 describes the current literature on the assessment, impact, and physiology of sleep fragmentation and the motivation and aims of this thesis.

Chapter 2 details the development of a KC detection algorithm using deep learning and probabilistic classification. The algorithm was validated using a dataset with manually defined KCs.

Chapter 3 examines the impact of environmental noise on sleep. Specifically, noise exposure dose-response curves were constructed to compare the occurrence of KCs, arousals, and awakenings between different types of environmental noise at realistic night-time sound pressure levels.

Chapter 4 outlines a novel quantitative way of measuring sleep quality using delta power, calculated using a Fourier-based method and spectral entropy. This new biomarker was then used to test for associations between sleep quality and all-cause mortality in a large cohort study.

Chapter 5 examines the association between sleepiness, poor sleep quality and sleep alertness/depth in two large cohorts using the odds ratio product, which is a novel EEG-marker of sleep alertness/depth.

Chapter 6 describes the co-occurrence of insomnia and obstructive sleep apnoea (OSA) in a large cohort study, and its association with sleep fragmentation and all-cause mortality.

Chapter 7 discusses the contribution of this thesis to the sleep research field, highlighting the strengths and limitations of this work, as well as possible future research directions.

CHAPTER 2. BEYOND K-COMPLEX BINARY SCORING DURING SLEEP: PROBABILISTIC CLASSIFICATION USING DEEP LEARNING

This chapter was published in the following journal article.

Citation: Lechat B, Hansen K, Catcheside P, Zajamsek B. Beyond K-complex binary scoring during sleep: probabilistic classification using deep learning. *Sleep*. 2020;43(10). Doi: https://doi.org/10.1093/sleep/zsaa077

Author Contributions: study design—BL, BZ.; analysis of data— BL; interpretation of findings— BL, KH, PC and BZ; preparation of initial manuscript— BL; revision of final manuscript— BL, KH, PC and BZ.

Summary

Background: K-complexes are a recognized EEG marker of sensory-processing and a defining feature of stage 2 sleep. K-complex frequency and morphology may also be reflective of sleep quality, aging and a range of sleep and sensory processing deficits. However, manual scoring of K-complexes is impractical, time-consuming and thus costly and currently not well-standardized. Although automated K-complex detection methods have been developed, performance and uptake remain limited.

Methods: The proposed algorithm is based on a deep neural network and Gaussian process, which gives the input waveform a probability of being a K-complex ranging from 0 to 100%. The algorithm was trained on half a million synthetic K-complexes derived from manually scored sleep stage 2 K-complexes from the Montreal archive of sleep study containing 19 healthy young participants. Algorithm performance was subsequently assessed on 700 independent recordings from the Cleveland Family Study using sleep stage 2 and 3 data.

Results: The developed algorithm showed an F1 score (a measure of binary classification accuracy) of 0.78 and thus outperforms currently available K-complex scoring algorithms with F1 = 0.2–0.6. The probabilistic approach also captured expected variability in KC shape and amplitude within individuals and across age groups.

Conclusions: An automated probabilistic KC classification is well suited and effective for systematic KC detection for a more in-depth exploration of potential relationships between KCs during sleep and clinical outcomes such as health impacts and daytime symptomatology.

2.1 Introduction

KCs (see section 1.1.1c) are conventionally scored manually based on rather vague and non-specific scoring guidelines defining a KC as a "well-delineated negative sharp wave, of at least 0.5 sec length, that stands out of the EEG background, usually followed by a positive phase" (Iber et al., 2007). Similar to other manually scored EEG features, this leads to a relatively low inter-rater agreement of around 60–70% (Devuyst et al., 2010). However, more systematic and reliable KC scoring could potentially be clinically useful given that multiple clinical conditions such as OSA (Nguyen et al., 2016, Parekh et al., 2019), insomnia (Forget et al., 2011), Alzheimer's disease (De Gennaro et al., 2017) and restless leg syndrome during sleep (Montplaisir et al., 1996) impact on KC occurrence and characteristics. Consequently, KCs could potentially be a clinically useful biomarker of sleep problems and daytime excessive sleepiness. However, previous studies have been limited by a small sample size given that large-scale scoring of KCs is time consuming and hence costly due to KC's high occurrence rate.

Some KC detection algorithms are based on quantifying human–defined features such as peak-to-peak amplitude, latency and duration of KCs in the time domain (Bankman et al., 1992, Devuyst et al., 2010, Erdamar et al., 2012), while others are based on EEG power in the frequency domain within the 0–4 Hz frequency band (Richard and Lengelle, 1998, Parekh et al., 2015).These algorithms have a clean-cut binary accept/reject output based on arbitrarily pre-defined threshold values in both the time and frequency domains. However, defining clean-cut threshold values is challenging due to waveform variability and noise present in EEG recordings, which has a marked effect on threshold values and corresponding algorithm performance. Furthermore, KC morphology differs greatly between individuals according to age (Crowley et al., 2002), genetics (Gorgoni et al., 2019) and sleep stage (Amzica and Steriade, 1997b, Massimini et al., 2004), compounding the difficulty for selecting optimal threshold values for clinically diverse populations.

More recently, a deep learning inspired algorithm was developed for scoring KCs based on the sleep stage 2 EEG time series with a scorer–algorithm agreement of 60% (Chambon et al., 2019). However, the algorithm is based exclusively on the open–source Montreal archive of sleep study (MASS) database

containing scored KCs from 19 young participants (O'Reilly et al., 2014). KCs were scored by a single scorer, potentially limiting generalizability. Probabilistic scoring, through Bayesian statistics (Leibig et al., 2017), can quantify uncertainties related to the training data selection and labelling accuracy (Begoli et al., 2019). These issues are particularly pressing in sleep research, where databases are usually small (O'Reilly et al., 2014) and inter-scorer agreement is low (Devuyst et al., 2010). Thus, uncertainty-informed decisions are more appropriate than "clean-cut" decisions and are likely to improve diagnostic performance. Only a few studies have attempted to manually score KCs in sleep stage 3 and consequently very little attention has been given to automated KC detection in that stage (Devuyst et al., 2010, Chambon et al., 2019). Despite the absence of manually annotated KCs in stage 3 sleep, automated scoring in that stage is warranted.

The main aim of this Chapter was to develop and validate a well-performing detection algorithm based on probabilistic classification better suited to more widespread uptake and use in sleep research and medicine than current approaches. The output of this approach comprises information on the overall quality of an EEG recording as well as KC morphology differences within and between individuals. The decision-processes underpinning the algorithm were also examined by comparing original and manually modified input effects on the output probability of the algorithm.

2.2 Methodology

2.2.1 Recordings

The algorithm was developed on an open access database MASS with 19 sleep polysomnography recordings from healthy individuals (8 males and 11 females) within an age range 18–33 and mean ($\pm$ SEM) age of 23.6 $\pm$ 3.7 years (O'Reilly et al., 2014). Data were recorded using a Grass Model 12 system with 19 EEG channels (C3, C4, Cz, F3, F4, F7, F8, O1, O2, P3, P4, Pz, T3, T4, T5, T6, Fp1, Fp2, Fpz) referenced to linked-ear electrodes (M1 and M2). The EEG time series were filtered with high– and low–pass first order filters with cut–offs of 0.30 Hz and 100 Hz, respectively. Sleep stages were scored by a single scorer according to Rechtschaffen and Kales sleep scoring rules (Kales and

Rechtschaffen, 1968). KCs were scored on the C3 channel using the American Academy of Sleep Medicine (AASM) manual (Iber et al., 2007).
Further assessment was performed by applying the algorithm to the Cleveland Family Study (CFS) dataset (Redline et al., 1999) from the National Sleep Research Resources (Dean et al., 2016) which contains 735 PSG recordings (406 males and 329 females), from participants aged between 6 and 88 years old. Only polysomnography recordings with at least 75% of artefact-free EEG, as recorded by the human-expert scorer, were included in this analysis. Only the C3 channel referenced to linked-ear electrodes (M1 and M2) was kept for analysis, as the algorithm was developed on C3. Any PSG recordings with a fallen C3 or reference EEG lead were excluded. Raw EEG data were recorded at a 256 Hz and 128 Hz for MASS and CFS dataset, respectively. Furthermore, MASS data were resampled offline at a sampling frequency of 128 Hz.

2.2.2 Algorithm workflow

The algorithm was designed to detect KCs during stage 2 sleep on a C3 channel using 3 steps involving data pre-processing, balancing and classification using deep learning (Figure 2-1).
The first pre-processing of EEG segments step, selects peaks greater than $\alpha = 15\ \mu$V as potential KCs and other peaks as non-KCs (nKCs) (Figure 2-1A). A data segment of $d = 3$ seconds is retained before and after each selected peak. The parameters α and d are arbitrarily defined and deliberately well below the KC features criteria of peak-to-peak amplitude $\geq$75 μV (Bastien and Campbell, 1992) and 0.5 second duration (Iber et al., 2007). An amount, N, of $2 * d$ long KCs and nKCs samples is then decomposed using the discrete wavelet transform (DWT) with a symlet 3 wavelet (Mallat, 1989). Wavelet decomposition returns 128 coefficients c for a DL network input and they represent energy within the 0-1, 1-2, 2-4, 4-8 and 8-16 Hz bands.
The second step balances KCs and nKCs wavelet coefficients due to highly imbalanced EEG recordings with many fewer KCs than non–KC waveforms given that ~95% of EEG data is KC free (Figure 2-1B). The MASS dataset, for example, was found to contain 160,000 nKCs and 7,535 KCs. DL requires large and balanced datasets for training (Goodfellow et al., 2016) and hence KCs were synthesized instead of randomly selecting equal sized KC and nKC dataset, which would be small due to the small number of KCs (Roy et al.,

2019). Balancing is achieved either through creation of new KCs by adding Gaussian noise (μ=0 and σ=0.4) to existing KCs (Wang et al., 2014) or as follows (Zhang et al., 2018):

$$KC_{synt} = KC_1 * \phi + KC_2 * (1 - \phi) \quad \text{Eq 3.1}$$

where the coefficient ϕ is drawn from a beta distribution (Johnson et al., 1995) ($\alpha = 20$ and $\beta = 1$, with α and β parameters controlling the shape of the distribution, analogous to mean and SD descriptions of the shape of a normal distribution), and KC_1 and KC_2 are randomly selected existing KCs. The balanced dataset was used only during the training phase containing approximately 500,000 KCs and 500,000 nKCs.

The last step operates the classification of EEG segments as KC or nKC using deep kernel learning (DKL) (Wilson et al., 2015, Wilson et al., 2016) which is a combination of deep neural network (DNN) and Gaussian processes (Rasmussen, 2006). The DNN consists of 5 layers with a max-norm constraint, including fully–connected linear layers followed by batch normalization (Ioffe and Szegedy, 2015), rectified linear unit (Nair and Hinton, 2010) and drop out layer (Srivastava et al., 2014). The number of units per layer is 1000-1000-500-256-b, where b is manually adjusted. A Gaussian process is fit on each dimension of b with a radial basis function (RBF) kernel. Because of the non-Gaussian likelihood, Gaussian process training was approximated using stochastic variational inference (SVI) (Wilson et al., 2016). SVI is an approximation that leverages inducing point methods (Snelson and Ghahramani, 2006) via performing training on an $M = 1000$ data sub-set. The model was optimised using stochastic gradient descent (SGD) with Nesterov momentum.

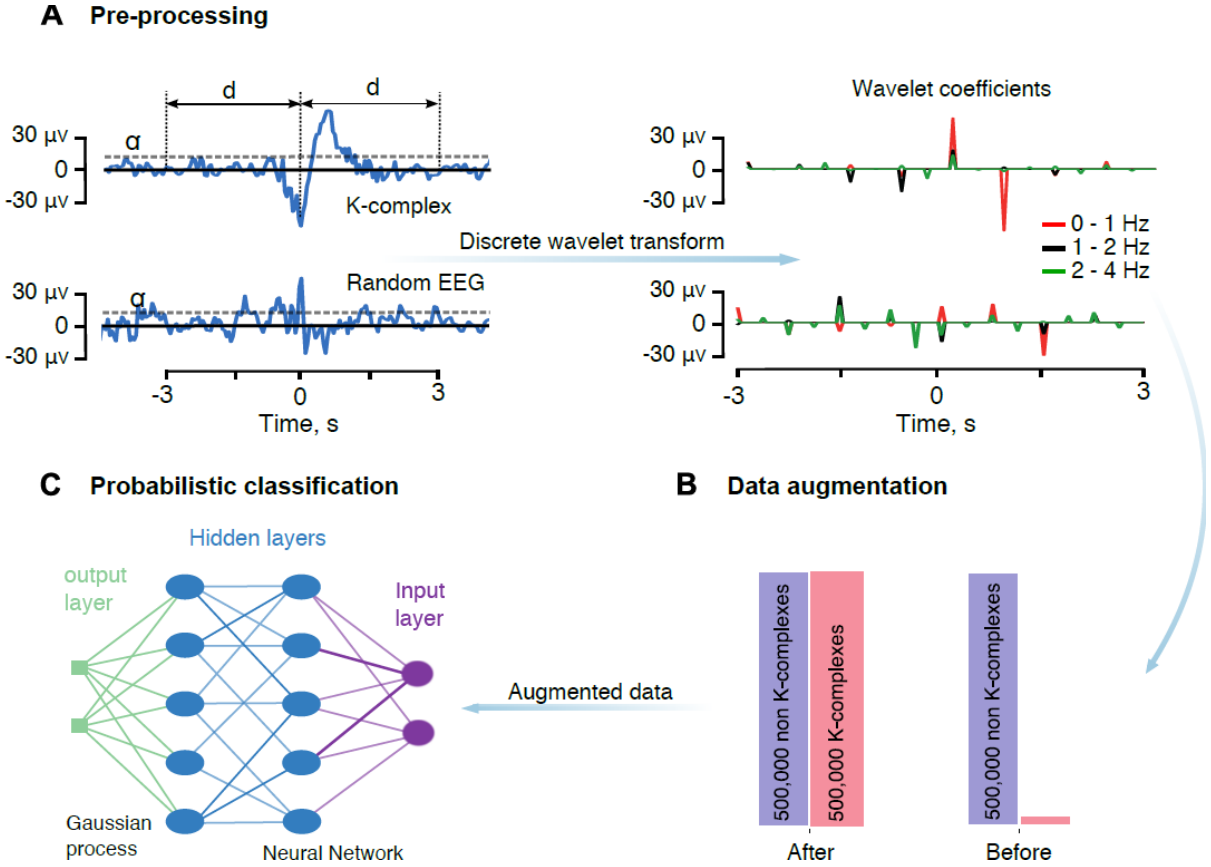

Figure 2-1: K-complex detection algorithm workflow. A) The pre-processing step using a discrete wavelet transform B) dataset balancing and C) classification using deep learning and Gaussian processes. Reproduced by permission of Oxford University Press https://doi.org/10.1093/sleep/zsaa077.

The algorithm was developed in Python, using MNE-python (Gramfort et al., 2013) for EEG processing, PyWavelets (Lee et al., 2019) for the DWT, and Pytorch (Paszke et al., 2019) and GPytorch (Gardner et al., 2019) for the classification model. The algorithm is available under a common license rule at https://github.com/Adelaide-Institute-for-Sleep-Health/K-complex_algorithm.

2.2.3 Hyper-parameter tuning and training

Hyper parameters in deep learning are parameters that cannot be learned from the data, and therefore require tuning. A five–split cross validation (Kohavi, 1995) was used to adjust the following hyper-parameters: learning rate, b, drop–out rate and momentum. The splits were organized with 10, 5 and 4

participants for training, validation and testing respectively. The most efficient combination was determined by the best average performance metrics on the validation set, and this combination was then tested using the testing set.
The most important hyper-parameters, presented in Table 2-1, were tuned using grid-search on all 144 possible combinations of hyper-parameters. DKL is computationally expensive to train and therefore only a subset of hyper–parameters and their values can be systematically tuned while the remaining parameters are adjusted manually. The best performance was achieved with a learning rate of 0.1, $b = 16$, a drop out rate of 0.7 and a momentum of 0.95.

Table 2-1: The most important hyper-parameters of the algorithm and their selected range for fine-tuning using cross-validation.

Hyper-parameter	Learning rate	b[1]	Drop-out rate	Momentum
Value	[0.5,0.1,0.05,0.01]	[32,16,8,4]	[0.5,0.7,0.9]	[0.9,0.95,0.99]

[1]Bottleneck size of the neural network (input of the Gaussian process)

2.2.4 Model evaluation

Models were evaluated using the true positive rate (TPR), $= \frac{TP}{TP+FN}$, the positive predictive value (PPV), $PPV = \frac{TP}{TP+FP}$, and the F1 score calculated as follows:

$$F1 = 2\frac{TPR \times PPV}{TPR + PPV} \qquad \text{Eq 3.2}$$

where TP is the number of positively scored (human) and detected (algorithm) KCs, FN is the number of positively scored but negatively predicted KCs and FP is the number of negatively scored but positively predicted KCs. The F1 score ranges between 0 and 1 where the higher the score, the better the algorithm performance. In practical terms, the F1 score is the harmonic mean between the PPV and the TPR.

2.2.5 Further evaluation of the algorithm

The size of KCs varies with age, reaching its maximum in adolescence and then steadily decaying with age (Crowley et al., 2002, Crowley et al., 2004). The

difference in median scored probability of KCs as a function of age using the CFS dataset (N = 599) was therefore studied, expecting that older individuals would show a smaller median KC probability and smaller KC peak-to-peak amplitude (Crowley et al., 2002).

As an indication of algorithm robustness to artefacts, KCs were automatically scored in all sleep stages in the CFS dataset, and mean KCs densities were compared across sleep stages. Assuming that KCs are not present in sleep stage 1 or REM sleep, a mean KCs density near 0 is expected for these sleep stages.

To differentiate slow waves from KCs, an up–state and down–state (N550) ratio was computed on all scored KCs during sleep stage 2 and 3. Assuming that down–states and up–states have the same amplitude (Mak-McCully et al., 2015), the ratio between up– and down– states for slow waves is expected to be around 100%, while the ratio for KCs is expected to be smaller. The ratio between the N550 component and the up-state peak which was defined as the maxima between 350 and 800 ms before the N550 component (Mak-McCully et al., 2015) was therefore calculated.

To further validate the algorithm performance in slow wave sleep, 500 KCs and 500 non-KCs, as scored by the algorithm, were randomly selected from slow wave sleep periods of participants in the CFS dataset and manually reviewed. Five researchers, experienced with sleep scoring, from the Adelaide Institute for Sleep Health participated in the review process and were prompt to classify a waveform as either being a KC, as defined by the AASM manual, or a random EEG segment. Scorers were blinded to the algorithm output. Inter-scorer agreement and algorithm/scorer agreement were calculated using the F1 score. Scoring "consensus" was defined as KCs scored by at least three scorers out of 5.

2.3 Results

2.3.1 Overall performance

The proposed algorithm with F1 = 0.78 outperforms published algorithms (Lajnef et al., 2015, Chambon et al., 2019) by a substantial margin (Table 2-2), where an EEG segment was classified as a KC with a probability $\geq$ 50%.

Table 2-2: Mean and standard deviation of the performance metric of the new algorithm (testing set, N = 4) compared with existing algorithm.

Algorithm	F1	± F1	TPR	± TPR	PPV	± PPV
Chambon et al. (2019)	0.60	*n.a.*	0.58	*n.a.*	0.64	***n.a***
Lajnef et al. (2015)	0.2	*n.a.*	0.19	*n.a.*	0.21	***n.a***
New	0.78	0.03	0.86	0.02	0.72	0.06

n.a indicates not available. F1, F1-score; TPR, true positive rate; PPV, positive predictive value.

The probability cut-off threshold can be increased in order to increase specificity (i.e. the PPV) and the resulting algorithm performance is shown in Table 2-3. With an increasing probability threshold, the PPV also increases, while the TPR decreases and the F1-score remains constant. Choosing a threshold is therefore a matter of how specific the algorithm needs to be with respect to the scorer, which does not necessarily improve the algorithm overall performance.

Table 2-3: Mean and standard deviation of the performance metric (testing set, N=4), as a function of several probability threshold cut-offs.

Probability threshold	F1	± F1	TPR	± TPR	PPV	± PPV
50	0.78	0.03	0.86	0.03	0.72	0.06
60	0.78	0.02	0.83	0.04	0.74	0.06
70	0.78	0.02	0.80	0.05	0.77	0.05
80	0.78	0.02	0.75	0.05	0.80	0.04

F1, F1-score; TPR, true positive rate; PPV, positive predictive value.

2.3.2 Probabilistic assessment

The interpretation of algorithm output is intuitive due to the probabilistic approach as shown in Figure 2-2. In this figure, KCs are ordered in increasing scored probability from top to bottom where the red color represents a positive voltage peak around 80 μV (P900) and the blue color indicates a negative voltage peak (N550) around -80 μV. The time values on the x-axis are aligned ($t = 0$ s) with the N550 peak. The KC scoring probability is clearly related to the magnitudes of the characteristic N550 and P900 peaks (Figure 2-2A and C). This means that smaller KCs, or KCs in the presence of significant noise,

will be scored with a lower probability than larger and well-defined KCs (Figure 2-2C).

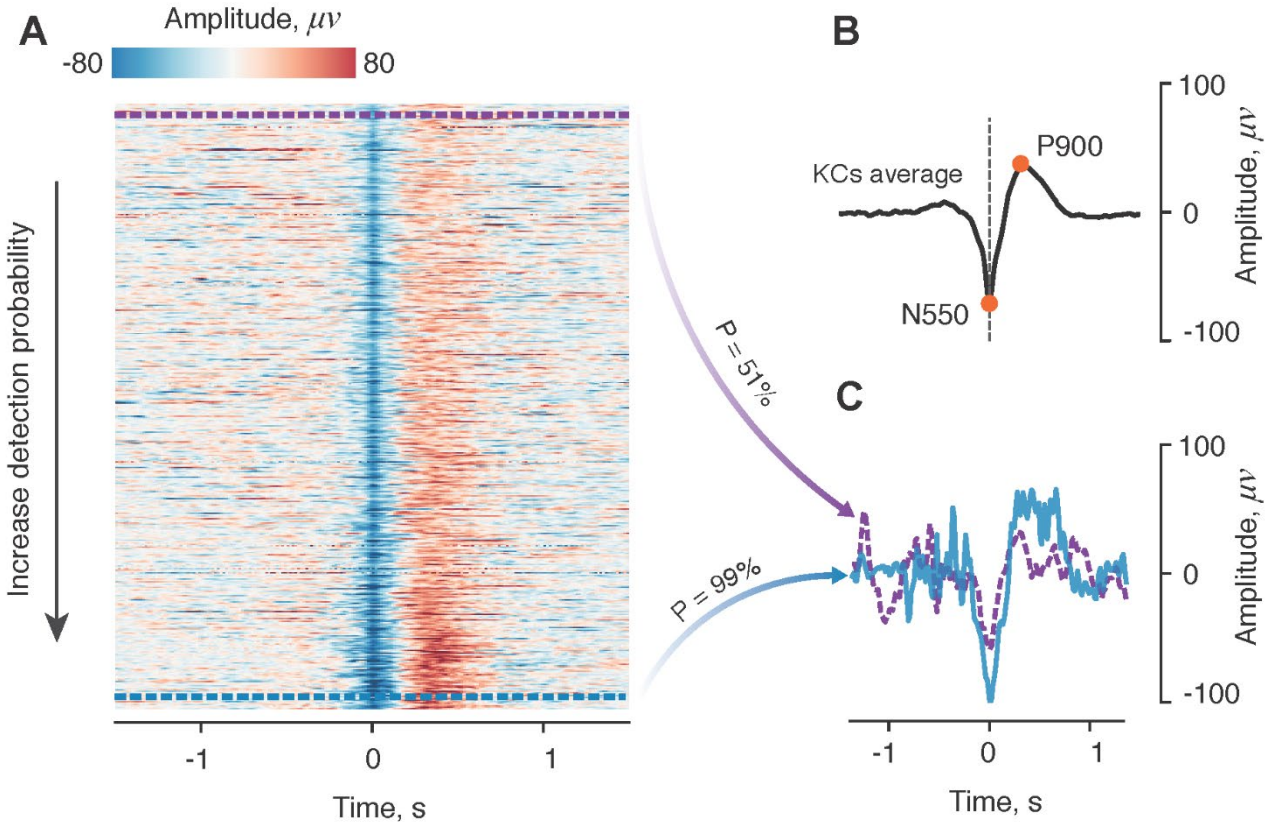

Figure 2-2: An example of varying K-complex probabilities of one participant from the MASS dataset during stage 2 sleep. A) KC probability map where each line represents a KC ordered from highest to lowest probability (bottom to top). B) Grand averaged KC. C) Typical low and high probability KC waveform. Reproduced by permission of Oxford University Press https://doi.org/10.1093/sleep/zsaa077.

2.3.3 Input perturbation

Deep neural networks are generally difficult to interpret. However, by manually modifying the input of the algorithm, it is possible to gain insights into the decision-making processes, as shown in Table 2-4. Table 2-4 shows differences in automatic scoring performance between the full model (with all DWT coefficients) and altered models with certain DWT coefficients removed. As an example, if coefficients between 2 and 4 Hz are removed, the F1-score drops from 0.82 to 0.56, which shows that this DWT frequency band is important in the algorithm decision making process. Furthermore, the mean probability of

scored KCs overnight drops by 48%, signifying the low algorithm confidence on scored KCs, resulting in 640 fewer KCs scored compared to the full model. The absence of individual DWT frequency band coefficients from the DNN input can have a marked effect on overall algorithm performance. With respect to the baseline F1 score obtained with all frequency bands present, the 1 to 8 Hz frequency bands are the most important for KC scoring since their absence drastically worsens the algorithm performance. The slow oscillations captured by the 0 to 1 Hz frequency band only slightly change the classification performance while the 8 to 16 Hz frequency band contains information useful for KC rejection since 160 more KCs were falsely identified when that frequency band was removed from the input into DNN compared to the full model.

Table 2-4: DNN input frequency bands perturbation effect on algorithm performance.

		Frequency bands				
Alteration	None	0–1 Hz	1–2 Hz	2–4 Hz	4–8 Hz	8–16 Hz
F1	0.82	0.81	0.69	0.56	0.77	0.79
Mean probability difference[1]	0	0%	-21%	-48%	-9%	+3%
± KCs[2]	0	+7	-184	-640	-141	+160

[1]A negative difference indicates that the algorithm is less confident if the given frequency band is removed. [2]Quantifies how many more/less k-complexes are scored if a given frequency band is removed.

2.3.4 KC probability and age

The algorithm appears to be robust to noise, with only a few KCs detected in REM sleep and stage 1 sleep; Median (IQR); 0.19 (0.08, 0.35) and 0.17 (0.08, 0.31) KC/min, respectively. For comparison, slow wave sleep and sleep stage 2 showed KC frequencies of 3.9 (2.6, 5.4) and 1.0 (0.7, 1.5) KC/min, which is in line with previously reported KC frequency of occurrence (Halasz et al., 2014). Changes in KC size and probability with age derived from application of the algorithm to the CFS data set is shown in Figure 2-3. The population aged between 18 and 33 years showed a median probability (Figure 2-3A) of 92.5% which then drops at a rate of (Mean (95% CI)); 1.2 (0.99, 1.35)% per 10 years (linear regression controlled for sex and BMI, $p<0.001$) along with an expected reduction in KC peak-to-peak amplitude with age (Figure 2-3B). Indeed, in a

linear regression model controlled for sex and BMI, the peak-to-peak amplitude of the averaged KC overnight dropped by (Mean (95% CI)) 8.5 (7.6, 9.5) μV per 10 years. Therefore, it is likely that the observed reduction in median probability in older individuals is due to changes in the KC shape, since the algorithm was trained on young and healthy individuals. The difference between the median probabilities could thus be interpreted as a measure of probability uncertainty between algorithm development and algorithm application. Although the algorithm was developed on a population between 18 and 33 years old, good classification performance, with an expected reduction in KC size and probability, appears to be retained even in older population datasets, likely attributable to a combination of algorithm design and a large and diverse training dataset.

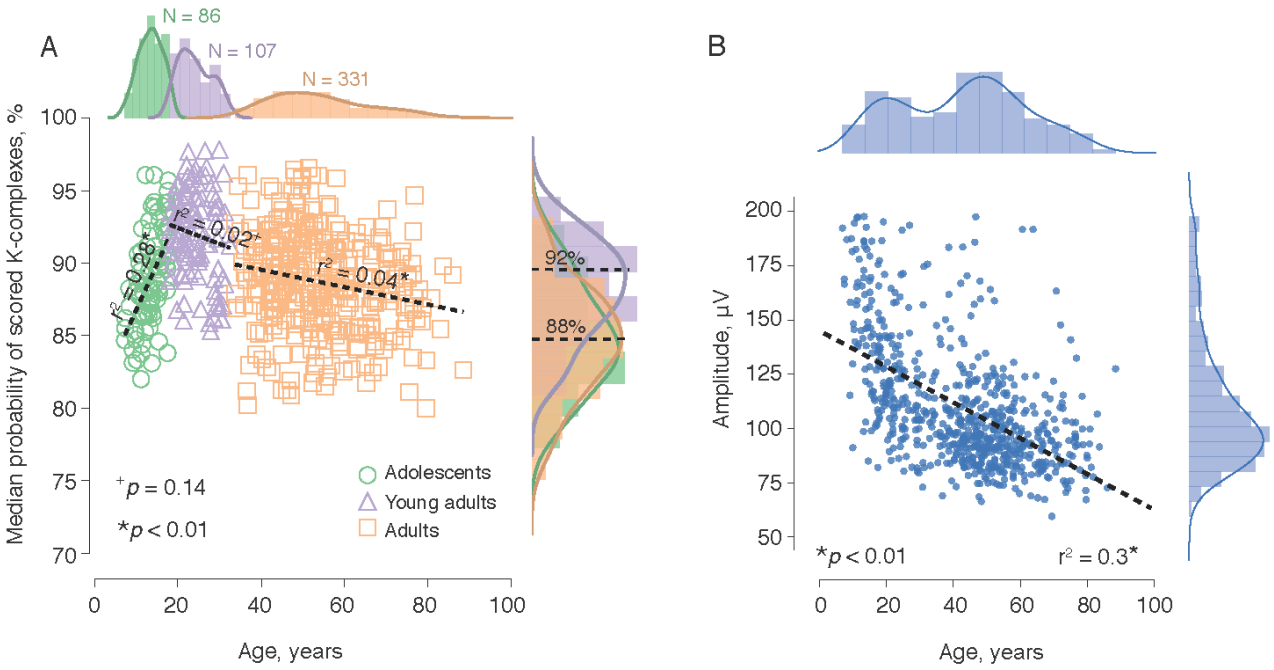

Figure 2-3: (A) Variation in automatically scored K-complex median probability across age on the CFS dataset. (B) Variation of the peak-to-peak amplitude of the grand average KC, per participant, across age on the CFS dataset. Reproduced by permission of Oxford University Press [https://doi.org/10.1093/sleep/zsaa077.](https://doi.org/10.1093/sleep/zsaa077)

2.3.5 KC scoring in sleep stage 3

The scoring of KCs in sleep stage 3 was remarkably good (Figure 2-4) as indicated by the absence of a large positive up–state peak before the N550 peaks, which increases the confidence that the scored waveforms are true KCs and not delta waves.

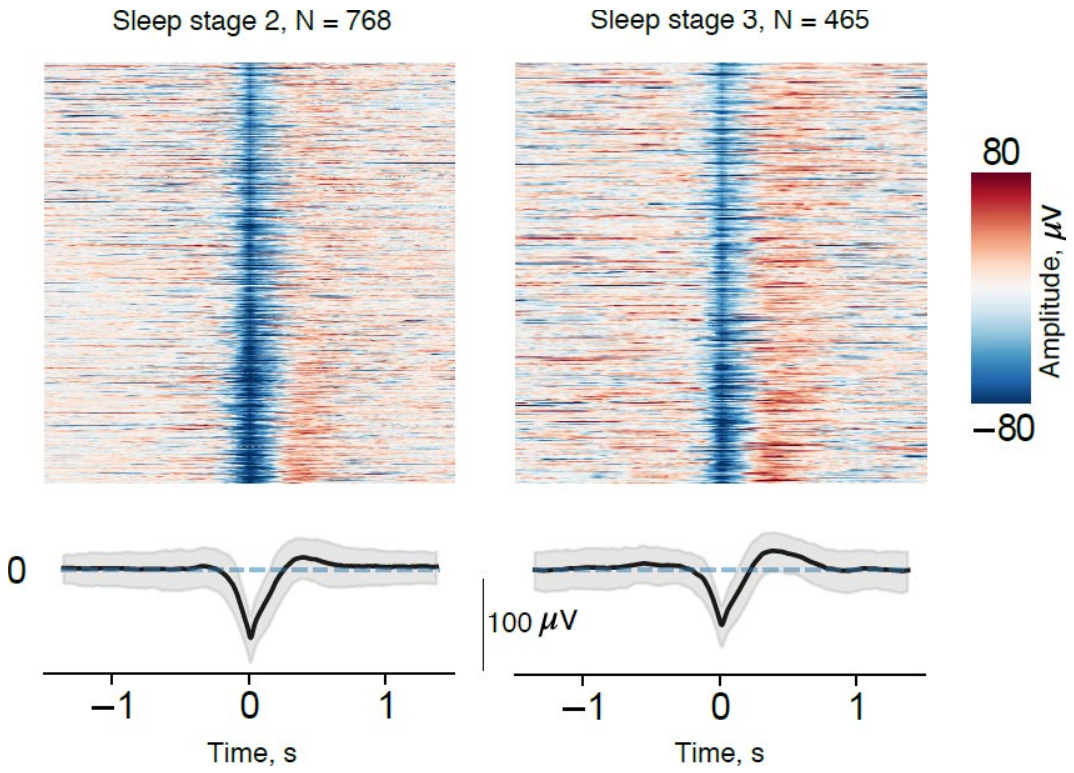

Figure 2-4: Comparison of automatically scored K-complexes in sleep stage 2 and 3 with a KC map (top), where each line represents a KC, and the grand average KC and its standard deviation (bottom). Reproduced by permission of Oxford University Press https://doi.org/10.1093/sleep/zsaa077.

The ratio was calculated for all KCs within one participant and then averaged per sleep stage. Figure 2-5 also shows that the mean ratios from sleep stage 3 (mean ± SD; 49 ± 10%) and sleep stage 2 (32 ± 9%) are similar and well below 100%, as expected for a slow wave ratio, supporting that the majority of scored waveforms are likely to be KCs.

Using an arbitrary limit of 80% as an indication of a slow wave, 10 ± 7% of KCs from sleep stage 2 and 20 ± 8.5% of KCs from sleep stage 3 were rejected as they were more likely to be a slow wave. Although the algorithm was developed exclusively on N2 sleep data, it also seems to perform remarkably well on sleep stage 3 data, where scoring of KCs is inherently more difficult in the

presence of strong slow wave oscillations with similar amplitude and frequency characteristics.

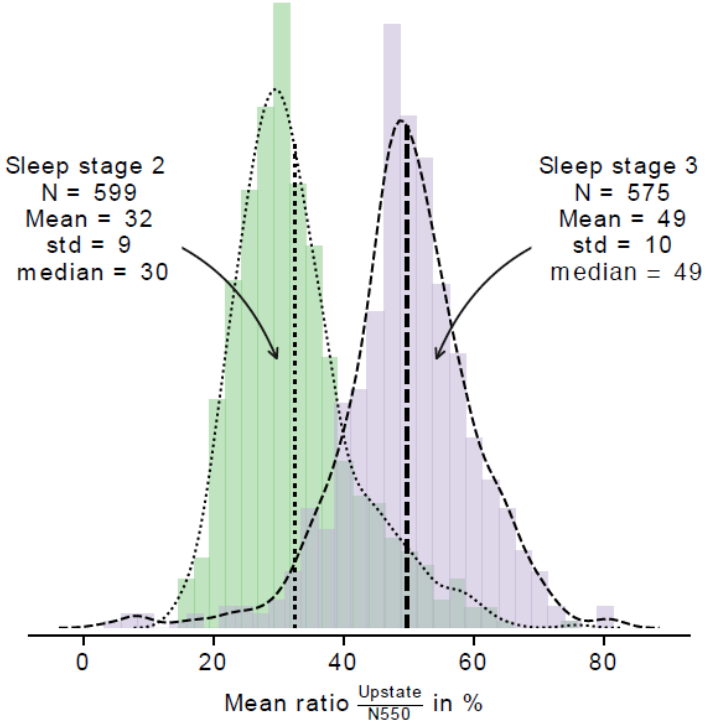

Figure 2-5: Distribution of the up-state/N550 ratio in sleep stage 2 and 3 from the CFS datasets. Vertical lines represent the mean. Reproduced by permission of Oxford University Press https://doi.org/10.1093/sleep/zsaa077.

Manual review of KCs and random segments confirmed good scoring of the algorithm in slow wave sleep with an F1-score between the algorithm and the scoring consensus of 0.52. When comparing individual scores with the algorithm, agreement was similar (mean F1-score ± SD; 0.51 ± 0.13). The relatively low F1-score is not surprising since inter-scorer agreement was similar (0.50 ± 0.10). Of note, the inter-scorer agreement found in sleep stage 3 is similar to previously published inter-scorer agreements for sleep stage 2 (Devuyst et al., 2010). Pairwise F1-scores between scorers is shown in Table 2-5, demonstrating highly variable inter-scorer agreement with the lowest and highest pairwise F1-score ranging from 0.33 and 0.70 respectively. A total of 719 waveforms out of 1000 were scored as a KC by at least one of the five scorers. From these KCs, only 44%, 20% and 6% were scored by at least 3, 4 and 5 scorers, respectively. The probability of scored KCs increased along with

inter-scorer agreement (mean, (95% CI)); 80 (77, 83) % for KCs scored by one scorer versus 88 (84, 93) % and 90 (86-95)% for 2 and 5 scorers, respectively.

Table 2-5: Pairwise F1-score of manual scoring (500 K-complexes and 500 non-K-complexes) in slow wave sleep.

		Scorers				
		A	**B**	**C**	**D**	**E**
Scorers	**E**	0.54	0.48	0.36	0.50	1
	D	0.48	0.33	0.53	1	
	C	0.58	0.70	1		
	B	0.49	1			
	A	1				

2.3.6 Clinical relevance and further work

Figure 2-6 shows the overnight variation in KC density, calculated over a 5-minute period for 4 participants from the CFS dataset. This shows a young and healthy individual (curve one), with a high amplitude overnight KC density function that consistently follows traditionally scored sleep cycles, supporting that KCs occur more frequently and then dissipate over time with each subsequent sleep cycle. As a group, all participants in the CFS dataset show a substantial (Mean (95%CI)) 42.5 (38.6, 46.4) % decrease (1 sample t-test $p<0.001$, N=585) in KC density from the first to the second half of the night. Curve 2 in Figure 2-6, representing a 48-year-old adult with no recorded cardiovascular disease, hypertension, medication use, a normal BMI and no excessive alcohol intake, shows a similar pattern but lower amplitude. In the group data, the mean KC density overnight, as shown by a linear regression model controlled for sex and BMI, showed a decrease of 0.32 KC/min (0.32 (0.28, 0.37); $p <0.001$; N=583) for each 10-year increase in age.

Curve 3 and 4 in Figure 2-6 are examples of participants with sleep apnea and alcoholism, respectively, showing reduced KC density. Indeed, in the group data, a linear regression model controlling for age, sex and BMI showed that alcoholism was associated with 0.42 (0.13, 0.78) KC/min decrease ($p = 0.006$, N=583). Furthermore, a decrease of 0.16 (0.02, 0.32) KC/min was associated ($p = 0.02$, N=152) with an increase of 10 in AHI, in a linear regression model

controlled for age, sex, and BMI. However, this association needs to be interpreted carefully since the AHI was only available for 25% of the dataset. Collectively, these data support the potential value of KC density as a marker of normal versus abnormal sleep health after adjustment for clear age effects. This marker could be substantially more informative than traditional measures of sleep.

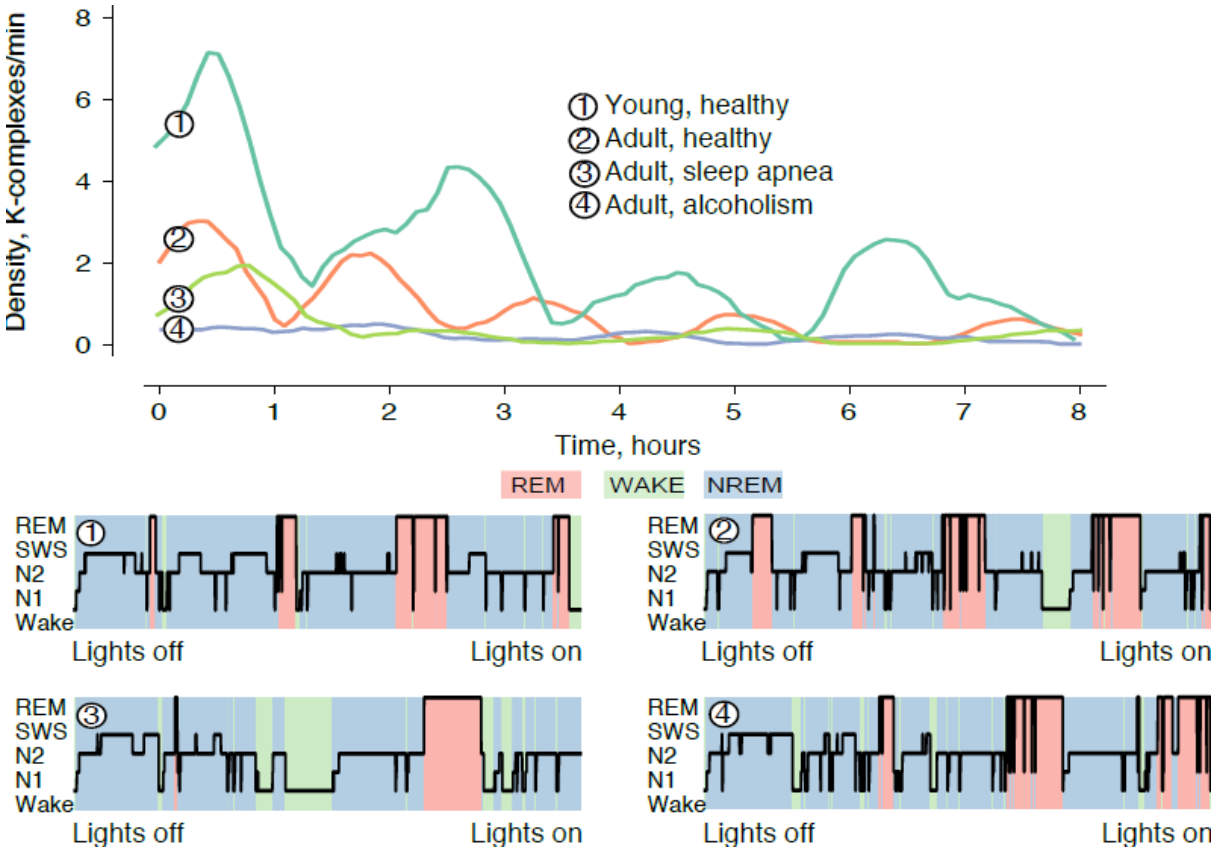

Figure 2-6: Variation in K-complex density overnight for 4 individuals from the CFS dataset, with their respective hypnograms. Reproduced by permission of Oxford University Press https://doi.org/10.1093/sleep/zsaa077.

2.4 Discussion

This Chapter presents a validated and high-performing KC detection algorithm based on probabilistic classification methods. The automatic probabilistic scoring is well suited for KCs with variable amplitudes and waveform features, between and within individuals. The developed algorithm shows major promise as an effective tool for exploring relationships between KCs and clinical outcomes.

Varying KC probabilities within participants is in accordance with fMRI and intracellular-EEG findings (Amzica and Steriade, 1997a, Jahnke et al., 2012), which attribute the origin of the KC to different parts of the cortex. The KC seems to be strongest in the pre-frontal cortex (Colrain, 2005) and relatively weaker in posterior and lateral scalp regions. The proposed algorithm could test this hypothesis by comparing scored probabilities across channels, in which case a decrease in probabilities from the frontal to the parietal site (for example, C3/C4), would be expected. On the other hand, Mak-McCully et al. (2014) (2015) showed that evoked and spontaneous KCs can be quasi-synchronous over much of the cortical surface in humans. Therefore, more uniformly distributed probabilities across channels might also be anticipated. Of note, since KCs can occur locally (Mak-McCully et al., 2014), and since the Montreal archive of sleep study only provides KC scoring on the C3 channel, meaningful comparisons of the algorithm scoring to other channels such as C4, F4 are problematic. Future studies with both manual and automated KC scoring of multiple channels, ideally combined with high-density EEG recordings to better define KC source localization and relationships with manual scoring are required.

While the KCs detected in sleep stage 3 had a slightly higher up-state/N550 ratio, this remained significantly lower than 100%. The difference in ratio between sleep stage 2 and 3 remains unclear, but it could be due to earlier KCs components, such as P200 or P400 (Colrain, 2005), which could be stronger in sleep stage 3 than in sleep stage 2. Mak-McCully et al. (2014) reported a higher KC occurrence rate in sleep stage 3 than in sleep stage 2 and Crowley et al. (2002) showed that KCs are smaller in elderly people, consistent with the findings in this Chapter. Hence, since the CFS dataset is imbalanced towards older people, and assuming that the EEG noise is similar across age, the higher up-state/N550 ratio might be due to a lower signal-to-noise ratio in sleep stage 3 in elderly people. Nonetheless, the algorithm still appears to perform well in sleep stage 3 and is therefore likely to be a practical and convenient tool for automated exploration of large clinical trials datasets for answering applied and fundamental research questions.

Furthermore, the reported inter-scorer F1-score of 0.50 in sleep stage 3 is inline with previously reported inter-scorer agreement of around 50% (Bremer

et al., 1970, Devuyst et al., 2010). The averaged F1-score between the algorithm and individual scoring was of 0.51, which is in line with inter-scorer agreement. Application to larger datasets, with KCs scored by multiple scorers, would likely be useful to better tune the algorithm to greater variability in KC shape and to reduce the impact of inter-scorer disagreement on algorithm performance beyond the current algorithm systematically exposed to data from only 19 healthy individuals. Nevertheless, probabilistic scoring was able to quantify both uncertainties, with a lower probability attributed to noisier KCs, and a higher probability to well-defined KCs and KCs where scorers agree on scoring.

Conclusion

This Chapter presents a high-performing and publicly available KC detection algorithm based on DNN and Gaussian processes. The main strengths of this approach include a probabilistic output and reliable automated KC detection essential to support systematic large-scale analysis not possible with traditional manual human scoring. The algorithm outperforms state-of-art algorithms previously reported in the literature. The probabilistic approach also helps to investigate the overall quality of an EEG recording, and to examine and deal with underlying EEG differences in sample populations; key features likely to be necessary for large scale systematic studies of KCs and their relationships with clinical outcomes. Finally, algorithm performance in sleep stage 3, was consistent with previously published findings supporting that KCs are also present and detectable in sleep stage 3 with appropriate methods.

While the Cleveland Family study contained some participants with sleep apnea, further work remains required to systematically test algorithm findings in clinical datasets. Alcoholism, neurological disorders and/or restless leg syndrome are known to be associated with differences in shape and density of KCs, and hence clearly warrant further investigation.

CHAPTER 3. K-COMPLEXES ARE A SENSITIVE MARKER OF NOISE–RELATED SENSORY PROCESSING DURING SLEEP

This Chapter was published in the following journal article.

Citation: Lechat B, Hansen K, Micic G, Decup F, Dunbar C, Liebich T, Catcheside P, Zajamsek B, K-complexes are a sensitive marker of noise–related sensory processing during sleep: A pilot study. *Sleep.* 2021 (In press). Doi: https://doi.org/10.1093/sleep/zsab065

Author Contributions: study design—BL, KH, GM, FD, CD, TL, PC, BZ data collection—BL, FD, CD, TL; analysis of data— BL; interpretation of findings— BL, KH, PC and BZ; preparation of initial manuscript— BL; revision of final manuscript— BL, HK, GM, FD, CD, TL, PC, BZ.

Summary

Study Objectives: The primary aim of this Chapter was to examine dose-response relationships between sound pressure levels and K-complex occurrence probability for wind farm and road traffic noise. A secondary aim was to compare K-complex dose-responses to manually scored EEG arousals and awakenings.

Methods: Twenty-five participants underwent polysomnography recordings and noise exposure during sleep in a laboratory. Wind farm and road traffic noise recordings of 20-sec duration were played in random order at 6 SPLs between 33 - 48 dBA during established N2 or deeper sleep. Noise periods were separated with periods of 23 dBA background noise. K-complexes were scored using a validated algorithm. K-complex occurrence probability was compared between noise types controlling for noise SPL, subjective noise sensitivity and measured hearing acuity.

Results: Noise-induced K-complexes were observed in N2 sleep at SPLs as low as 33 dBA (Odds ratio, 33 dBA vs 23 dBA, mean (95% confidence interval); 1.75 (1.16, 2.66)) and increased with SPL. EEG arousals and awakenings were only associated with noise above 39 dBA in N2 sleep. K-complexes were 2 times more likely to occur in response to noise than EEG arousals or awakenings. Subjective noise sensitivity and hearing acuity were associated with K-complex occurrence, but not arousal or awakening. Noise type did not detectably influence K-complexes, EEG arousals or awakening responses.

Conclusion: These findings support that K-complexes are a sensitive marker of sensory processing of environmental noise during sleep and that increased hearing acuity and decreased self-reported noise sensitivity increase K-complex probability.

3.1 Introduction

A report from the World Health Organization (2011) estimated that at least one million healthy life years are lost every year due to environmental noise in western Europe alone, with annoyance and sleep disturbance the main contributing factors. Sleep disturbance, including increased rates of awakenings due to transportation noise is well established (see Basner and McGuire (2018) for a comprehensive review). However, evidence surrounding sleep disturbance from others noise types, such as wind farm noise, are only emerging and warrants further studies.

The primary markers of sleep disturbance from environmental noise (see section 1.1.3b for more details) are micro-arousals and awakenings (Basner et al., 2008, Elmenhorst et al., 2012, Jalali et al., 2016, Smith et al., 2016, Basner and McGuire, 2018, Rudzik et al., 2018, Smith et al., 2020). These studies suggest that environmental noises of higher SPLs are more likely to elicit awakenings and micro-arousals, above certain SPL thresholds (which may be between 33 and 38 dBA) (Basner and McGuire, 2018). However, more subtle changes, such as sub-cortical autonomic responses including heart-rate acceleration (Griefahn et al., 2008), peripheral vasoconstriction (Catcheside et al., 2002) and KCs (Colrain, 2005) are observed with noises that do not necessarily elicit cortical arousals or awakenings.

McGuire et al. (2016) found that a large portion of variance, between 40 - 60%, in noise-induced sleep disturbance measured using traditional markers of arousals and awakenings, is due to inter-individual differences other than age and sex, suggestive of physiological differences in auditory processing in participants without clinically relevant hearing loss. Hearing acuity appears likely to explain some of this variance (Lee et al., 2018), but other factors also appear likely to contribute. For example, high noise sensitivity, a psychological trait contributing to increased noise reactivity, has been associated with a lower EEG amplitude of noise-event related potential components (Kliuchko et al., 2016), suggesting an association between auditory processing and noise sensitivity. However, evidence to support an influence of noise sensitivity on physiological responses during sleep are currently restricted to macro-structural sleep parameters such as the probability of evoking an awakening (Marks et

al., 2008), sleep onset latency, wake after sleep onset, or total sleep time (Marks and Griefahn, 2007).

In the context of noise, the likelihood of evoking a KC ranges from 10% (Forget et al., 2011) to 80-90% (Bastien and Campbell, 1992, Colrain et al., 1999, Nicholas et al., 2006, Colrain et al., 2010). Multiple demographic, behavioural and clinical factors can affect the rate of evoked KC production, such as age (Colrain et al., 2010), sleep pressure (Nicholas et al., 2002), alcoholism (Colrain et al., 2009), neuropathology (Crowley et al., 2005) and OSA (Afifi et al., 2003, Nguyen et al., 2016). Furthermore, the likelihood of evoking a KC is also dependent on stimulus characteristics, such as SPL or noise type (Bastien and Campbell, 1992, Colrain et al., 1999, Forget et al., 2011). However, most auditory evoked KC studies have used relatively high SPL and short duration simple stimuli (usually pure tones) to maximize KC occurrence (Colrain, 2005); which are not representative of real-world environmental noise.

Thus, the primary aim of this pilot study, designed to help inform a larger trial, was to investigate potential "dose" or exposure-level response relationships between different types of environmental noises including road traffic and wind farm noise at different SPLs, and traditional (arousals, awakenings) as well as more subtle (KC) markers of noise-related sensory disturbance during sleep. A secondary aim was to investigate the potential influence of hearing acuity and subjective noise sensitivity on exposure-response relationships.

3.2 Methods

3.2.1 Participants

Twenty-five healthy individuals (11 males, mean ± SD 26.5 ± 16.4 years; 14 females, 24.1 ± 9 years) were recruited for an overnight polysomnography study. Hearing acuity was assessed by a qualified audiologist and consisted of clinical history, hearing threshold measurement between 125 and 8000 Hz, ear tympanometry, otoscopy and acoustic reflex assessments. Mean hearing threshold across 125 to 8000 Hz was calculated for each participant, and participants were categorized into high and low hearing acuity groups based on the hearing thresholds group median of 3.9 dB in hearing level (dB HL). Self-reported noise sensitivity was obtained via the 21-question Weinstein noise-sensitivity scale (WNSS) (Weinstein, 1978), which ranges from 0 (noise insensitive) to 105.

Participants were categorized as noise-sensitive if their WNSS score exceeded 54, the mean value reported by Weinstein (Weinstein, 1978). Participants provided informed consent and were reimbursed for their time. The study was approved by the Flinders University Social and Behavioral Research Committee.

Participants were recruited using recruitment posters posted on websites, public notice boards, and word of mouth. Participants who met the eligibility described below were enrolled by study personnel. Study inclusion required participants to have a BMI < 30 kg.m^{-2}, to be non-smokers, free from health problems that may affect sleep and to report healthy sleep, defined as a score < 6 on the PSQI and an average of >85% sleep efficiency based on self-reported wake versus sleep opportunity time. Participant with insomnia or excessive daytime sleepiness were excluded, assessed using the insomnia severity index (excluding insomnia severity index > 8) and the Epworth sleepiness scale (excluding ESS > 10), respectively. Finally, self-reported onset and offset seep times were required to be within 2 hours of each other on weekday vs weekend nights.

3.2.2 **Experimental procedure**

Participants were exposed to block-randomized 20-sec environmental noise and background noise (control) samples, with an inter-stimulus interval of 20-sec. Noise samples were continued throughout periods of consolidated N2 or deeper sleep, except in the event of awakening (EEG arousals ≥15-sec), in which case the noise battery was paused at the end of any currently playing stimulus, and only recommenced after N2 sleep was re-established. Noise samples were played at 6 different SPLs ranging from 33 to 48 dBA in 3 dBA increments. Controls consisted of quiet background noise at 23 dBA (Figure 3-1). Noise stimuli included two road traffic noise samples recorded near ($<$ 100 m) and away ($>$ 700 m) from a busy road, and three types of WFN. Two WFN samples, commonly referred to as 'swish' and 'thumping', included amplitude modulation, a periodic variation in noise amplitude due to blade rotation. These noise samples were measured at short- ($\approx$ 700 m) and long-range ($\approx$ 3 km) from a wind farm, respectively. For comparison, the third WFN sample was a modified long-range wind farm noise with amplitude modulation removed via filtering

(sixth degree notch filter centered at 46 Hz). Environmental noise was reproduced using RME Babyface Pro sound card, LabGruppen C 10:4X amplifier and a Krix Pheonix V2.1 loudspeaker. The loudspeaker was placed next to participants' bed, approximately 2 meters away (Figure 3-1). Equivalent SPL in dBA were measured over a 20-sec period at the participants' head location.

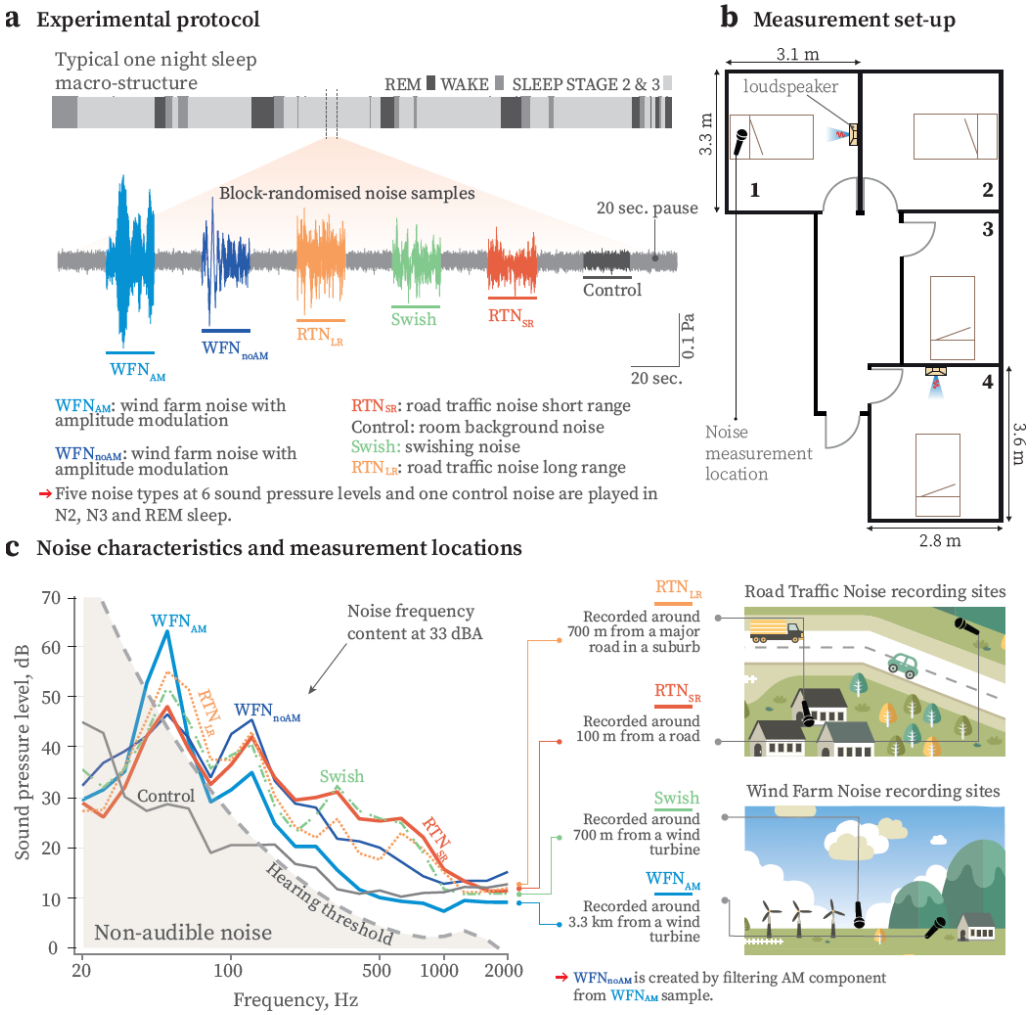

Figure 3-1: A, experimental protocol showing an example of block-randomised noise stimuli. B, noise characteristics and measurement locations. Reproduced by permission of Oxford University Press
https://doi.org/10.1093/sleep/zsab065

3.2.3 Sleep recordings

Participants undertook full polysomnography including EEG using the 10-20 placement system (EEG; F3, F4, C3, C4, Cz, O1 and O2 referenced to M1 or M2), left and right EOG, chin EMG, limb movements, ECG and finger pulse

oximetry measurements. Electrodes were fitted and refitted to achieve impedances <5 kΩ where possible, although 2 participants showed 6 and 9 kΩ (group mean ± SD; 3 ± 2 kΩ). Signals were amplified and recorded using Grael 4K Polysomnography (Compumedics Ltd.) at a 512 Hz sampling frequency using Profusion 4 EEG acquisition software.

3.2.4 EEG processing and K-complex scoring

Sleep stages, awakenings and arousals were scored by a single scorer, blinded to noise conditions, according to American Academy of Sleep Medicine manual sleep scoring guidelines (Iber et al., 2007). An arousal or awakening was considered to have been evoked by a noise if it occurred anywhere during the noise presentation (i.e., 20-sec window from noise onset). Arousals and awakenings were grouped as a single event given that their incidence was low (Table A2). Only noise events presented during N2 and N3 sleep were analyzed since KCs are not generally considered to occur in REM sleep (Colrain, 2005).

The EEG time series were down-sampled to 128 Hz, for faster processing, and filtered with high– and low–pass first order filters with cut–offs of 0.30 Hz and 35 Hz, respectively. For the KC analysis only the C3 signal referenced to M2 was used. For automated detection of KCs, an algorithm based on DWT, DNN and Gaussian probabilistic classification approach was used with an established sensitivity of 0.86, a precision of 0.72; and a global F1-score of 0.78 (Lechat et al., 2020), following training on C3 electrode data from the MASS (O'Reilly et al., 2014). This algorithm scores a waveform with a probability of being a KC ranging from 50% to 100% where larger amplitude and more well-defined KCs are attributed higher probability. A KC was considered to be evoked by a noise if it occurred within 2 seconds of noise onset and had an algorithm scored probability of being a KC (hereafter referred to as "scoring threshold") greater than 50%. A 2 second window was chosen based on the expected short latency from stimulus onset (Crowley et al., 2004, Colrain, 2005, Willoughby et al., 2020), and to reduce spontaneous KC detection over the remainder of the stimulus exposure. The probability of occurrence of a KC to each noise (type and SPL) was defined based on the proportion of noise presentations that evoked a KC.

3.2.5 Statistical analysis

Descriptive statistics included Fisher's exact test for categorical variables and analysis of variance for continuous variables. The association between KC probability of occurrence and SPL was examined using mixed effects logistic regression with participant number as a random effect, each with a separate intercept (Model 1). Noise type effects on KC response probability were tested based on the interaction between noise type and SPLs (Model 2). Three other models with interactions were constructed to examine possible modulation of KC responses to noise SPL by hearing acuity (Model 3), noise sensitivity (Model 4), or both (Model 5). Results were also compared against more traditional methods, by repeating analyses with evoked arousal and awakening as dependent factors in similar models. Results are reported as odds ratios (ORs) with their respective 95% confidence intervals (CI). Summary graphs for each model are presented with marginal probabilities and ORs.
Statistical analysis was performed using the computing environment R (R Core Team, 2019) with lme4 (Bates et al., 2015) open source package for logistic analysis.

3.2.6 Sensitivity analysis

A sensitivity analysis was performed to further examine findings. The scoring threshold of KCs was elevated from 50% (main analysis) to 95% (in 10% increments up to 90%). By augmenting the scoring threshold, the scoring becomes more "conservative" as shown previously (Lechat et al., 2020). The association between KC probability of occurrence and SPL (Model 1) was then re-examined for these more conservative scorings. Significant interactions found in the main analysis were further examined by increasing the KC scoring threshold from 50% to 75%.

3.3 Results

3.3.1 Participant characteristics

The analysis included 21 participants (Table 3-1) following exclusion of two participants due to technical polysomnography failure (due to timing system failure the onset of noises played could not be determined accurately) and a further two whose age was more than double the group mean (given that age

has a strong effect on the frequency and the shape of the KC, (Crowley et al., 2002, Crowley et al., 2004) likely to confound).

Eight participants were categorized as noise-sensitive, but there were no significant differences in hearing acuity or hearing acuity categories between the noise sensitive and non-sensitive groups. The number of presented noise stimuli was also not different between noise sensitivity groups. The mean number of noise presentations was around 27 and 21 samples per SPL in N2 and N3 sleep, respectively (Table A1). Between 1 to 5% of all noise presentations evoked an EEG arousal (EEG changes $>$ 3-sec and $<$ 15-sec) and only 1 to 4% of all noise presentations evoked an awakening (EEG changes $\geq$ 15-sec, Table A2). Given low frequency of occurrence, particularly for awakenings, EEG arousals and awakenings were only considered together in subsequent analyses.

Table 3-1: Participant characteristics stratified by noise sensitivity group. Data are presented in n (%) for categorical variables, median [IQR] for non-normally distributed variables and mean (SD) for normally distributed continuous variables.

	All	Noise non-sensitive	Noise sensitive	p-value
n	21	13	8	
Age	22.00 [21.00, 22.00]	22.00 [22.00, 23.00]	20.50 [19.00, 22.00]	0.011
sex: male (%)	12 (57.1)	8 (61.5)	4 (50.0)	0.673
Noise sensitivity	50.00 [44.00, 59.00]	44.00 [36.00, 50.00]	59.50 [59.00, 67.00]	<0.001
Hearing acuity in dB HL*	3.90 [2.00, 6.80]	4.09 [3.40, 9.09]	3.45 [1.98, 4.93]	0.404
Total sleep time, in hours	7.57 [7.00, 7.78]	7.57 [7.09, 7.78]	7.58 [6.94, 7.84]	0.942
REM, in %	18 [15, 20]	19 [15, 22]	16 [13, 18]	0.128
N1, in %	7 [5, 8]	6 [5, 8]	8 [6, 9]	0.051
N2, in %	43 (6)	42 (7)	45 (4)	0.328
N3, in %	22 [18, 26]	20 [18, 24]	25 [20, 26]	0.515
Wake, in %	7 [7, 11]	7 [7, 12]	7 [7, 8]	0.885
Arousal index events/hours	4.82 (2.08)	4.90 (2.11)	4.69 (2.18)	0.836

*Calculated using mean hearing threshold between 125 and 8000 Hz

3.3.2 K-complexes in N2 sleep

In Model 1, and regardless of noise type, KCs were more likely to occur with increasing SPLs (overall effect, $\chi^2 = 96.8$, $df = 6$, $p < 0.001$, Figure 3-2). Even at 33 dBA, 10 dBA higher than the background noise, the probability of KC occurrence to noise presentation was 1.75 times greater (Figure 3-2B) than the background noise, increasing to 4.6 times greater at 48 dBA, although the absolute probability of occurrence remained low. The interaction between noise type and SPL (Model 2) was not significant ($\chi^2 = 24.7$, $df = 20$, $p = 0.21$).

However, these results should be interpreted cautiously given the mean number of noise presentations stratified by noise SPLs and types, was only around 6.

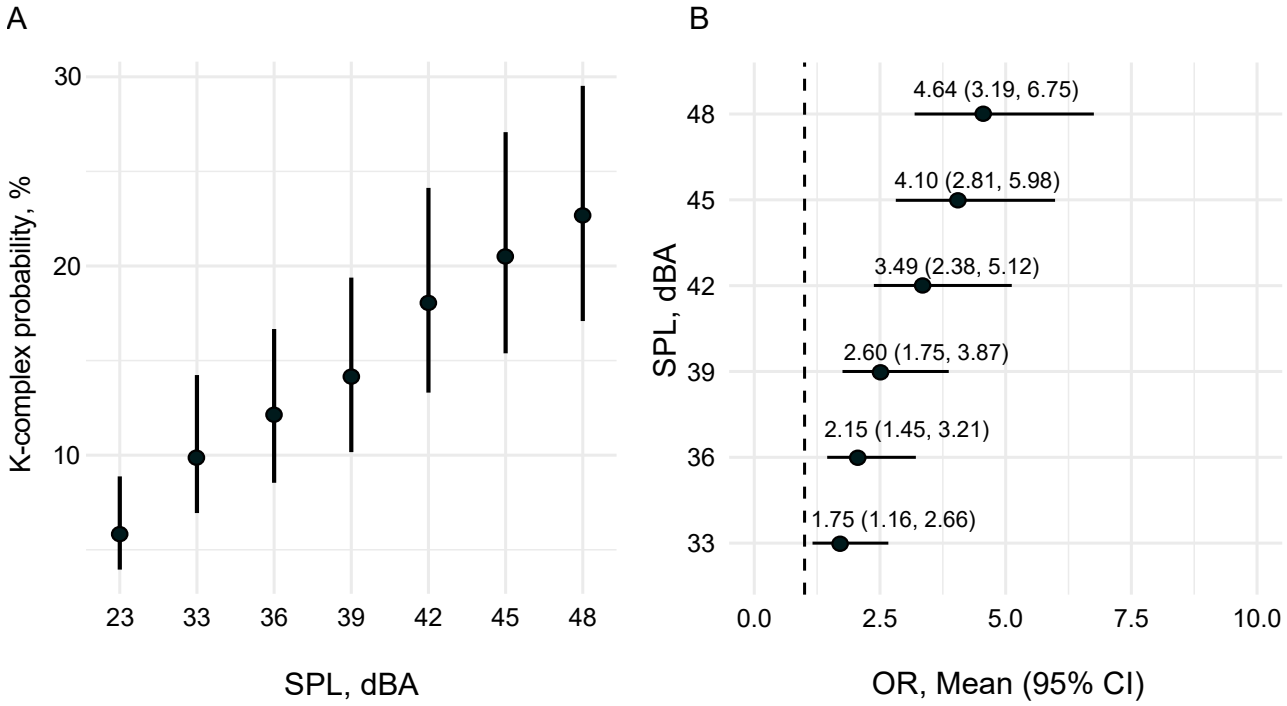

Figure 3-2: Association between noise SPL and K-complex response in N2 sleep. A, probability of occurrence of a K-complex at a given noise level. B, odds ratio (95% CI) of evoking a K-complex at a given SPL compared to background noise. Reproduced by permission of Oxford University Press https://doi.org/10.1093/sleep/zsab065

The overall interaction effect between hearing acuity group and SPL was significant ($\chi^2 = 13.2$, $df = 6$, $p = 0.04$, Figure 3-3A), where KC response occurrence was significantly greater than background noise only at SPLs $\geq$ 42

dBA for the low hearing acuity group (Figure 3-3B), but at all SPLs in the high hearing acuity group (Figure 3-3C).

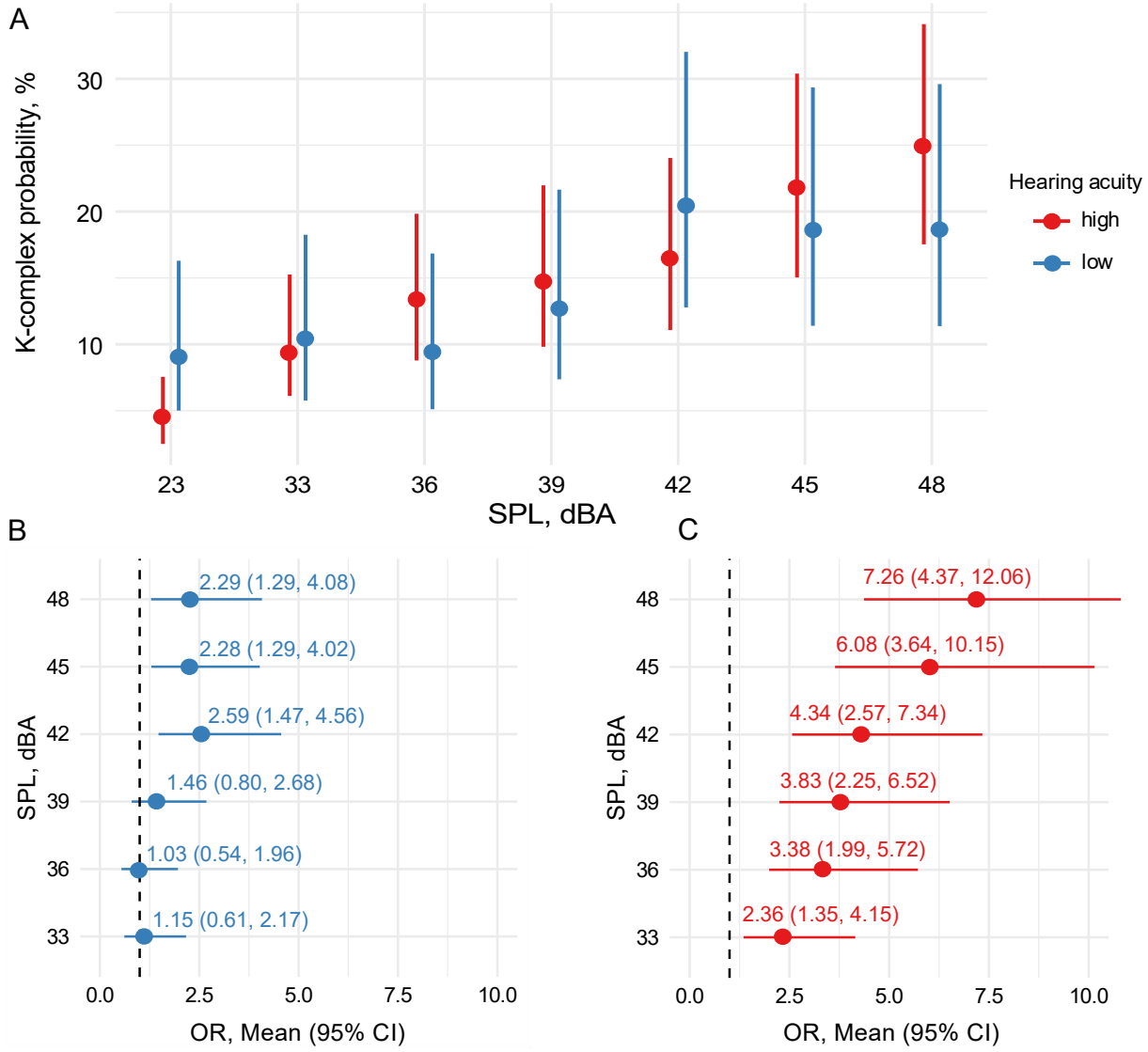

Figure 3-3: Association between noise SPL and the K-complex response in N2 sleep for participants with hearing thresholds ≤ 3.9 dB HL (red) or > 3.9 dB HL (blue). A, probability of occurrence of a K-complex at a given noise level. B and C, odds ratio (95% CI) of evoking a K-complex at a given SPL compared to background noise. Reproduced by permission of Oxford University Press https://doi.org/10.1093/sleep/zsab065

The odds of a KC response to noise stimuli was dependent on noise sensitivity (Model 4, $\chi^2 = 15.7$, $df = 6$, $p = 0.01$) and was lower for noise sensitive compared to low noise sensitivity participants (Figure 3-4A). Noise sensitive participants only showed KC responses above background noise levels for noise SPLs ≥ 39 dBA (except for 42 dBA) whereas non-noise sensitive participants

showed significantly elevated odds of KCs at all SPLs above background noise (Figure 3-4C).

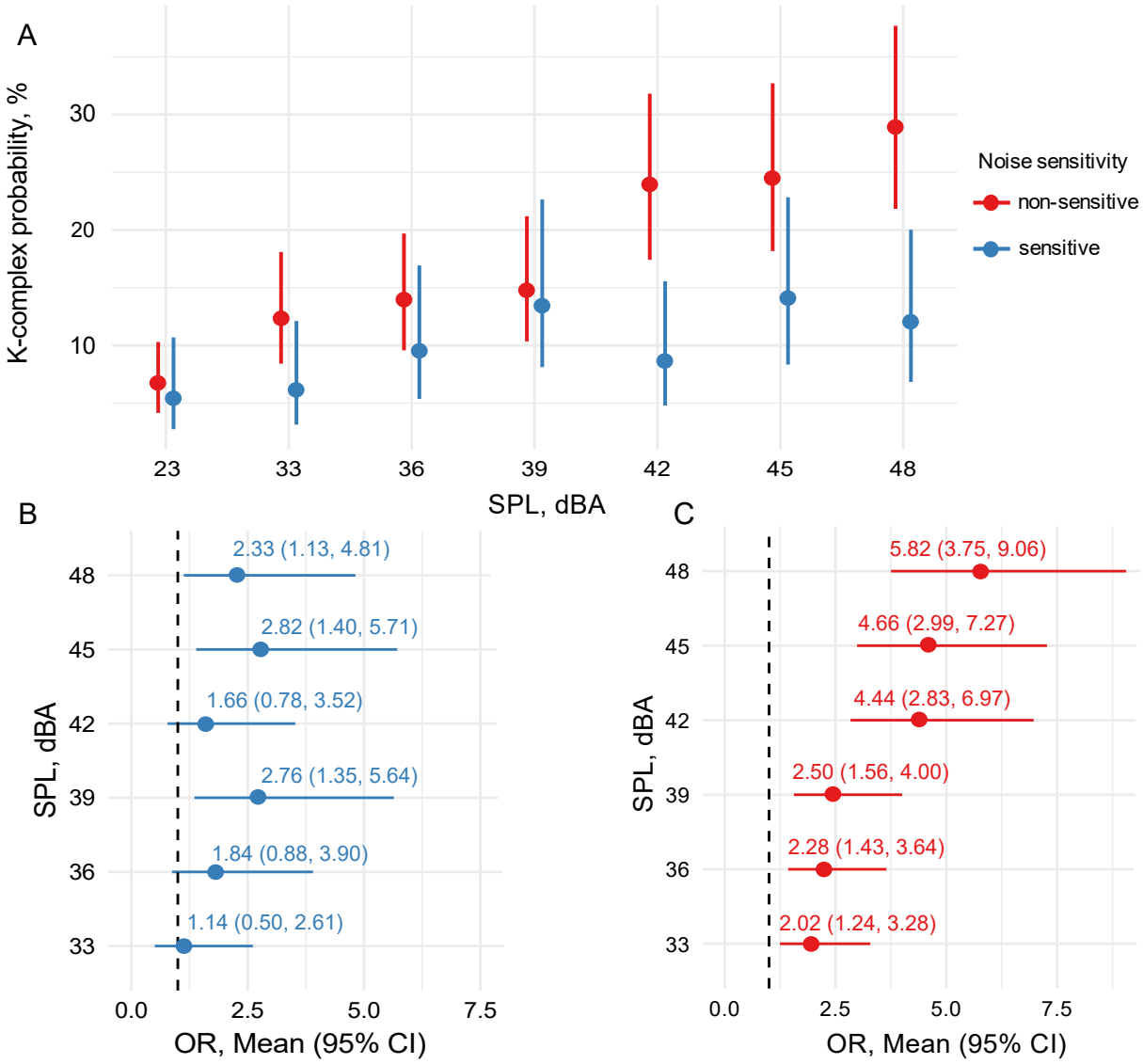

Figure 3-4: Association between noise SPL and the K-complex response in N2 sleep for participant with a Weinstein noise sensitivity score ≤ 54 (red) or > 54 (blue). A, probability of occurrence of a K-complex at a given noise level. B and C, odds ratio (95% CI) of evoking a K-complex at a given SPL compared to background noise. Reproduced by permission of Oxford University Press https://doi.org/10.1093/sleep/zsab065

Interactions between noise SPL and noise sensitivity or hearing acuity remained significant in Model 5, while the interaction between noise sensitivity and hearing acuity was not significant ($\chi^2 = 3.40$, $df = 6$, $p = 0.75$). Thus,

the effect of noise sensitivity and hearing acuity on the dose-response relationship between KC occurrence and noise SPL remained similar in Model 5 (Figure A1) compare to model 3 and 4 shown in Figure 3-3 and Figure 3-4, respectively.

The estimated association between noise SPL and KC responses remained similar in sensitivity analysis. An increased scoring threshold resulted in fewer scored KCs across all SPLs, and thus the marginal probability of evoked KCs for each SPL decreased. However, the increase in marginal probability compared to background noise remained significant for all levels at all thresholds (except 33 and 36 dBA at 90 and 95%), as shown in Table A3. Therefore, the effect of noise SPL on the KC responses is very consistent across different KC scoring thresholds. An additional advantage of automatic scoring at different scoring thresholds is that KC classification uncertainty can be estimated. For example, pooled estimates across all scoring thresholds >50% indicate a 19.5 ± 2.8% probability of an evoked KC at 48 dBA. Finally, interactions between noise SPL and hearing acuity ($\chi^2 = 12.5$, $df = 6$, $p = 0.05$), and noise sensitivity ($\chi^2 = 14.3$, $df = 6$, $p = 0.03$), remained largely similar in sensitivity analyses.

3.3.3 K-complexes in N3 sleep

In comparison to N2 sleep, higher SPLs were needed to elicit a KC in N3 sleep (Figure A2). Overall probabilities for noise-evoked KCs remained similar to N2 sleep, but the probability of spontaneous non-evoked KCs occurring during background noise was around 22% in N3 versus around 6.6% in N2 sleep. Thus, the resulting ORs are generally smaller and only KC responses to noise stimuli with SPLs ≥ 39 dBA were significantly higher than KC responses during background noise (apart from 42 dBA, Figure A2). The interaction between noise SPLs and hearing acuity was not significant in N3 sleep (overall effect, $\chi^2 = 2.01$, $df = 6$, $p = 0.91$).

In N3 sleep the effect of noise SPL on the KC-response was different between noise sensitivity groups ($\chi^2 = 15.3$, $df = 6$, $p = 0.02$), where noise-evoked KC responses were only apparent in the non-noise sensitive group and were not significantly different from background noise at any SPL in the noise sensitive group (Figure A3). Similar to findings with N2 sleep, sensitivity analysis revealed that different scoring thresholds had little impact on the dose-response

relationship between KC probability of occurrence and noise SPL (Table A3). Furthermore, the interactions between noise sensitivity and SPL also remained significant in sensitivity analysis ($\chi^2 = 13.08$, $df = 6$, $p < 0.041$).

3.3.4 Arousals

The effect of SPL on arousal and awakening occurrence was significant ($\chi^2 = 27.32$, $df = 6$, $p < 0.001$), but with a smaller effect size compared to the KC occurrence (Figure 3-5A). Furthermore, the effect was significant only for noise played at an SPL $\geq$ 39 dBA (Figure 3-5B). Noise-sensitivity ($\chi^2 = 3.33$, $df = 6$, $p = 0.76$) and hearing acuity ($\chi^2 = 2.35$, $df = 6$, $p = 0.88$) do not effect arousal/awakening responses to noise, although these results need to be interpreted with caution given that only 114 noises (3.3% of all noise presentations) evoked an arousal or awakening. In N3 sleep, there was no effect of SPL on arousal/awakening ($\chi^2 = 3.89$, $df = 6$, $p = 0.69$).

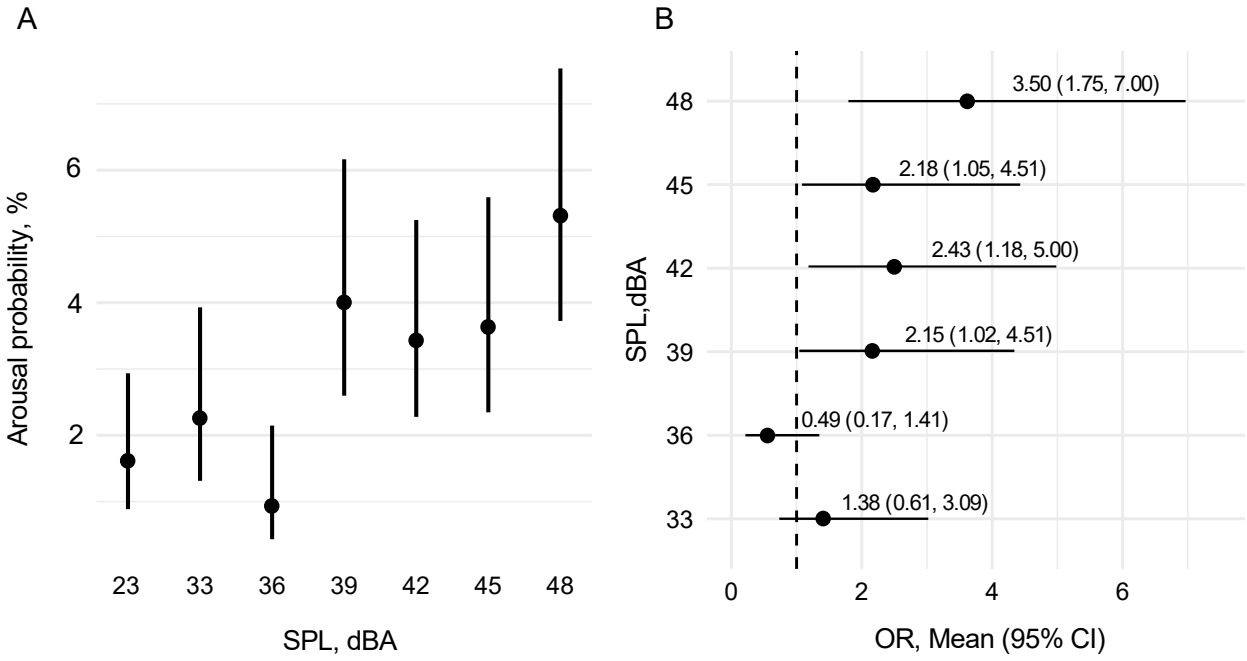

Figure 3-5: Association between noise SPL and arousal/awakening response in N2 sleep. A, probability of occurrence of an arousal/awakening at a given noise level. B, odds ratio (95% CI) of evoking an arousal/awakening at a given SPL compared to background noise. Reproduced by permission of Oxford University Press https://doi.org/10.1093/sleep/zsab065

Co-occurrence of KCs and arousals was low. A total of 21% (N = 24) evoked arousals in N2 sleep were preceded by a KC (between the noise onset and

within 10 seconds before the arousal) and only 0.7% of all noise presentations that evoked a KC were followed by an arousal. At 23 dBA (control) and 48 dBA (loudest noise) only one (out of 692, 0.1%) and 7 noises (out of 597, 1.2%) evoked a KC followed by an arousal, respectively.

3.4 Discussion

This Chapter shows that KCs are a substantially more sensitive marker of sensory processing to environmental noise during sleep compared to more traditional EEG arousal and awakening responses. Two factors likely to explain some of the inter-individual differences to noise-evoked EEG responses during sleep were identified, including hearing acuity and noise sensitivity. Participants with low hearing acuity needed higher SPL noise to elicit KCs and KCs were more likely to be elicited in non-noise sensitive compared to noise-sensitive participants. KCs and arousals or awakenings rarely co-occurred, perhaps more suggestive of largely independent responses to noise rather than a hierarchy of responses, and the concept that KCs may help to suppress arousals and preserve sleep in response to repetitive acoustic stimuli during sleep.
KCs, especially the N550 component, and slow waves share common neuronal sources. Slow waves are formed by a succession of up– and down– states reflecting synchronous firing of cortical neurons and periods of neuronal silence, respectively (Amzica and Steriade, 1997b, Amzica and Steriade, 1997a, Cash et al., 2009, Nir et al., 2011). Within the synaptic homeostasis hypothesis (Tononi and Cirelli, 2006, Tononi and Cirelli, 2014), the central role for slow wave sleep is to downscale synaptic strength to optimize daytime functioning and capacity for new learning without excessive synaptic potentiation. Since KCs reflect an up–state–deprived slow oscillation (Cash et al., 2009), and thus neuronal silence, they have been hypothesized to be a marker of a gating mechanism to incoming stimuli that may help to preserve sleep homeostasis (Halasz et al., 2014, Halasz, 2016). With that in mind, the relative absence of noise-evoked KCs in the noise-sensitive group could potentially indicate a greater propensity to disturbed sleep, or alternatively a higher stimulus threshold required to elicit KCs. Given no evidence to support increased arousals in the noise-sensitive group a higher stimulus threshold appears more likely.

Since sub-cortical processing of auditory stimuli continues during sleep (Bastuji et al., 2002, Campbell and Colrain, 2002), KCs could reflect a mechanism to help counterbalance incoming noise to suppress thalamo-cortical sensory transmission. Noise sensitive participants, who may be more reactive to noise, could produce fewer KCs in response to noise leading to reduced suppression of sensory input (Tononi and Cirelli, 2006). Although speculative, this could theoretically delay the onset to slow wave sleep after a noise event since the brain would tend to be in a chronically more "activated" state.

The dose-response relationships observed between arousal rates and SPLs are largely in accordance with previously reported rates of awakenings/shifts to wake or N1 sleep (Basner et al., 2008, Elmenhorst et al., 2012, Basner and McGuire, 2018), which range from 5 to 10% for noise-events at SPLs from 40 dBA to 50 dBA, respectively. Furthermore, the evoked KC probability of around 25% for noise played at 48 dBA in this Chapter is similar to KC elicitation rates of between 25% and 55% reported by Franzen et al. (2012) in one of the most relevant KC studies to have also used relatively long and complex sounds rather than pure tones. The relatively lower probability in this Chapter appears likely to reflect a lower SPL since SPL has a strong effect on evoked KC rates (Colrain, 2005).

Several limitations in this Chapter warrant consideration. Firstly, the low number of repetitions per noise type and SPL does not allow for a meaningful analysis of the effect of different acoustic characteristics on the KC or arousal and awakening responses. Secondly, although studying young and healthy individuals has several advantages, such as reducing the risk of confounding through co-morbidities or age effects on sleep EEG, this limits generalisability to the broader population. Thirdly, the participants were mainly non-noise-sensitive and/or mildly noise sensitive, thus the effect of high noise sensitivity on noise-related sleep disturbance requires further research in a larger study. Finally, although the algorithm showed good agreement with consensus manual scoring irrespective of sleep stages (Lechat et al., 2020), lower inter-scorer agreement in N3 sleep (around 50%) introduces more uncertainty regarding findings in N3 sleep.

Conclusion

Given growing concerns regarding potential adverse health effects of environmental noise on sleep (World Health Organization, 2011, Basner and McGuire, 2018), understanding mechanisms underpinning noise-induced sleep disturbance is important for guiding decisions around public policy and noise guidelines for noise exposure levels during sleep. KCs are clearly a sensitive marker of sensory processing of environmental noise exposure during sleep. Possible long-term effects on sleep and daytime functioning related to the absence versus presence of KCs is uncertain and clearly warrants further research. Remarkably strong interactions between subjective noise-sensitivity and KC-response rates, with an almost two-fold reduction in KC-response occurrence in noise-sensitive participants, and with no corresponding changes in arousal rates, supports the value of KC-responses as an objective marker of sleep-effects from environmental noise. Future studies are clearly warranted to further examine relationships between KC occurrence rates during sleep and subjective outcomes, including different phenotypic responses to environmental noise exposure during sleep.

CHAPTER 4. A NOVEL EEG DERIVED MEASURE OF DISRUPTED DELTA WAVE ACTIVITY DURING SLEEP PREDICTS ALL-CAUSE MORTALITY

This Chapter was under review at the time of thesis submission.

Citation: Lechat B, Hansen K, Melaku Y.A, Vakulin A, Micic G, Adams R.J, Appleton S, Eckert DJ, Catcheside P, Zajamsek B, Sleep delta wave activity fragmentation and mortality: re-analysis of the sleep heart health study. 2021 [AnnalsATS, submitted 11/03/2021, ID: White-202103-315OC]
Author Contributions: study design—BL, KH, BZ and PC.; analysis of data— BL, YAM; interpretation of findings— BL, KH, YAM, AV, GM, RJA, SA, DJE, PC, BZ; preparation of initial manuscript— BL; revision of final manuscript— BL, KH, YAM, AV, GM, RJA, SA, DJE, PC, BZ.

Summary

Study Objectives: Conventional markers of sleep disturbance, based on manual electroencephalography scoring, may not adequately capture important features of more fundamental electroencephalography-related sleep disturbance. This Chapter aimed to determine if more comprehensive power-spectral measures of delta wave activity during sleep are stronger independent predictors of mortality than conventional sleep quality and disturbance metrics.

Methods: Power spectral analysis of the delta frequency band and spectral entropy-based markers to quantify disruption of electroencephalography delta power were performed to examine potential associations with mortality risk in the Sleep Heart Health Cohort (N = 5804). Adjusted Cox proportional hazard models were used to determine the association between delta wave activity disruption at baseline and all-cause mortality over an ~11y follow-up period.

Results: Disrupted delta electroencephalography power during sleep was associated with a 32% increased risk of all-cause mortality compared with no fragmentation (hazard ratios 1.32 [95% confidence interval 1.14, 1.50], after adjusting for total sleep time and other clinical and life-style related covariates including sleep apnoea. The association was of similar magnitude to a reduction in total sleep time from 6.5h to 4.25h. Conventional measures of sleep quality, including wake after sleep onset and arousal index were not predictive of all-cause mortality.

Conclusion: Delta wave activity disruption during sleep is strongly associated with all-cause mortality risk, independent of traditional potential confounders. Future investigation into the potential role of delta sleep disruption on other specific adverse health consequences such as cardiometabolic, mental health and safety outcomes has considerable potential to provide unique neurophysiological insight.

4.1 Introduction

OSA is a common sleep-breathing disorder estimated to affect 936 million adults worldwide (Benjafield et al., 2019). OSA has been associated with an increased risk of hypertension (Grote et al., 2000, Lavie et al., 2000, Nieto et al., 2000), cardiovascular disease (CVD) (Shahar et al., 2001) and all-cause mortality (Young et al., 2008, Punjabi et al., 2009, Kendzerska et al., 2014a). In the Sleep Heart Health Study (SHHS) (Quan et al., 1997) Punjabi et al. (2009), showed that clinical categories of the AHI were a significant independent predictor of all-cause mortality but that a marker of sleep fragmentation, the arousal index was not. Similarly, Shahar et al. (2001) using the same dataset found no association between the arousal index and increased risk of adverse cardiovascular events. Thus, potential underlying mechanisms for a higher risk of mortality/cardiovascular events from sleep-breathing disorders might be due to adverse respiratory disturbance effects on the cardiovascular system, blood pressure and/or hypoxemia, rather than sleep disruption *per se* (Shahar et al., 2001, Punjabi et al., 2009). However, in (Kendzerska et al., 2014a), the total number of awakenings overnight and the total number of periodic leg movements were associated with all-cause mortality; suggesting that sleep fragmentation may be related to adverse health outcomes. The lack of association between clinical outcomes and sleep fragmentation metrics in the other cohorts could also reflect the arbitrary nature of manual EEG scoring rules, developed around practical constraints of traditional paper-based recordings. Conventional sleep staging, respiratory events and particularly arousal scoring also show poor inter- and intra-scorer reliability (Ruehland et al., 2011, Ruehland et al., 2015) which may further confound potential associations with adverse outcomes.

Prerau et al. (2017) emphasized that conventional sleep scoring is purely time-based while some EEG dynamics are only visible through time-frequency analysis. These authors showed how multi-taper spectral analysis could help to quantify EEG dynamics at different timescales, from micro-events to full night EEG recordings. Power spectral analysis of sleep EEG signals is a potentially useful marker of overall sleep quality particularly in sleep disorders, such as OSA (D'Rozario et al., 2017a, Appleton et al., 2019) or insomnia (Krystal et

al., 2002). However, mortality risk associations with spectral analysis derived metrics have not been extensively studied to date.

Delta wave activity is one of the most fundamental EEG features of sleep and indicates strongly synchronous neuronal "down states" of relative neuronal inactivity and "up states" as activity resumes (Nir et al., 2011). Delta wave activity is a key feature of deep sleep (Nir et al., 2011), and cyclically dissipates over the course of normal sleep. In addition, delta activity progressively increases over the course of extended wake, with strong temporal alignment and marked decrements in performance during wake suggestive of localized sleep (Vyazovskiy et al., 2011). Thus, overnight power spectral analysis of delta wave activity could provide useful markers of sleep quality and sleep disturbance.

Measurements of entropy characterize the level of order in a complex signal and are used widely in EEG signal processing (Abasolo et al., 2006). Given the ubiquity of sleep and its importance for normal brain function and health, it was hypothesized that spectral entropy of delta activity would provide a useful marker of sleep quality predictive of health outcomes. Accordingly, this Chapter aimed to investigate if a comprehensive entropy-based measure of sleep delta wave activity fragmentation is a stronger predictor of mortality than conventional polysomnography derived sleep metrics (e.g. wake after sleep onset, total sleep time, arousal index etc.) in a large population cohort, independent of sleep-apnea and other traditional clinical and life-style related covariates.

4.2 Methods

4.2.1 Study design and participants

The SHHS was a cohort study of cardiovascular and cerebrovascular consequences of sleep-disordered breathing. The study design and methodology are described elsewhere (Redline et al., 1998). Full unattended overnight sleep studies (Compumedics P Series System; Abbotsford, Victoria, Australia) from a total of 6,204 participants were pooled from different population-based studies, from which 5,804 are available through an open access dataset from the National Sleep Research Resource (Dean et al., 2016). A second PSG recording was obtained from a subset (N = 3,295) of the participants between January 2001 and June 2003, from which 2,647 are available through the National Sleep

Research Resource. Data from both studies will be used in this Chapter. For both visits, C3–A2 and C4–A1 EEG channels, EOG, a single ECG, chin EMG, nasal thermocouple, oxygen saturation and thoracic/abdominal signals were recorded.

Sleep and EEG arousals were scored according to standard criteria at the time (Kales and Rechtschaffen, 1968, Guilleminault et al., 1992). Apneas were scored as a $\geq$ 75% reduction in the breathing amplitude lasting at least 10 sec as recorded via the thermocouple signal. Hypopneas were identified if the breathing amplitude of thermocouple or thoracic/abdominal band signals decreased by $\geq$ 30% for at least 10 sec. The AHI was defined as the number of apneas and hypopneas, associated with $\geq$ 4% reduction in oxygen saturation per hour of sleep, using an AHI cut-off $<$5 vs $\geq$5 events/hr to define normal vs OSA respectively (Redline et al., 1998).

4.2.2 EEG power spectral analysis

The multi-taper technique (Prerau et al., 2017) was used which minimizes the uncertainty in the spectral estimate across frequencies by multiplying the original signal with multiple orthogonal windows called tapers. This approach reduces windowing artifacts compared to traditional quantitative EEG. Technical details regarding multi-taper method can be found in Prerau et al. (2017). The absolute power was calculated for each 5-second window in delta, theta, alpha, sigma and beta frequency bands (0.5 - 4.5, 4.5 - 8, 8 - 12, 12 - 15, and 15 - 32 Hz, respectively). The EEG power spectral analysis was primarily performed on the C3 channel with a signal quality score, recorded by the human-expert scorer, of $\geq$3 indicating that at least 50% of the EEG signals were artefact-free. The C4 channel was used in sensitivity analysis.

The variation of absolute power in the delta frequency band was captured in a density function. The average power of wake and sleep stage 1 (but not REM) was set to zero, since slow wave activity predominantly occurs in sleep stage 2 and 3 and to help reduce movement artefacts in wake and light transitional sleep. A weighted moving average Gaussian window (120 points, standard deviation = 10) was applied to a density function to reduce stochastic noise. The shape of the final delta wave density function was quantified using spectral entropy, an information measure that determines the degree of uniformity of the distribution (Inouye et al., 1991). The spectral entropy was

computed by calculating the Shannon entropy of the power spectrum (calculated using fast Fourier transform) of the delta wave density function. More details on the power spectral analysis and the calculation of the entropy can be found in APPENDIX B.

Entropy-based measures of the density function are shown in Figure 4-1, including examples of normal sleep (mid-tertile of the entropy distribution function), and abnormal sleep (upper and lower tertiles), along with fragmented sleep due to rapid fluctuations in the delta band and prolonged awakening. A low spectral entropy could arise from a lack of the normal cyclical distribution of delta wave activity during the night (e.g. only one full sleep cycle followed by relatively stable N1/N2 or REM), or through very little slow wave sleep across the whole night. A high spectral entropy could arise from highly fragmented sleep (e.g. rapid and recurrent shifts from deep sleep to wake/N1) with frequent awakenings or arousals during the night. The spectral entropy therefore encapsulates sleep fragmentation both with a higher spectral entropy, for higher frequency fluctuations in delta power, and a lower spectral entropy for short or absent fluctuations.

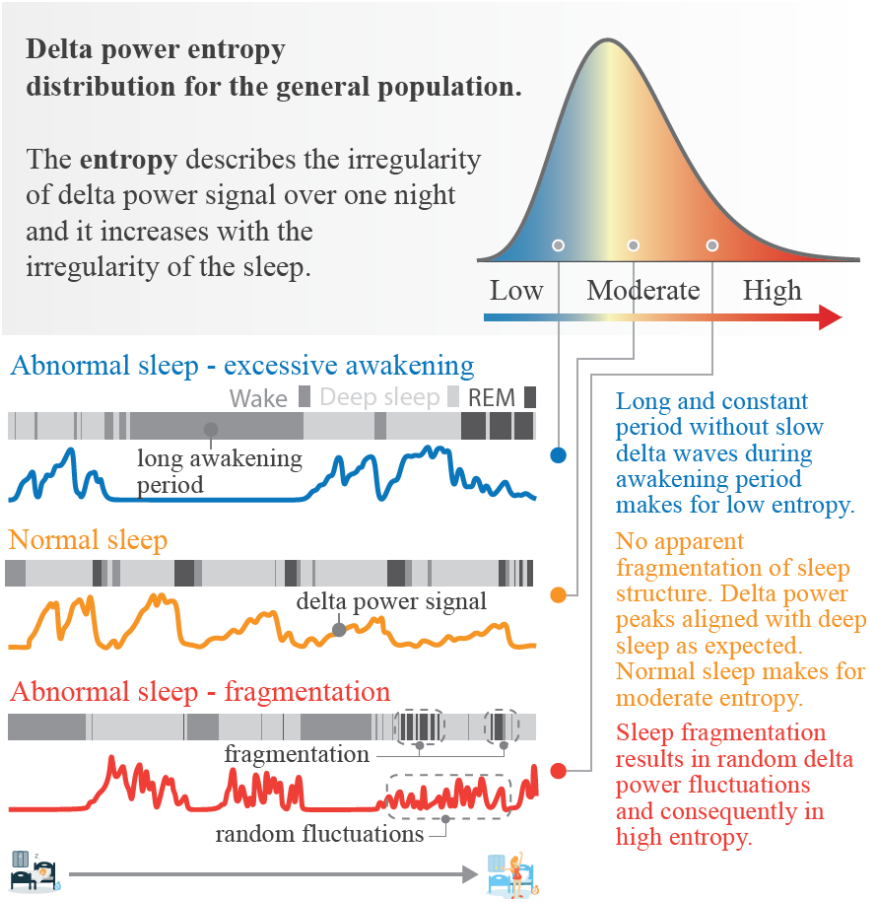

Figure 4-1: Three examples of overnight variations in delta power activity, with their respective hypnograms.

4.2.3 Sleep parameters

The association between mortality-risk and sleep was assessed using traditional polysomnography measures of sleep quality including wake after sleep onset, total sleep time, arousal index, percent of time spent in NREM and REM sleep stages. More specific quantitative EEG measures were also assessed including spectral entropy of the delta density function and overnight (NREM sleep stage) mean of delta, theta, alpha, sigma and beta frequency bands.

4.2.4 Potential confounders

Questionnaires were used to determine baseline and follow-up characteristics including socio-demographics (age, sex and educational status), behavioral factors (alcohol intake, smoking status, physical activity) and participants' BMI. Medical history (hypertension; CVD; diabetes and general health perception) was determined during a baseline examination no more than five years before

the first polysomnography study. Physical activity and general health perception were assessed using the SF36 questionnaire. CVD cases included physician reported angina, heart attack, heart failure, stroke, or if the participant ever underwent coronary bypass surgery, coronary angioplasty or any other heart surgery. AHI and the percentage of sleep time spent with oxygen saturation $<$ 90% were used to control for sleep-related breathing disorders (Kendzerska et al., 2014a).

4.2.5 Outcome assessment

Death from any cause, up until 2011, was identified in another study (Punjabi et al., 2009) using follow-up interviews, written annual questionnaires, telephone contact with study participants or next-of-kin, surveillance of local hospital records and community obituaries and linkage with the Social Security Administration Death Master File. For this analysis, all-cause mortality was used as the primary outcome.

4.2.6 Statistical analysis

Chi-square tests (categorical variables) and analysis of variances (continuous variables) were used to investigate potential differences in baseline characteristics between included and excluded participants. Kaplan-Meier survival estimates and log-rank tests were used for visual interpretation of the associations between spectral entropy (divided into tertiles) and mortality risk.

Hazard ratios (HRs) and 95% confidence interval (CIs) were determined using Cox-regression models to assess the association between sleep parameters and all-cause mortality. Arbitrary cut-offs for continuous variables were omitted in favor of restricted cubic spline transformations better suited to non-linearity. Thus, HRs for continuous variables were used to compare the 5th and 95th percentiles to that of the 50^{th} percentile by using the 50^{th} percentile as the reference. Proportional hazard assumptions for each variable were tested using Schoenfeld residuals (Schoenfeld, 1982).

Multivariate imputation by chained equations (Azur et al., 2011, Buuren and Groothuis-Oudshoorn, 2011) was used to generate five complete datasets to account for missing variables, with a predictive mean-matching imputation model used for continuous variables, logistic regression for binary variables and

polytomous logistic regression for categorical variables. Coefficients and standard errors of the 5 complete datasets were pooled using Rubin's rules (Rubin, 1987). Findings are reported for pooled HRs and CIs. Confounders were identified based on previous literature (Punjabi et al., 2009, Kendzerska et al., 2014a, Melaku et al., 2019) and included in all models. Total sleep time and the percentage of time spent in REM sleep were included as confounding factors, since an index of sleep quality was hypothesized to have additive information.

The association between sleep disruption, as defined with the spectral entropy of the delta density function, and all-cause mortality risk (primary outcome) was assessed in (Model 1). Seven additional models (secondary outcomes) were separately constructed using the arousal index (Model 2), wake after sleep onset (Model 3), and mean delta, theta, alpha, sigma and beta frequency bands in NREM (Model 4, 5, 6, 7 and 8) to compare the new metric with more conventional sleep and quantitative EEG analysis methods. Interactions between predictor variables and age, BMI, sex, AHI, percent of time with oxygen saturation less than 90%, total sleep time and time spent in REM sleep were also examined. When a significant interaction between sleep fragmentation and a continuous variable was observed, the continuous variable was transformed into quartiles for easier interpretation. Sex stratified and age stratified (>70 and <70 years old) models were also investigated, as reported previously (Punjabi et al., 2009). Predictive performance of the model was assessed using the Harell C-index and Somers' D indices (Newson, 2010) corrected for optimism using bootstrapping. The models were compared to a model containing only the confounder variables using a likelihood ratio test.

4.2.7 Sensitivity analysis

In addition to the main analyses, four sensitivity analyses were performed to further test and validate the findings. First, participants with CVD at baseline (Sensitivity analysis 1; S1) were excluded. Second, participants who died in the first three years were removed (S2), to account for unidentified acute terminal illnesses that might have disrupted sleep (e.g. cancer). Third, the EEG quality was elevated to at least 75% of artefact-free EEG signals (S3). Finally (S4), the analysis was on C4 EEG channel, to help test if the shape of the delta-density function on C3 and C4 were similar.

4.3 Results

4.3.1 Baseline characteristics

The analysis cohort included 5096 participants (88.0% of the recruited participants), from which 2470 (48.4%) participants took part in the follow-up visit. Six hundred and eighty (11.7%) and 177 (6.7%) participants did not meet the EEG quality criteria in the first and second polysomnography visit, respectively. An additional 28 (0.4%) participants had missing data on the primary outcome variable. Compared to included participants, those excluded had a higher AHI and percent of total sleep time spent with oxygen saturation less than 90% (Table 4-1). Included participants also showed a lower proportion of CVD and hypertension. Alcohol consumption (N=407, 7.9%), the SF36 (general health: N=441, 8.7%; physical activity: N=428, 8.4%) and educational status (N=470, 9.2%) were the top 3 missing variables.

There were a total of 69,943 person-years with a mean follow-up of 10.9 years included in the analysis. 1,124 participants died during the follow-up period (men: 601, women: 523). The crude mortality rate for all participants was 16.1 (95% CI: 15.1,17.1) per 1000 person-years.

Table 4-1: Baseline characteristics of study participants.

		Included	Excluded	p-value
N		5096	690	
Demographics				
Age, y		63 (11)	64 (11)	0.001
Sex	Female	2686 (52.7%)	338 (49.7%)	0.152
Years spent in education				
	< 10 years	383 (8.2%)	50 (7.9%)	0.95
	11 - 15	2419 (51.7%)	321 (51.0%)	
	16 - 20	1677 (35.9%)	233 (37.0%)	
	> 20	197 (4.2%)	26 (4.1%)	
Behavioral factors				
Alcohol consumption*, drinks/day		0 [0, 3]	0 [0, 3]	0.42
Smoking Status				
	Never	2413 (47.7%)	281 (47.7%)	0.03
	Former	2178 (43.0%)	309 (45.8%)	
	Current	470 (9.3%)	84 (12.5%)	
Physical Activity		3.13 (1.41)	2.89 (1.55)	<0.001
General health		2.92 (0.76)	2.85 (0.78)	0.03
Anthropometric				
BMI, kg/m²		28.1 (5.0)	28.7 (5.7)	0.004
Medical history				
Hypertension	Yes	2147 (42.1%)	319 (46.9%)	0.02
CVD	Yes	856 (16.8%)	141 (20.7%)	0.02
Diabetes	Yes	347 (7.1%)	58 (9.1%)	0.09
PSG variables				
AHI		9.9 (12.8)	12.4 (18.4)	<0.001
% TST SpO_2 < 90%		3.2 (9.5)	5.6 (15.0)	<0.001
Time Asleep, h		6.04 (1.04)	5.66 (1.21)	<0.001
Time N1, %		5.4 (3.9)	6.2 (4.5)	<0.001
Time N2, %		56.2 (11.5)	59.8 (12.9)	<0.001
Time N3, %		18.5 (11.8)	15.2 (11.8)	<0.001
Time REM, %		19.9 (6.1)	18.8 (7.3)	<0.001
Arousal Index, /h		19.4 (13.6)	19.1 (10.3)	0.59

Data are reported in Mean (SD) for continuous variables and n (%) for categorical variables.

BMI, body mass index; CVD, cardiovascular disease; AHI, Apnea hypo-apnea index.

* data reported as median [IQR]

4.3.2 Sleep disruption and all-cause mortality

Kaplan-Meier survival curves (Figure B1) showed a lower survival probability for participants within the lowest and highest tertile of spectral entropy (compared to the middle tertile). In the multivariate Cox-regression models, a low spectral entropy remained associated with an increase in mortality risk (Figure 4-2). The magnitude of this effect was similar to a 2.25h reduction (6.5h to 4.25h) in total sleep time.

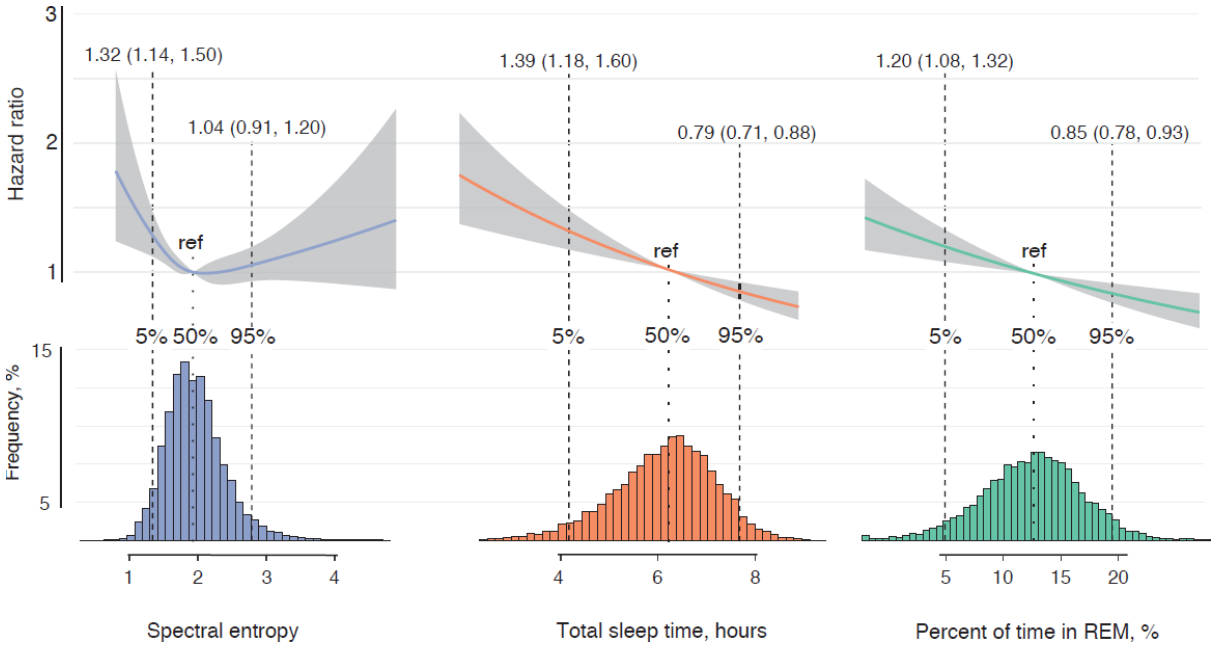

Figure 4-2: Summary hazard ratio for the associations between sleep-related variables (spectral entropy, total sleep time, and percentage of total sleep time spent in REM sleep) and mortality-risk. Hazard ratios (95% confidence interval) compare the 50th to 5th and 95th percentiles of the population.

The interaction between total sleep time and spectral entropy was also significant (χ^2=13.0, p <0.001). Low spectral entropy was associated with increased mortality risk for those sleeping more than 5.5h (Figure 4-3) while a higher spectral entropy was associated with increased mortality risk in those sleeping less than 5.5h. However, the interaction effect was not significant when a more conservative EEG quality cut-off was used (sensitivity analysis 3). The model was well validated (all-optimism < 2%) and well specified (C-index = 0.80, D

= 0.59). The estimated association between spectral entropy and all-cause mortality remained consistent in sensitivity analyses (Table B1).

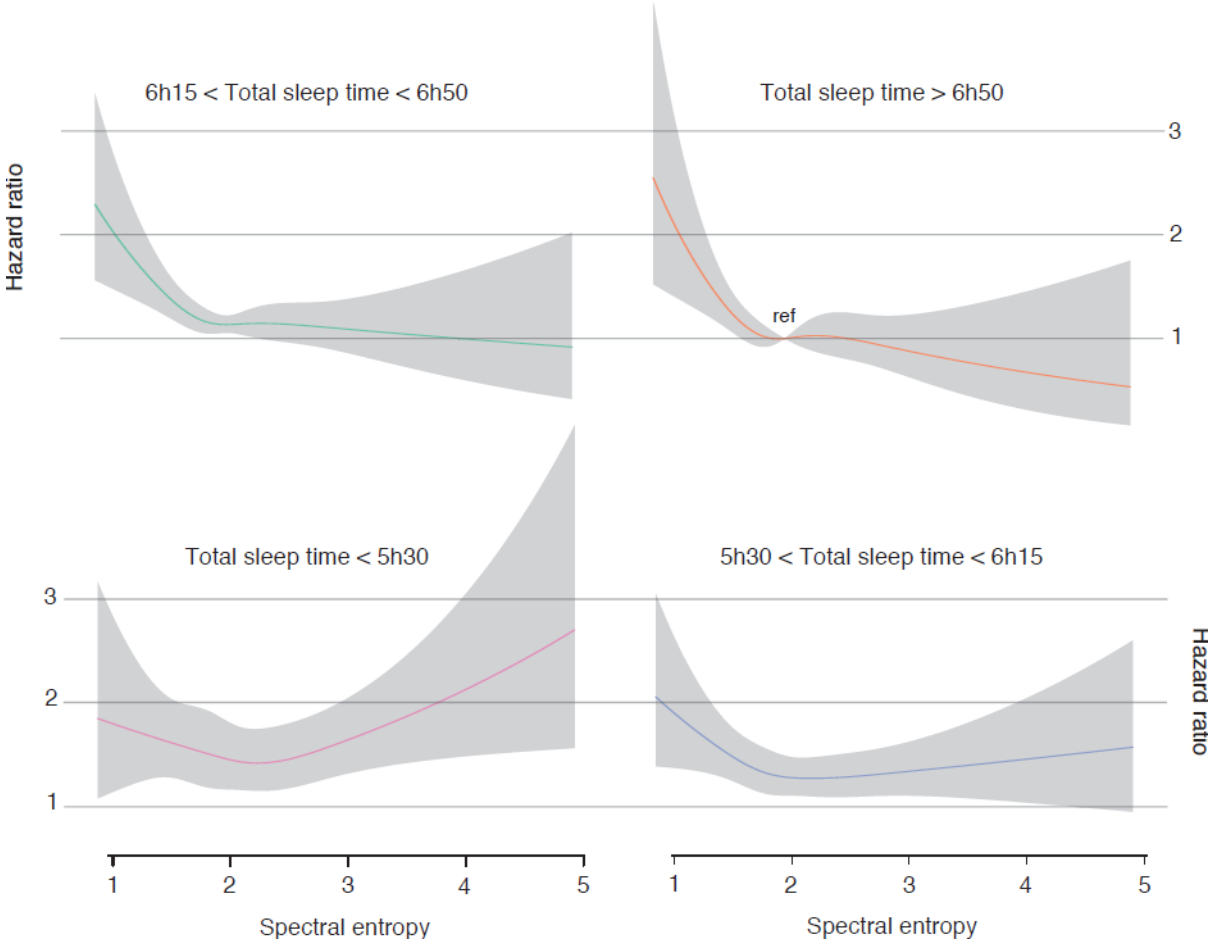

Figure 4-3: Association between spectral entropy and all-cause mortality risk for quartile-based (with interaction term) subgroup analysis of total sleep time models.

Sex-stratified and age-stratified associations between low spectral entropy (5th vs. 50th percentiles) and all-cause mortality risk are shown in Figure 4-4 and Table B2. The association was similar although slightly stronger (p=0.02) for females (HR, 1.40; 95% CI, 1.17,1.70) than males (HR, 1.24; 95% CI, 1.00,1.53) especially for individuals aged less than 70 years old (Table B2). There were no significant associations between high spectral entropy and all-cause mortality.

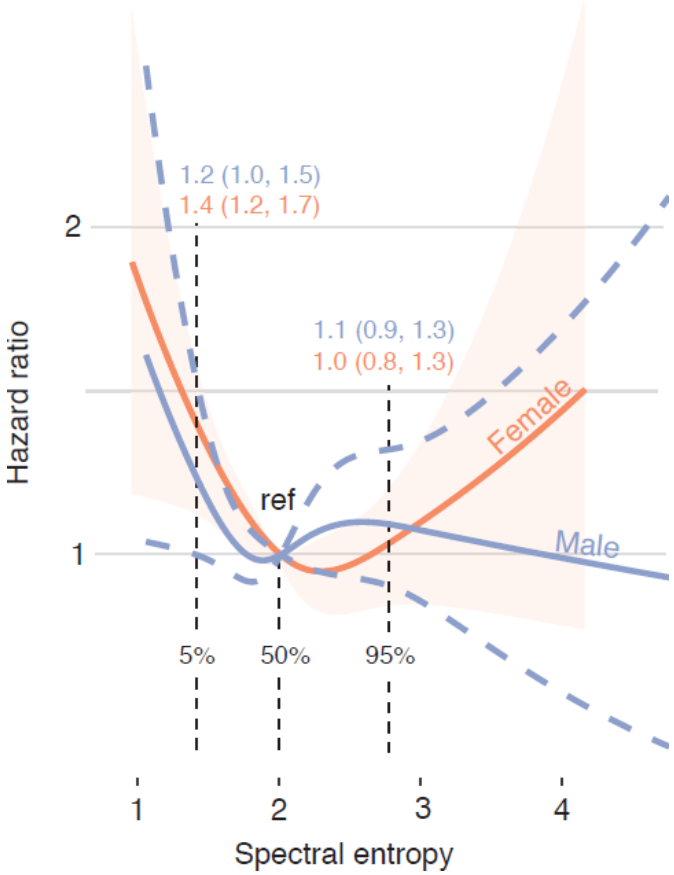

Figure 4-4: Association between spectral entropy and all-cause mortality risk for sex-stratified models. Dashed lines and shaded area represent 95% confidence intervals for male (blue) and female (red), respectively. Hazard ratios (95% confidence interval) compare the 50th to 5th and 95th percentiles of the population.

The fit for the model without the spectral entropy was significantly worse as shown by a likelihood ratio test (Table 4-2) indicating that sleep disruption, defined through the shape of the delta power density function overnight, increased the model fit. Wake after sleep onset, arousal index and the mean power of delta, theta, alpha, sigma and beta frequency bands were not associated with an increase or decrease in mortality-risk (Figure B4 and Table B3). Furthermore, the predictive model fit with traditional (Table 4-2) metrics was significantly worse than a model with spectral entropy, as shown with a lower C- and D-index and a higher Akaike information criterion.

Table 4-2: Likelihood ratio test comparing the predictive performance of the fully adjusted models with spectral entropy and different markers of sleep quality. A lower Akaike information criterion (AIC) and a higher C- and D-index indicates better model fit. The likelihood ratio test tests if the difference in model fit is significant.

	Model	AIC	C-index	D-index	Likelihood ratio test	
					χ^2	p-value
	Core* + spectral entropy[#]	17269.3	0.798	0.597	N/A	N/A
vs	Core	17273.9	0.797	0.594	22.8	0.006
vs	Core + wake after sleep onset	17272.6	0.797	0.595	19.5	0.01
vs	Core + arousal index	17274.4	0.797	0.594	21.4	0.006

*Core is a model containing age, sex, % of total sleep time with less than 90% of oxygen saturation, apnoea-hypopnoea index, Diabetes, SF36 raw physical score, SF36 general health score, % of time spent in REM, total sleep time, alcohol intake, educational status, smoking status, hypertension, cardiovascular disease.
[#]Both interaction terms (spectral entropy x TST and spectral entropy x % of total sleep time with less than 90% of oxygen) were present in the model.

4.3.3 Sleep disruption and sleep breathing disorder

AHI was not associated with mortality-risk (Figure B2) and there was no interaction between spectral entropy and AHI (χ^2=1.9, p=0.16). However, the percentage of total sleep time with oxygen saturation less than 90% was associated with an increase in mortality-risk (Figure B2, 18.2% vs. 0.2 %; 1.24 [1.13, 1.36]). In addition to the total sleep time-spectral entropy interaction previously described, the interaction between spectral entropy and the percent of time spent with oxygen saturation less than 90% was also significant (χ^2=12.80, p=0.0017). The interaction effect (using quartile of % of sleep time spent with oxygen saturation less than 90%) is shown in Figure B3. The interaction remained significant in all sensitivity analyses. In the highest quartile of percentage of sleep time spent with oxygen saturation less than 90%, the association was weaker than in the lowest quartiles.

4.4 Discussion

This Chapter used a large open-access study cohort to examine the association between sleep disruption and all-cause mortality, using both traditional objective sleep quality assessments and more advanced EEG power-spectrum and entropy-based markers of sleep disturbance. The results suggest that delta wave fluctuations were predictive of all-cause mortality independently of other relevant clinical covariates. In contrast, more conventional sleep metrics, such as wake after sleep onset, arousal index and average power across EEG frequency bands were poorly predictive of mortality risk.

These findings provide support that power-spectral analysis of EEG-frequency bands contains more sensitive and useful information than current manually derived measures of sleep. Changes in mean EEG power have previously been associated with increased severity of OSA (Appleton et al., 2019) and insomnia (Krystal et al., 2002). However, in this Chapter, the mean power in any given frequency band in REM/NREM was not predictive of all-cause mortality. This supports a higher value for metrics designed to encapsulate core features of sleep homeostatic processes, particularly in the low frequency band power distribution across the night. Fluctuations in absolute band power of other frequencies (e.g. alpha, sigma, theta and beta) were beyond the scope of this Chapter and remain to be studied in further studies. An immediate challenge is that representing such complex interactions between time and EEG delta frequency distribution overnight with a single number results in substantial information loss that may undermine relationships between sleep quality and relevant clinical and health outcomes. While in this Chapter fluctuations in delta-power overnight provided a robust metric more predictive of all-cause mortality than mean overnight delta power and other covariates, this clearly does not rule out that any number of alternative metrics may be potentially more informative.

Slow wave activity during NREM sleep is thought to reflect synchronous low-level activity of large neuronal populations and to be a fundamental marker of sleep homeostasis and stability (Nir et al., 2011). While previous research has primarily focused on the role of slow oscillations in memory consolidation (Stickgold, 2005, Marshall et al., 2006, Rasch et al., 2007, Maingret et al., 2016). some evidence supports that slow wave sleep is also involved in systemic

metabolic regulation and tissue growth and repair. For example, the suppression of slow wave sleep has been associated with decreased insulin levels, suggesting a potential role of slow wave sleep in modulating glucose regulation (Tasali et al., 2008). Reduced slow wave sleep has also been associated with an increase in the protein β-amyloid, a potential factor in the development of Alzheimer's disease (Ju et al., 2017). Furthermore, growing evidence suggests that slow waves are involved in cardiovascular regulation (Javaheri and Redline, 2012, Silvani and Dampney, 2013). Brindle et al. (2018) showed that the quantity of slow wave sleep may moderate the effect of cardiovascular reactivity on carotid intima-media thickness, a sub-clinical marker of cardiovascular disease. Slow wave sleep duration is also associated with new cases of hypertension (Javaheri et al., 2018). Multiple associations between slow wave sleep and a wide range of cardio-metabolic outcomes support the concept that delta wave disruption may contribute to increased mortality risk. The lack of association between high spectral entropy and all-cause mortality is consistent with studies showing no association with the arousal index . High spectral entropy was also associated with the arousal index and wake after sleep onset in this study. It is possible that with sufficient slow wave activity (no low entropy) and time spent in NREM and REM sleep that recurrent arousals/awakening (high entropy) are not necessarily associated with all-cause mortality.

Methodological considerations

Several limitations of the current study warrant consideration. Firstly, the SHHS began almost 25 years ago, and it is likely that factors not captured within the available data and those that change over time also contribute to all-cause mortality. For example, socio-economic factors appear to be stronger predictors of all-cause mortality than behavioural characteristics, metabolic and chronic conditions, medication and health service utilization (Melaku et al., 2019). Socio-economic factors were not captured in the current analyses. Thus, reproduction of findings in other cohorts, better adjusting for these factors is clearly warranted. Secondly, the SHHS participants were pooled from existing trials examining the effect of sleep-disordered breathing on cardiovascular complications. Thus, participation and survival biases might contribute to these findings. Furthermore, although a recent meta-analysis found no consistent evidence to increased mortality risk with insomnia (Lovato and Lack,

2019), insomnia could potentially also play an important role in helping to explain mortality effects, through sleep and delta wave disruption effects or more extended overnight periods of wake. Replication in unselected populations without sleep disorders (and/or with insomnia) would help to clarify the role of sleep disruption in all-cause mortality. Thirdly, the metric is a single marker of slow wave sleep disruption and refinements using different criteria and cut-offs may increase predictive utility, particularly where replication in large independent samples is clearly required. Furthermore, although confounders were chosen based on recent published analysis (Punjabi et al., 2009, Kendzerska et al., 2014a, Melaku et al., 2019), as in any cross-sectional study the possibility of uncontrolled confounders inevitably remains. Finally, these results rely on a single night of home-based polysomnography, where first night effects and night-to-night variability may influence the distribution of delta wave sleep across the night. These effects remain to be studied but appear unlikely to explain associations with all-cause mortality.

The reported association between all-cause mortality and AHI differs from the initial report from the SHHS (Punjabi et al., 2009), most likely reflecting multiple methodological differences. Firstly, this Chapter contained 1000 fewer participants than the initial report, since 637 participants were not available through National Sleep Research Resources, and some participants were excluded due to poor EEG quality not compatible with power-spectral analysis. Secondly, models contained both percentage of sleep time spent with oxygen saturation less than 90% and AHI, similar to a more recent study (Kendzerska et al., 2014a). However, results are in accordance with a more recent prospective study of 10,000 participants (Kendzerska et al., 2014a) where the percentage of sleep time spent with oxygen saturation less than 90%, but not AHI, was associated with an increased mortality risk. Similar to other recent findings, an association between all-cause mortality and multiple sleep symptoms, such as percentage of time spent in REM (Zhang et al., 2019a, Leary et al., 2020) and total sleep time (Kendzerska et al., 2014a); or as defined by Mazzotti et al. (2019), the disturbed sleep group (although the outcome was cardiovascular event incidence instead of mortality) was also found.

Conclusion

This Chapter supports the concept that sleep disruption contributes to mortality and other adverse health outcomes through multiple pathways. These may include reduced REM sleep, sleep disruption, cardiovascular system stress through exaggerated intra-thoracic pressure swings, increased sympathetic nervous system activity, episodic hypoxia and sleep disruption impacts on daytime functioning, accident risks and health. Finally, upon further validation of this metric, and ideally in combination with other informative metrics, these new tools could be implemented within standard clinical sleep medicine software tools to quantify sleep disruption, underlying pathophysiological mechanisms and clinical outcome risks to better inform targeted treatment and management decisions for sleep problems.

Open source software

The code developed for the presented analysis is written in Python 3 (Python Software Foundation, https://www.python.org/). This code will be available under a common license rule at https://github.com/Adelaide-Institute-for-Sleep-Health/ upon publication of this Chapter; and also include some other analysis published recently (Lechat et al., 2020, Scott et al., 2020). To facilitate uptake a user-interface (see Figure B5) was developed using Qt for python and Plotly (Plotly Technologies Inc., 2015) for interactive visualization.

CHAPTER 5. A NOVEL EEG MARKER PREDICTS PERCEIVED SLEEPINESS AND POOR SLEEP QUALITY

This Chapter was under review at the time of thesis submission.

Citation: Lechat B, Hirotsu C, Appleton S,Younes M, Adams R.J, Vakulin A, Hansen K, Zajamsek B, Wittert G, Catcheside P, Heinzer R, Eckert D.J. A Novel EEG marker predicts perceived sleepiness and poor sleep quality. 2021 [Sleep, submitted 23/06/2021, ID: SLEEP-2021-0526]

Author Contributions: study design—BL, CH, RH, PC and DJE.; data collection— CH, SA, RH, RJA and GW ;analysis of data— BL, MY; interpretation of findings— BL, CH, SA, MY, RJA, AV, KH, BZ, GW, PC, RH; preparation of initial manuscript— BL; revision of final manuscript— BL, KH, YAM, AV, GM, RJA, SA, DJE, PC, BZ.

Summary

Study Objectives: To determine if a novel EEG-derived continuous index of sleep depth/alertness, the odds ratio product (ORP), predicts self-reported daytime sleepiness and poor sleep quality compared to traditional polysomnography metrics in two large population-based cohorts.

Methods: ORP, generated using continuous three second intervals of EEG, ranges from 0 (very deep sleep) to 2.5 (fully alert). $ORP_{wake,}$ reported to be high in hyperarousal states, and the difference between ORP_{wake} and ORP_{NREM} (ΔORP) were measured in the HypnoLaus cohort (N = 2162: 1106 females, 1056 males). ORP_{wake} was also quantified in the MAILES cohort (N = 754 males) as a validation dataset. Logistic regression, controlled for age, body mass index, sex, total sleep time, apnoea/hypopnea index and depression was used to examine associations between ORP_{wake}, ΔORP and traditional polysomnography measures with excessive sleepiness (ESS > 10) and poor sleep quality (PSQI > 5) and insomnia sleep symptoms.

Results: Hyperarousal (high ORP_{wake}) was associated with a ~30% increase in poor sleep quality in both HypnoLaus (odds ratio, OR, and 95% CI) 1.28 (1.08, 1.51) and MAILES 1.36 (1.11, 1.68). High ORP_{wake} was also associated with a ~20% decrease in excessive daytime sleepiness in the combined dataset OR 0.81 (0.69, 0.95). High ΔORP was associated with a decrease in perceived poor sleep quality and excessive daytime sleepiness in HypnoLaus. No associations were detected using traditional polysomnography markers of sleep quality.

Conclusions: ORP, a novel EEG-derived metric, predicts perceived sleepiness and poor sleep quality whereas traditional polysomnographic metrics do not. ORP may also provide unique insight into physiological hyperarousal and the propensity for insomnia.

5.1 Introduction

Excessive daytime sleepiness and poor sleep quality are key markers of sleep difficulties (Balkin et al., 2008) and sleep disorders (Zeman et al., 2004, Slater and Steier, 2012, Sforza et al., 2015). Poor sleep and sleepiness are bidirectionally associated with multiple adverse health outcomes including depression, obesity, anxiety and increased risk of psychiatric conditions and motor vehicle accidents (Johns, 1993, Johns, 1994, Lyznicki et al., 1998, Backhaus et al., 2002, Grandner et al., 2006, Elwood et al., 2011, Tsapanou et al., 2015).

While alternate measures have been proposed (Adams et al., 2016), perceived sleep quality and sleepiness are commonly measured using the Pittsburgh sleep quality index (Buysse et al., 1989) and the Epworth sleepiness scale (Johns, 1991), respectively. PSQI and ESS provide complementary information about healthy and restorative sleep and daytime consequences (Buysse et al., 2008). However, relationships between subjective sleep quality and/or sleepiness and objectives markers of sleep quality derived from traditional gold-standard polysomnography are weak and inconsistent (Buysse et al., 2008, Sforza et al., 2015, Adams et al., 2016).

Poor association between clinical outcomes and polysomnography measures of sleep time and fragmentation may reflect the rather arbitrary and subjective nature of electroencephalography (EEG)-based sleep scoring rules. Traditional polysomnography metrics, which have relatively poor inter- and intra-scorer reliability (Ruehland et al., 2011, Ruehland et al., 2015), were developed around practical constraints of paper-based recordings, rather than underlying pathophysiological processes more likely to be related to clinical outcomes. More recently, the odds ratio product (ORP), a novel EEG-derived metric, has been developed to provide a continuous index of sleep depth and alertness (Younes et al., 2015, Younes, 2017). Following sleep deprivation and sleep restriction, the ORP decreases (deeper sleep). ORP increases (indicating a more alert state) as spontaneous breathing periods resume in patients in the intensive care unit (Dres et al., 2019). Given inconsistent findings between traditional objective sleep measures and subjective sleep quality, this Chapter aimed to determine if ORP-based metrics predict self-reported daytime sleepiness and poor sleep quality.

5.2 Methods

5.2.1 Study design

Perceived sleep quality, sleepiness and polysomnography data were acquired from two community-based cohort studies, the HypnoLaus cohort (2162 participants) and the MAILES cohort (754 participants) (Grant et al., 2014, Heinzer et al., 2015). Signal processing and statistical modelling used the HypnoLaus cohort as a development sample and the MAILES cohort as a validation sample. Further analyses were conducted using both datasets combined.

5.2.2 Cohort overview

Cohort data collection included multiple clinical assessments and a home-based polysomnography (level 2) between 2009 and 2012 for HypnoLaus and 2002 and 2011 for MAILES. Home-based polysomnography was conducted with multiple EEG leads (C3, C4, F3, F4, O1, O2 referenced to M1/2) in the HypnoLaus cohort (Titanium, Embla Flaga, Reykjavik, Iceland) and with a single EEG lead (F3-M2) in the MAILES cohort (Embletta X100, Embla Systems, Thornton, CO, USA). Data from both cohorts included electrooculogram (EOG), chin electromyography (EMG), nasal pressure, thoracic and abdominal effort bands, oximetry and body position signals. Manual scoring of sleep stage and apnoeic events were performed using the 2007 American Academy of Sleep Medicine alternative criteria (Iber et al., 2007) in both cohorts. A more detailed overview of the cohort protocols is described in their respective main papers (Grant et al., 2014, Heinzer et al., 2015). All data were collected as part of research protocols that were approved by the local institutional human research ethics committees of the coordinating institutions. Informed written consent was obtained from each individual prior to participation.

5.2.3 EEG processing

ORP values were generated in non-overlapping three second epochs according to previously published methodology (Younes et al., 2015). Fast Fourier transform was performed in each epoch and the total power in four frequency ranges was calculated: 0.33-2.33 Hz (slow delta), 2.67-6.33 Hz (range 2, includes theta

and fast delta), 7.0-14.0 Hz (alpha/sigma) and 14.0-35.0 Hz (beta). The probability of being awake in each three second epoch based on these absolute power metrics was then derived using a look-up table as described previously (Younes et al., 2015). This probability was then divided by 40 to give an ORP value ranging from 0 (deep sleep) to 2.5 (fully awake/alert). ORP values correlate well with the visual appearance of the EEG across the night and there is an excellent correlation between average ORP in 30 seconds epochs and the probability of an arousal or an awakening occurring in the next 30 seconds epoch (Younes et al., 2015; Younes, 2017; Younes et al., 2020). ORP values from C3 and C4 were averaged in the HypnoLaus study, while ORP in the only available electrode (F3) is reported for the MAILES participants. EEG frequency changes are large and vary across the cortex during sleep. Noise reduction strategies applied averaging of C3 and C4 in the HypnoLaus cohort but were not possible or comparable in the MAILES cohort. Accordingly, ORP analysis was limited to wakefulness only in the MAILES validation cohort.

5.2.4 Questionnaires

Anthropometric information, PSQI, ESS, and medication use were recorded at the time of the polysomnography study in both HypnoLaus and MAILES. The Centre for Epidemiologic Studies Depression Scale (Lewinsohn et al., 1997) (CES-D) was used to assess depression symptoms in HypnoLaus also at the time of the polysomnography study. Beck's depression inventory (Lasa et al., 2000) scale and CES-D were used in MAILES and this information was collected between one and three years before the overnight sleep study. For the purpose of the current study (and for uniformity across cohorts), depression was defined as the use of any anti-depressant medication and/or a depression score above the recommended clinical cut-offs for each questionnaire.

5.2.5 Sleep parameters

Associations between sleepiness and objective sleep quality were assessed using traditional polysomnography metrics of sleep quality (wake after sleep onset, total sleep time, apnoea-hypopneas index, arousal index, sleep efficiency and sleep onset latency) as well as ORP based metrics: the mean overnight ORP_{wake}, hypothesized to reflect sleep propensity during wake epochs and reported to

be high in hyperarousal states (Younes and Giannouli, 2020), and the difference between ORP_{wake} and ORP_{NREM} (ΔORP). Low ΔORP reflects a combination of low ORP_{wake} (high sleep propensity during wake epochs) along with high ORP_{NREM} (light sleep). This combination typically occurs in the presence of sleep disrupting influences that interfere with progression to deep sleep.

5.2.6 Study outcome

The primary aims of this Chapter were to investigate potential relationships between traditional and novel polysomnography markers and perceived excessive daytime sleepiness/poor sleep quality as defined by an ESS > 10 and a PSQI > 5, respectively. These cut-offs were selected as they are the most commonly used definitions in clinical settings (Buysse et al., 1989, Johns, 1991, Johns, 1992, Johns, 1994, Backhaus et al., 2002). Furthermore, when an association was detected between poor sleep quality and a polysomnographic predictor, potential relationships between the polysomnographic predictor and each component of the PSQI sub-scale were investigated. Each PSQI component consists of a score ranging from 0 to 4 and encompasses multiple key aspects of a healthy sleep including a) sleep medication intake, b) subjective sleep quality and c) daytime dysfunction.
Finally, potential associations between ORP_{wake} and difficulty in initiating and maintaining sleep (DIMS) symptoms, generally found in people with insomnia were investigated. DIMS was defined as a sleep onset latency > 30 minutes three times a week or more (PSQI question 5a) and/or "wake up in the middle of the night or early mornings" at least 3 times a week.

5.2.7 Statistical analyses

The Chi-square test for categorical variables and Kruskal-Wallis test for continuous variables were used to identify potential differences in baseline characteristics between the HypnoLaus and MAILES cohort.
Logistic regression was used to investigate associations between the ORP_{wake}, ΔORP and traditional polysomnography measures (wake after sleep onset, arousal index and sleep onset latency) with sleepiness and poor sleep quality. For each polysomnography predictor, five models were constructed: 1) an unadjusted model, 2) a model adjusted for age and sex, 3) and 4) were additionally adjusted for BMI, and total sleep time and AHI, respectively and 5) a

model that included adjustment for all these covariates as well as depression. Multiple imputation using additive regression, bootstrapping, and predictive mean matching was used to generate 20 complete datasets to account for missing variables. Coefficients and standard errors of the 20 complete datasets were pooled using Rubin's rules (Rubin, 1987). Arbitrary cut-offs for polysomnographic variables were omitted and non-linear associations were tested using restricted cubic spline transformations. Thus, the results are expressed as odds ratio (ORs) and 95% confidence intervals (CIs) comparing the 75th against the 25th percentiles of the population.

In addition to the main analyses, two sensitivity analyses to further validate the findings were conducted. In the first sensitivity analysis, participants who reported using benzodiazepine were removed, while in the second analysis participants with less than 30 minutes of wakefulness data during the overnight sleep study were removed.

5.3 Results

5.3.1 Baseline characteristics

Participants from the HypnoLaus and MAILES cohorts had similar BMI, ESS, PSQI and age range, although the median age was 3 years less in HypnoLaus (Table 5-1). Participants in the MAILES cohort had worse objective sleep quality, as reflected by lower total sleep time, longer sleep onset latency and greater wake after sleep onset (Table 5-1). The proportion of participants diagnosed with depression was higher in HypnoLaus. A total of 685 (37%) and 339 (46%) participants reported poor sleep quality and 261 (13.2%) and 88 (11.8%) participants reported excessive daytime sleepiness in HypnoLaus and MAILES, respectively. Participants in MAILES had a higher proportion of sleep disturbances in the 3rd, 4th and 5th components of the PSQI questionnaire compared to HypnoLaus (Table 5-1). However, participants in MAILES reported lower proportions of sleep medication use (PSQI component 6). Data from a total of 192 (8.9%) and 93 (12.3%) participants were rejected for ORP analysis because of poor EEG quality in HypnoLaus and MAILES, respectively.

Table 5-1: Baseline participant characteristics in HypnoLaus and MAILES. Continuous variables are summarized as median [IQR] and categorical variable as n (%).

	All	HypnoLaus	MAILES	p-values
Demographics				
n	2916	2162	754	
Age (years)	58 [50, 69]	57 [49, 68]	60 [52, 69]	< 0.001
Sex: Male (%)	1810 (62.1)	1056 (48.8)	754 (100.0)	< 0.001
BMI (kg/m^2)	26 [24, 29]	26 [23, 29]	28 [26, 31]	< 0.001
Depression: Yes (%)	347 (13.2)	284 (14.8)	63 (8.8)	< 0.001
Anti-depressant use: Yes (%)	167 (5.8)	129 (6.1)	38 (5.0)	0.322
PSQI	5 [3, 7]	4 [3, 7]	5 [3, 8]	< 0.001
ESS	6 [3, 8]	6 [3, 9]	5 [3, 8]	0.068
PSG parameters				
Total sleep time (min)	396 [352, 438]	404 [357, 446]	377 [339, 413]	< 0.001
AHI (#events/h sleep)	6 [2, 14]	4 [1, 11]	10 [6, 20]	< 0.001
Sleep efficiency (%)	86 [77, 91]	88 [79, 92]	80 [73, 87]	< 0.001
Wake after sleep onset (min)	62 [37, 104]	58 [34, 99]	71 [47, 114]	< 0.001
Sleep onset latency (min)	12 [5, 22]	11 [5, 21]	14 [7, 24]	< 0.001
Arousal index (# events/h sleep)	18 [13, 25]	19 [14, 26]	17 [13, 22]	< 0.001
ORP_{wake}	2.08 [1.93, 2.19]	2.05 [1.90, 2.17]	2.14 [2.04, 2.24]	< 0.001
ΔORP	N/A	1.03 [0.5, 1.19]	N/A	N/A
PSQI sub-components				
Sleep quality: >= 2 *	519 (18.5)	387 (18.8)	132 (17.6)	0.488
Sleep latency: > 30min	594 (21.6)	420 (31.1)	174 (23.2)	0.23
Sleep duration: < 6hours	309 (10.6)	193 (8.9)	116 (15.4)	< 0.001
Sleep efficiency: < 75 %	485 (17.3)	316 (15.4)	169 (22.4)	< 0.001
Sleep disturbances: >= 2 *	822 (30.5)	463 (23.7)	359 (48.6)	< 0.001
Medication: > 1/week	353 (12.7)	298 (14.7)	55 (7.3)	< 0.001
Daytime Dysfunction: >= 2 *	374 (13.6)	284 (14.1)	90 (12.0)	0.163

* these PSQI components consisted of a scale that ranged from 0 to 4.

5.3.2 Predictors of daytime sleepiness and perceived poor sleep quality

Unadjusted models suggest an association between ORP_{wake}, ΔORP, wake after sleep onset (WASO), sleep efficiency and sleep onset latency with excessive daytime sleepiness and/or poor sleep quality in both HypnoLaus and MAILES cohorts (Table 5-2). Associations between polysomnography predictors and perceived sleepiness and poor sleep quality were in the opposite direction. For example, higher sleep efficiency was associated with a 10% decrease in the odds of having perceived poor sleep quality but a ~30% increase of having excessive daytime sleepiness. Positive associations between perceived sleepiness and poor sleep quality were detected for ΔORP in the HypnoLaus cohort. Furthermore, ORP_{wake} was associated with both perceived sleepiness and poor sleep quality in HypnoLaus and MAILES, although the association was inconsistent across outcomes. Specifically, a high ORP_{wake} was associated with perceived poor sleep quality and a decrease in the odds of excessive daytime sleepiness.

Table 5-2: Unadjusted univariate logistic regression association (OR, Mean (95% CI)) between excessive daytime sleepiness, poor sleep quality and polysomnography predictors.

	Epworth sleepiness scale (ESS) > 10		**Pittsburgh sleep quality index (PSQI) > 5**	
	HypnoLaus	MAILES	HypnoLaus	MAILES
Wake after sleep onset	0.73 (0.61, 0.88)	0.95 (0.71, 1.28)	1.12 (1.01, 1.25)	1.43 (1.17, 1.75)
Total sleep time	0.94 (0.80, 1.10)	1.28 (0.96, 1.73)	1.02 (0.91, 1.14)	0.96 (0.80, 1.16)
Sleep onset latency	0.81 (0.71, 0.94)	0.97 (0.78, 1.19)	1.11 (1.04, 1.19)	1.34 (1.16, 1.56)
Sleep efficiency	1.32 (1.11, 1.59)	1.15 (0.84, 1.58)	0.90 (0.80, 1.00)	0.65 (0.52, 0.80)
Arousal index	0.96 (0.83, 1.12)	0.99 (0.76, 1.29)	1.00 (0.90, 1.11)	1.14 (0.96, 1.36)
Apnoea/hypopnoea index	0.98 (0.88, 1.09)	0.99 (0.78, 1.26)	1.00 (0.92, 1.07)	1.07 (0.92, 1.25)
ORP wake	0.78 (0.66, 0.93)	0.74 (0.57, 0.97)	1.38 (1.20, 1.59)	1.39 (1.14, 1.69)
ΔORP	0.80 (0.66, 0.95)	na	0.90 (0.79, 1.03)	na

In the fully adjusted models, a higher ORP_{wake} was consistently associated with poor perceived sleep quality in both HypnoLaus and MAILES (Table 5-3), that is, the probability of having perceived poor sleep quality increased with the ORP_{wake} score (Figure 5-1). In contrast, a higher ORP_{wake} was also consistently associated with a decrease in the odds of being excessively sleepy in MAILES and in the combined dataset (Table 5-3).

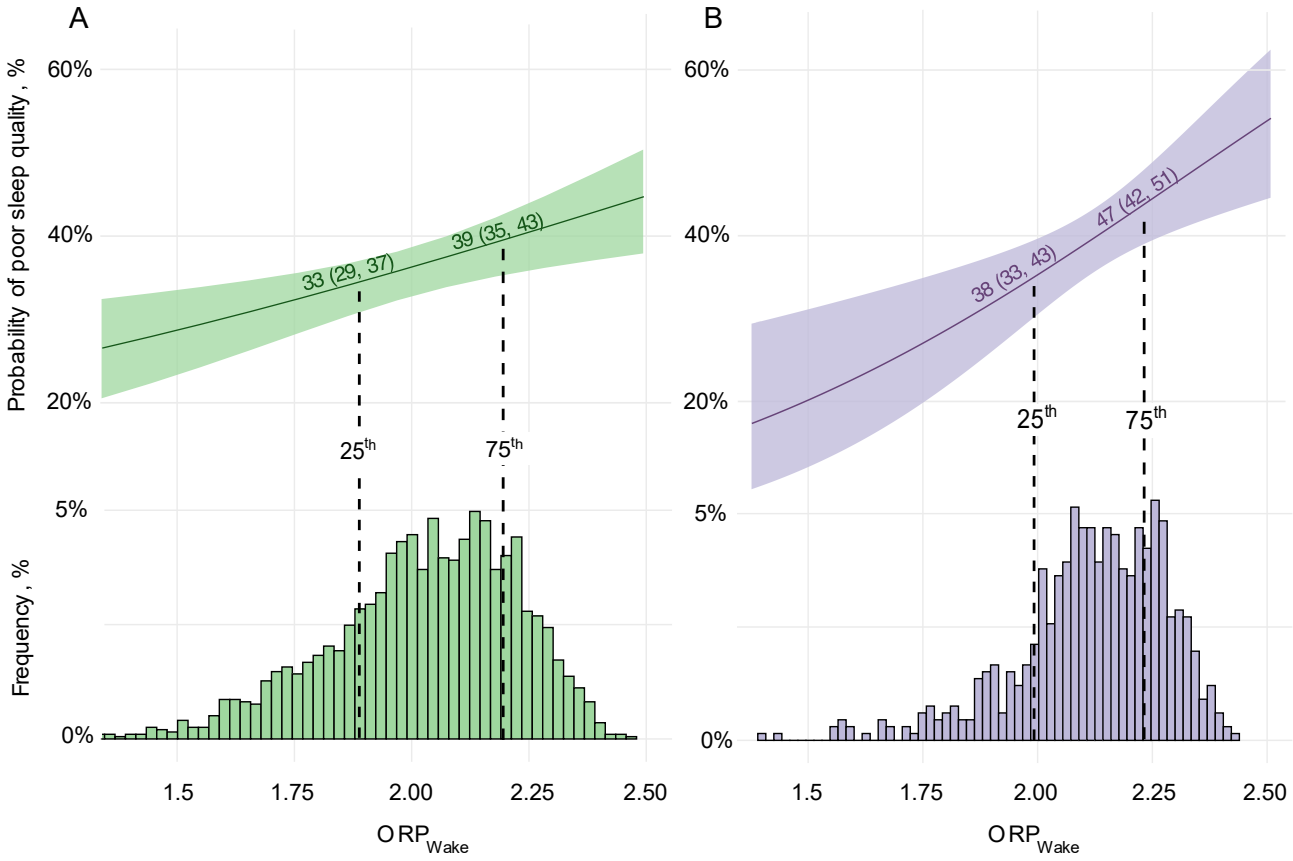

Figure 5-1: Marginal probability of reporting poor sleep quality as a function of ORP_{wake} in (A) HypnoLaus and (B) MAILES.

Estimated associations between ORP_{wake} and excessive daytime sleepiness/perceived poor sleep quality were similar in the pooled versus separate cohort datasets, usually with a narrower confidence interval in the pooled dataset (Table 5-3). There was considerable consistency in observed associations (excessive daytime sleepiness and poor sleep quality) across model 1 to 5 in both cohorts, although the confidence intervals were larger in the latter adjusted models. Furthermore, none of the traditional polysomnography predictors were

consistently associated with perceived poor sleep quality or excessive daytime sleepiness in the fully adjusted models (Table 5-4).

Table 5-3: Adjusted associations (OR, Mean (95% CI)) between subjective measures of sleepiness and poor sleep quality excessive daytime sleepiness with ORP_{wake}.*

	Epworth sleepiness scale (ESS) > 10			**Pittsburgh sleep quality index (PSQI) > 5**		
	HypnoLaus	MAILES	Combined	HypnoLaus	MAILES	Combined
Model 1	0.78 (0.66 – 0.93)	0.74 (0.57 – 0.97)	0.77 (0.66 – 0.89)	1.38 (1.20 – 1.59)	1.39 (1.14 – 1.69)	1.44 (1.28 – 1.61)
Model 2	0.91 (0.76 – 1.09)	0.75 (0.57 – 0.98)	0.84 (0.72 – 0.98)	1.32 (1.13 – 1.53)	1.39 (1.15 – 1.70)	1.43 (1.28 – 1.60)
Model 3	0.92 (0.77 – 1.10)	0.74 (0.56 – 0.97)	0.84 (0.72 – 0.98)	1.34 (1.15 – 1.56)	1.37 (1.12 - 1.67)	1.45 (1.29 – 1.62)
Model 4	0.88 (0.73 – 1.07)	0.75 (0.57 – 0.99)	0.83 (0.71 – 0.97)	1.36 (1.16 – 1.60)	1.39 (1.14 – 1.70)	1.48 (1.31 – 1.66)
Model 5	0.85 (0.70 – 1.04)	0.72 (0.54 – 0.96)	0.81 (0.69 – 0.95)	1.28 (1.08 – 1.51)	1.36 (1.11 – 1.68)	1.42 (1.26 – 1.62)

Model 1: Unadjusted; Model 2: Model 1, age and sex; Model 3: Model 2 and BMI; Model 4: Model 3 and total sleep time, apnoea-hypopneas index; Model 5: Model 4 and depression.
*Odds ratio (ORs) and 95% CIs compare the 75th against the 25th percentiles of the population.

Table 5-4: Multivariable adjusted association between conventional polysomnography markers of sleep quality and excessive daytime sleepiness/perceived poor sleep quality.*

	Epworth sleepiness scale (ESS) > 10			**Pittsburgh sleep quality index (PSQI) > 5**		
	HypnoLaus	MAILES	Combined	HypnoLaus	MAILES	Combined
Wake after sleep onset	0.86 (0.71, 1.04)	1.07 (0.75, 1.51)	0.91 (0.76, 1.08)	1.12 (0.97, 1.29)	1.60 (1.26, 2.03)	1.20 (1.07, 1.35)
Total sleep time	0.88 (0.74, 1.04)	1.21 (0.90, 1.63)	0.96 (0.83, 1.11)	0.93 (0.82, 1.06)	0.92 (0.76, 1.12)	0.93 (0.83, 1.03)
Apnoea-hypopnea index	1.04 (0.92, 1.18)	0.96 (0.75, 1.24)	1.03 (0.92, 1.16)	0.99 (0.91, 1.09)	1.11 (0.94, 1.31)	1.04 (0.95, 1.13)
Arousal index	1.06 (0.88, 1.29)	1.05 (0.75, 1.48)	1.07 (0.91, 1.25)	1.05 (0.91, 1.22)	1.12 (0.89, 1.41)	0.94 (0.84, 1.05)
Sleep efficiency	1.19 (0.97, 1.47)	0.98 (0.64, 1.48)	1.16 (0.95, 1.43)	0.87 (0.74, 1.02)	0.49 (0.37, 0.66)	na[1]
Sleep onset latency	0.80 (0.69, 0.98)	0.98 (0.79, 1.23)	0.83 (0.73, 0.93)	1.02 (0.95,1.10)	1.35 (1.15, 1.57)	na[1]

* Odds ratio (ORs) and 95% CIs compare the 75th against the 25th percentiles of the population. Models are adjusted for age, BMI, sex, total sleep time, apnoea-hypopnea index and depression.

[1] Data cannot be accurately summarised with one odds ratio (95% CI) due to a significant interaction between polysomnography predictors and cohort centre.

A greater ΔORP was associated with decreased odds of excessive daytime sleepiness in models 1-4 in the HypnoLaus cohort (Table 5-5). However, while the point estimate was of similar magnitude, the association was not significant when additionally controlled for depression. A greater ΔORP was also associated with improved perceived sleep quality in some but not all of the models (Table 5-5).

Table 5-5: Adjusted association (OR, Mean (95% CI)) of excessive daytime sleepiness and poor sleep quality with ΔORP for the HypnoLaus cohort.

	Epworth sleepiness scale (ESS) > 10	**Pittsburgh sleep quality index (PSQI) > 5**
Model 1	0.80 (0.66 – 0.95)	0.91 (0.80 – 1.04)
Model 2	0.81 (0.67 – 0.97)	0.86 (0.76 – 0.99)
Model 3	0.82 (0.68 – 0.99)	0.88 (0.76 – 1.00)
Model 4	0.80 (0.66 – 0.97)	0.85 (0.74 – 0.99)
Model 5	0.82 (0.68 – 1.00)	0.93 (0.80 – 1.08)

Model 1: Unadjusted; Model 2: Model 1, age and sex; Model 3: Model 2 and BMI; Model 4: Model 3 and total sleep time, apnoea-hypopneas index; Model 5: Model 4 and depression.

In the fully adjusted models, males were less likely to report poor sleep quality (Hypnolaus, ORs 95% CI, 0.57 (0.46, 0.71)) and depression was a predictor of poor sleep quality (Hypnolaus, 4.10 (3.22, 5.24), MAILES 2.32 (1.43, 3.77)) and sleepiness (HypnoLaus, 1.60 (1.17, 2.19), MAILES 2.25 (1.25, 4.04)). There were no significant interactions between polysomnography predictors and AHI, sex, age and depression. Sensitivity analysis did not change any of the main findings.

5.3.3 PSQI sub-components, insomnia symptoms and hyper-arousal

A higher ORP_{wake} was associated with a 70% increase in consumption of sleeping medications (Table 5-6) in HypnoLaus but not in MAILES, where there was a smaller proportion of medication users (7.5%, N=55) versus HypnoLaus (Table 5-1). Higher ORP_{wake} was also associated with a 24% and 26% odds of reporting high levels of sleep disturbances in HypnoLaus and MAILES respectively. Finally, associations between ORP_{wake} and almost all PSQI sub-components were observed in HypnoLaus and the combined dataset (Table 5-6). Conversely, this was not the case for any of the traditional polysomnography metrics.

Table 5-6: Adjusted association (OR, Mean (95% CI)) between PSQI sub-components, insomnia symptoms and ORP_{wake} in HypnoLaus, MAILES and pooled dataset.

		HypnoLaus	MAILES	Combined
PSQI	Sleep quality	1.16 (0.97 – 1.38)	1.19 (0.90 – 1.57)	1.21 (1.04 – 1.41)
	Sleep latency	1.23 (1.04 – 1.45)	1.24 (0.97 – 1.58)	1.29 (1.12 – 1.48)
	Sleep duration	1.26 (0.99 – 1.61)	1.03 (0.78 – 1.35)	1.26 (1.05 – 1.53)
	Habitual sleep efficiency	1.13 (0.93 – 1.36)	1.23 (0.96 – 1.58)	1.25 (1.07 – 1.47)
	Sleep disturbances	1.24 (1.05 – 1.45)	1.26 (1.03 – 1.54)	1.47 (1.29 – 1.67)
	Sleep medications	1.70 (1.37 – 2.12)	1.31 (0.87 – 1.96)	1.60 (1.31 – 1.94)
	Daytime dysfunction	0.77 (0.67 – 0.98)	1.08 (0.78 – 1.48)	0.87 (0.74 – 1.03)
Insomnia symptoms	SOL > 30 minutes 3/ times a week	1.56 (1.23 – 1.98)	0.92 (0.70 – 1.22)	1.39 (1.15 – 1.68)
	Frequent awakening/early awakening 3 times/week	1.13 (0.97 – 1.32)	1.08 (0.89 – 1.33)	1.25 (1.11 – 1.41)
	Difficulty initiating and maintaining sleep*	1.20 (1.04 – 1.38)	1.08 (0.88 – 1.32)	1.28 (1.14, 1.42)

*defined as sleep onset latency (SOL) > 30 minutes 3/ times a week OR Frequent awakening/early awakening 3 times/week.

In the HypnoLaus cohort, a greater ΔORP was associated with a 17 % reduction in odds of reporting high levels of sleep disturbances and a 31% reduction in odds of taking sleep medications more than once a week. ORP_{wake} was also significantly positively associated (28% increase) with difficulties initiating and maintaining sleep in the combined dataset and in HypnoLaus (20% increase); as well as some of the individual symptoms (Table 5-6).

5.4 Discussion

This Chapter supports that alertness during wakefulness periods during an overnight in-home sleep study, measured using the mean ORP, provides a marker of excessive daytime sleepiness, perceived poor sleep quality and difficulties initiating and maintaining sleep, while traditional polysomnography markers of sleep quality are not.

Predictors of daytime sleepiness and perceived poor sleep quality

Multiple cross-sectional and longitudinal studies have used various traditional polysomnography markers to investigate their potential predictive capacity for sleepiness and sleep quality (Buysse et al., 2008, Adams et al., 2016, Berger et al., 2020). Unfortunately, predictive performance from these traditional metrics has been inconsistent and poor (Buysse et al., 2008, Adams et al., 2016, Berger et al., 2020). Recent studies incorporating more detailed analyses of specific EEG signal components compared to traditional polysomnography have yielded more promising findings. For example, features of KCs are associated with mild cognitive impairment (Liu et al., 2020) and lapses in alertness as measured by the psychomotor vigilance task (Parekh et al., 2019). Similar findings have been reported with sleep spindles (Chatburn et al., 2013) and slow wave sleep (Torsvall and Akerstedt, 1988). These findings, combined with the current study results, support that more detailed polysomnography markers are superior predictors of sleepiness and poor sleep quality than traditional polysomnography markers. Indeed, novel EEG derived metrics such ORP have been used to quantify post-arousal sleep dynamics in patients with OSA (Younes and Hanly, 2016), excessive overnight wake time (Younes and Giannouli, 2020), and other important physiological metrics such as blood pressure (Kim et al., 2020). Thus, there is considerable scope to use more advanced signal processing techniques to identify novel clinically important predictors of disease consequences such as sleepiness which traditional metrics consistently fail to predict.

A potential novel marker of hyperarousal and insomnia

The current findings indicate that a high ORP_{wake} may be a novel biomarker of physiological hyperarousal. The theoretical concept of hyperarousal has gained wide-spread attention as a potential mechanism to explain, at least in part, the pathophysiology of insomnia. Hyperarousal is characterised by 24-h increased cognitive/emotional (e.g. ruminations about sleep, anxiety) (Harvey, 2002, Bonnet and Arand, 2010) and physiological arousal (e.g. heart rate, basal metabolic rate, core body temperature) (Bonnet and Arand, 1997, Nofzinger et al., 2004, Bonnet and Arand, 2010). In addition, beta EEG activity is increased during NREM and/or REM in people with insomnia (Freedman, 1986, Lecci et al., 2020). Recent evidence from the HypnoLaus dataset also suggests

that people with higher frequency EEG activation during sleep tend to underestimate their total sleep time (Lecci et al., 2020). Thus, an association between high ORP, a marker of alertness/sleep depth, and insomnia symptoms is in accordance with these previous findings. However, consistent with an earlier preliminary report (Freedman, 1986), the results also suggest that physiological hyper-arousal is measurable using EEG collected during wakefulness periods that occur throughout sleep. Previous investigation of spectral components of the EEG has only been derived in a small (~50 participants) dataset of people with insomnia rather than two much larger and independent population cohorts. Thus, this Chapter strongly supports the value of a high ORP_{wake} as a novel biomarker of physiological hyperarousal.

Importantly, the ESS assesses propensity of dozing off/falling asleep rather than feeling tired or fatigued and hasn't been associated with insomnia symptoms in previous report on MAILES (Adams et al., 2016). Thus, it is perhaps not surprising that a high $ORP_{wake,}$ marker of hyperarousal state, is associated with insomnia symptoms while a low ORP_{wake}, which measures the ability of dozing off while awake, is associated with excessive daytime sleepiness.

Methodological considerations

While the current study findings are novel and robust given consistent findings and cross validation in two large independent community samples, several limitations warrant consideration. Firstly, the MAILES dataset only included men. Thus, although no significant interactions with sex in the HypnoLaus cohort was found, the findings in women from HypnoLaus require further independent validation. Secondly, only F3 EEG was available from the home polysomnography in the MAILES cohort. Thus, it was not possible to reliably estimate sleep ORP, which requires cross-validation of two EEG sites. Thus, while the difference between NREM ORP and wake ORP and its prediction utility for identifying sleepiness in the HypnoLaus cohort is promising, this marker also needs independent validation in another cohort. Furthermore, while depression was used as a confounder in this analysis, antidepressant medications themselves could also directly or indirectly impact sleep quality and sleepiness. Future work beyond the scope of this thesis is clearly warranted to further examine potential interactive effects between sleep quality, sleepiness, depression, and medication.

Finally, the Epworth sleepiness scale may not be a reliable marker of daytime sleepiness (Chervin et al., 1997, Kendzerska et al., 2014b, Adams et al., 2016), and therefore potential association between ORP_{wake}, traditional marker of sleep quality and daytime sleepiness remained to be further studied using alternative assessment of daytime impairment.

Conclusion

The odds ratio product, a novel EEG-derived metric, predicts important outcomes of perceived sleepiness and poor sleep quality. No associations were observed using traditional polysomnographic metrics. These key findings are consistent across two large community cohorts even after adjustment of key potential confounders. In addition, the odds ratio product may provide unique neurophysiological insight into physiological hyperarousal and the propensity for insomnia.

CHAPTER 6. CO-MORBID INSOMNIA AND OBSTRUCTIVE SLEEP APNOEA IS ASSOCIATED WITH ALL-CAUSE MORTALITY

This Chapter was under review at the time of thesis submission.

Citation: Lechat B, Appleton S, Melaku Y.A, Hansen K, McEvoy R.D, Adams R.J, Catcheside P, Lack L, Eckert D.J, Sweetman A. Co-morbid insomnia and obstructive sleep apnoea is associated with all-cause mortality. [Submitted – ERJ, submitted 13/07/2021, ID: ERJ-01958-2021]

Author Contributions: study design—BL, SA and AS; analysis of data—BL; interpretation of findings— BL, SA,YAM, KH, RDM, RJA, PC, LL, DJE, AS; preparation of initial manuscript— BL; revision of final manuscript BL, SA,YAM, KH, RDM, RJA, PC, LL, DJE, AS.

Summary:

Study Objectives: Increased mortality and cardiovascular disease risks have been examined in people with insomnia and OSA separately. However, insomnia and OSA commonly co-occur and no studies have investigated the effect of COMISA on mortality and cardiovascular event risks. Thus, the aim of this Chapter was to assess the potential association between COMISA and all-cause mortality and cardiovascular event risks.

Methods: Insomnia was defined as difficulty falling asleep, maintaining sleep, and/or early morning awakenings from sleep at least 5 times a month and daytime impairment. OSA was defined as an AHI ≥15 events/h sleep. COMISA was defined if both conditions were present. Multivariable adjusted Cox proportional hazard models were used to determine the association between COMISA and all-cause mortality (n = 1210) and cardiovascular events (N = 1243) over 15 years of follow-up in the Sleep Heart Health Study (n = 5803).

Results: 5236 participants were included in the analysis. 2504 (47.8%) did not have insomnia/OSA, 374 (7.1%) had insomnia-alone, 2027 (38.7%) had OSA-alone, and 331 (6.3%) had COMISA. Compared to participants with no insomnia/OSA, COMISA was associated with a 32% (HR, 95% CI; 1.32 (1.06, 1.64)) and 38% (1.38 (1.11, 1.71)) increased risk of mortality and cardiovascular events, respectively. Insomnia-alone and OSA-alone were not associated with all-cause mortality risk or cardiovascular event risk.

Conclusions: COMISA is associated with increased risk of all-cause mortality and cardiovascular events. These results highlight the need to develop effective treatment approaches for COMISA.

6.1 Introduction

Insomnia and OSA are the two most common sleep disorders, each occurring in approximately 10 to 30% of the general population (Ohayon, 2009, Peppard et al., 2013, Heinzer et al., 2015). Insomnia is characterized by frequent difficulties initiating and/or maintaining sleep, and daytime impairments such as reduced energy, concentration difficulties, and feeling unrested. OSA is characterized by frequent brief narrowing and closure of the upper airway during sleep, resulting in transient reductions in oxygenation, cortical arousals, blood pressure surges and daytime sleepiness and fatigue (Chervin, 2000). Both insomnia and OSA alone contribute to increased risk of future psychiatric and medical conditions, reduced productivity and quality of life, and high healthcare utilization (Peppard et al., 2006, Baglioni et al., 2011, Natsky A., 2020).

Insomnia and OSA often co-occur within the same patient (Sweetman et al., 2017b, Sweetman et al., 2019, Zhang et al., 2019b). COMISA is associated with greater impairment of sleep (Bianchi et al., 2013) and daytime functioning (Krakow et al., 2001), and reduced productivity and quality of life (Bjornsdottir et al., 2012, Lang et al., 2017b), compared to individuals with either insomnia-alone or OSA-alone (Sivertsen et al., 2013, Anttalainen et al., 2019, Sweetman et al., 2019). Individuals with COMISA may also be at increased risk of CVD, compared to people with either disorder alone (Vozoris, 2012, Gupta and Knapp, 2014, Cho et al., 2018). Previous research has examined associations between both insomnia and mortality (Kripke, 2002, Bertisch et al., 2018, Lovato and Lack, 2019), and OSA and mortality (Marshall et al., 2014). However, no general population study has investigated potential associations between COMISA and mortality or cardiovascular event risks.

The SHHS (Quan et al., 1997) is a US-based population cohort study and has considerably advanced knowledge on the potential adverse health outcomes of OSA (Nieto et al., 2000, Shahar et al., 2001). In general, the SHHS has found that OSA may be associated with increased prevalence and incidence of CVD, and all-cause mortality (Shahar et al., 2001, Punjabi et al., 2009, Gottlieb et al., 2010, Redline et al., 2010). Among SHHS participants with OSA (AHI $\geq$ 15 events/h sleep), those with elevated daytime sleepiness were at increased risk of all-cause mortality and cardiovascular disease (Mazzotti et al., 2019).

Similarly, these data have shown that insomnia with objective short sleep duration is associated with increased risk of CVD but not all-cause mortality (Bertisch et al., 2018).

Although COMISA is a common and debilitating condition that is associated with greater morbidity compared to either insomnia-alone or OSA-alone, no general population study has investigated the association of COMISA with all-cause mortality or cardiovascular event risks. The aim of this Chapter was therefore to investigate associations between COMISA and all-cause mortality and cardiovascular event risk in a population-based cohort.

6.2 Methods

6.2.1 Study design and participants

The study design and methodology of the SHHS has been reported previously (Redline et al., 1998) and Chapter 4. Full overnight sleep studies from 6,441 participants were pooled from different population-based studies, of which 5,804 are available through an open access dataset from the National Sleep Research Resource (Dean et al., 2016).

Participants undertook home-based ambulatory polysomnography recordings in 1995- 1998 (Compumedics P Series System; Abbotsford, Victoria, Australia). Polysomnography included two electroencephalograms (EEG) (C4-M1, C3-M2), chin EMG, left and right EOG, ECG, nasal cannula, oro-nasal thermistor, two respiratory band signals (abdominal and thoracic), and finger pulse-oximetry. Sleep and EEG arousals were scored according to the standard criteria at the time (Kales and Rechtschaffen, 1968). Apnoea were scored as a $\geq$ 75% reduction in breathing amplitude lasting at least 10 sec as recorded via the thermocouple signal. Hypopnoeas were identified if the breathing amplitude of the thermocouple or thoracic/abdominal band signals decreased by $\geq$ 30% for at least 10 sec in association with $\geq$ 3% reduction in oxygen saturation or an arousal (Redline et al., 1998). The AHI was defined as the total number of apneas and hypopneas per hour of sleep.

6.2.2 Insomnia, OSA and COMISA

At the time of the polysomnography study, participants completed questionnaires assessing sleep habits and quality of life. Insomnia was defined according

to the presence of self-reported nocturnal sleep difficulties (difficulties falling asleep, waking up in the middle of the night and having difficulty returning to sleep, and/or waking up too early and being unable to resume sleep, at least 5 times per month) *and* daytime impairments including having little to no energy in the past 4 weeks; feeling unrested at least 5 times a month or feeling tired most/all of the time. This definition is similar to the one employed in the American National Health and Nutrition Examination Survey (Hayley et al., 2015). Initially insomnia was defined according to a frequency of at least 15 times per month. However, this definition resulted in a small number of participants in the COMISA group. As this could reduce power to identify a between-group difference in mortality-risk in fully adjusted models, we defined insomnia according to a frequency of at least 5 times per month in the primary analysis, and at least 15 times per month in sensitivity analyses

An AHI of ≥15 events/h sleep was used to define OSA. COMISA was defined if both conditions were present. Participants who did not meet criteria for either insomnia or OSA were categorized with no insomnia/OSA (reference group).

For the all-cause mortality analysis (primary outcome), the potential association between symptomatic OSA, defined as OSA with excessive daytime sleepiness (Epworth sleepiness score > 10) was investigated in supplementary analysis. Furthermore, given the previously published association between hypoxemia and all-cause mortality (Kendzerska et al., 2014a), a supplementary analysis was undertaken to assess potential additive risk of insomnia to hypoxemia (assessed using the % of time spent with less than 90% of oxygen saturation and all-cause mortality). EEG processing

A secondary aim of this Chapter was to investigate sleep fragmentation in COMISA. As such, the EEG processing (spectral entropy and mean NREM absolute powers) was identical to the one developed in Chapter 4. KCs densities (in N2 sleep and N3 sleep) were also calculated using the algorithm described in Chapter 2. Similarly to Chapter 4, participants with less than 50% of artefact free EEG were removed from this analysis.

6.2.3 Potential confounders

Questionnaires determined baseline characteristics including demographics (age, sex, race, educational and marital status), behavioural factors (smoking

status) and body mass index (BMI; kg/m^2). Medical history (hypertension; CVD; chronic obstructive pulmonary disease (COPD); diabetes and medication intake) was determined during an examination no more than five years before the baseline polysomnography study. Medication intake included benzodiazepines, tricyclic anti-depressants and any sleep medication intake more than 5 times a month. Pre-existing CVD cases were determined according to data provided by the parent study cohorts or by self-report at enrolment on the basis of physician reported angina, heart attack, heart failure, stroke, or if the participant ever underwent coronary bypass surgery and/or coronary angioplasty.

6.2.4 Outcome assessment

Death from any cause, up until 2011, was identified in a prior study (Punjabi et al., 2009) using follow-up interviews, written annual questionnaires, telephone contact with study participants or next-of-kin, surveillance of local hospital records and community obituaries and linkage with the Social Security Administration Death Master File. For this analysis, all-cause mortality was used as the primary outcome.

Cardiovascular events were determined by the parent study cohorts according to specific protocols described previously (Gottlieb et al., 2010); and included nonfatal and fatal events. Cardiovascular events occurring before and following baseline were investigated, and included myocardial infarction, myocardial infarction procedure, stroke, angina, coronary heart disease death, congestive heart failure, coronary artery bypass surgery and coronary angioplasty. If multiple cardiovascular events were observed, the closest event following the polysomnography study was retained, and others were disregarded.

6.2.5 Statistical analysis

Confounders and the statistical analyses closely followed previous SHHS reports (Punjabi et al., 2009, Bertisch et al., 2018). Distributions of covariates were summarized by sleep disorder group. Kaplan-Meier survival estimates and log-rank tests were used for visual interpretation of the crude probability of mortality over time. Hazard ratios (HRs) and 95% confidence intervals (CIs) were determined using Cox-regression models to compare risks in sleep disorder groups (insomnia-alone vs. OSA-alone vs. COMISA) and all-cause mortality

(primary outcome) and cardiovascular events (secondary outcome) relative to the reference group (no insomnia/OSA). Proportional hazard assumptions for each variable were tested using Schoenfeld residuals. A similar set of confounders were constructed across four models for all-cause mortality (primary outcome) and cardiovascular events (secondary outcome). The first models were unadjusted. The second models were adjusted for demographics and anthropometrics (age, BMI, race, sex), behavioral (smoking status) and polysomnography-measured total sleep time. The third models were additionally adjusted for pre-existing cardio-metabolic conditions, including diabetes, CVD, hypertension, lipid lowering medications and COPD. Finally, the last models were additionally adjusted for use of benzodiazepines, tricyclic anti-depressants and a binary variable constructed representing participants taking sleep medication more than 5 times a month. Differences between sleep disorder groups in all-cause mortality and cardiovascular event risks were studied using post-hoc comparison.

Two *A-priori* interactions between sleep disorder groups with sex and total sleep time were tested in both outcomes given previous evidence that prevalence of OSA may be under-estimated among women, and that total sleep time may modulate cardiovascular risk in participants with insomnia (Bertisch et al., 2018, Won et al., 2020). For both outcomes, sex-stratified and sex-by-age stratified models were also constructed (shown in APPENDIX C).

Associations between sleep disorder group and specific cardiovascular events (coronary heart disease, heart failure and stroke) were studied in a sub-analysis (APPENDIX C). Coronary heart disease was defined as any myocardial infarction, myocardial infarction procedure, coronary heart disease death, coronary artery bypass surgery and/or coronary angioplasty. Given lower numbers of specific cardiovascular events, only two models were constructed for this analysis. The first model was unadjusted, and the second model was adjusted for age, sex, BMI, race, smoking status, baseline CVD, hypertension and diabetes. Sex-stratified models were also constructed to further investigate potential sex-specific associations with specific cardiovascular events.

6.2.6 Sensitivity analyses

Four sensitivity analyses were conducted. First, for the all-cause mortality analysis, participants who died within the first two years were excluded. Second, for the cardiovascular event analysis, participants with baseline CVD were removed to more specifically examine incident cardiovascular events without potential confounding from prior CVD. Additionally, for both outcomes, the frequency of insomnia symptoms was increased from at least 5 nights per month to at least 15 nights per month. Third, although socioeconomic factors have been shown to influence insomnia and increase mortality and CVD-risk (Melaku et al., 2019), these were not included in primary analyses due to a large amount of missing data. Therefore, the final sensitivity analysis further adjusted for marital status and years of education.

6.3 Results

6.3.1 Baseline characteristics

The analysis sample included 5236 participants (90.2% of 5804) after exclusion of 114 (1.9%), 452 (7.8%) and 2 (< 0.1%) participants due to missing information on nocturnal insomnia symptoms, daytime symptoms, and all-cause mortality, respectively. A complete case dataset (with no missing variables for the fully adjusted model) had 4815 participants (92.0 % of the analysis sample), with baseline cardiovascular disease (N = 202, 3.9%) and diabetes (N = 238, 4.5%) accounting for most missing data. For the cardiovascular event analysis, the sample consisted of 4575 participants, since a further 661 (12.6%) participants were excluded due to missing cardiovascular event data (Gottlieb et al., 2010). Baseline characteristics of the analysis sample are reported in Table 6-1. 14% of all participants with OSA had co-occurring insomnia, and 46.9% of all participants with insomnia had OSA.

Table 6-1: Participant baseline characteristics.

	Overall	Reference	Insomnia	OSA	COMISA
N (%)	5236 (100)	2504 (47.8)	374 (7.1)	2027 (38.7)	331 (6.3)
Demographics					
Age, years	63 (11)	61 (11)	61 (12)	66 (11)	65 (11)
BMI, kg/m²	28 (5)	27 (4)	28 (6)	29 (5)	30 (6)
sex: female	2747 (52.5%)	1581 (63.1%)	293 (78.3%)	727 (35.9%)	146 (44.1%)
Race					
White	4418 (84.4%)	2079 (83.0%)	307 (82.1%)	1759 (86.8%)	273 (82.5%)
Black	463 (8.8%)	238 (9.5%)	28 (7.5%)	159 (7.8%)	38 (11.5%)
Other	355 (6.8%)	187 (7.5%)	39 (10.4%)	109 (5.4%)	20 (6.0%)
Smoking status (%)					
Never	2471 (47.3%)	1235 (49.5%)	183 (49.1%)	895 (44.3%)	158 (48.2%)
Current	497 (9.5%)	275 (11.0%)	55 (14.7%)	137 (6.8%)	30 (9.1%)
Former	2251 (43.1%)	986 (39.5%)	135 (36.2%)	990 (49.0%)	140 (42.7%)
Cardio-metabolic conditions					
Hypertension	2263 (43.2%)	901 (36.0%)	178 (47.6%)	978 (48.2%)	206 (62.2%)
Diabetes	374 (7.5%)	124 (5.2%)	27 (7.8%)	184 (9.4%)	39 (12.3%)
Baseline CVD	855 (16.3%)	313 (12.5%)	58 (15.5%)	390 (19.2%)	94 (28.4%)
Lipid lowering mediation	634 (12.1%)	262 (10.5%)	45 (12.1%)	274 (13.6%)	53 (16.1%)
COPD	59 (1.1%)	27 (1.1%)	9 (2.4%)	16 (0.8%)	7 (2.1%)
Medication intake					
Sleeping pills > 5 times a month	410 (7.8%)	168 (6.7%)	94 (25.2%)	91 (4.5%)	57 (17.3%)
Benzodiazepines	285 (5.5%)	127 (5.1%)	68 (18.2%)	59 (2.9%)	31 (9.4%)
Tricyclic anti-depressants (%)	145 (2.8%)	69 (2.8%)	26 (7.0%)	35 (1.7%)	15 (4.6%)
Sleep related covariates					
AHI, events/hours	14.8 (15.6)	5.1 (3.4)	5.4 (3.6)	26.3 (16.4)	28.2 (17.7)
TST90	3.5 (10.3)	1.3 (6.3)	1.6 (7.5)	6.0 (12.9)	7.6 (14.2)
Total sleep time, min	360 (65)	361 (61)	360 (74)	351 (65)	343 (66)
Wake after sleep onset, min	62 (44)	54 (39)	59 (44)	69 (46)	79 (53)
Sleep efficiency, %	83 (11)	84 (9)	83 (11)	81 (11)	78 (12)
Sleep onset latency, min	14 (20)	14 (20)	16 (22)	14 (20)	16 (23)
Arousal index, events/hours	19 (11)	15 (7)	15 (7)	24 (12)	25 (13)
Outcomes					
Cardiovascular events*	1243 (23%)	475 (19%)	68 (18%)	591 (29%)	109 (33%)
Death	1210 (23%)	486 (19%)	78 (21%)	542 (27%)	104 (31%)

Data are reported as mean (SD) if continuous and n (%) if categorical
AHI = apnoea/hypopnea index/hr, BMI = body mass index, COMISA = co-morbid insomnia and sleep apnoea, CVD = Cardiovascular disease, OSA = obstructive sleep apnoea, TST90 = percent of total sleep time spent with less than 90% of oxygen saturation, COPD = chronic obstructive pulmonary disease * Composite cardiovascular endpoint in a community sample with and without prevalent cardiovascular disease

6.3.2 **COMISA and sleep fragmentation** [a]

A further 622 participants were removed due to inadequate EEG quality needed for this analysis. Sleep fragmentation as a result of insomnia alone, OSA alone and COMISA compared to the control group was studied using linear regression controlled for age, BMI and sex. These results are presented in Table 6-2 and showed that COMISA participants tended to have more fragmented sleep than insomnia alone or OSA alone participants. Specifically, COMISA participants showed higher wake after sleep onset, lower total sleep time and higher spectral entropy (Table 6-2). All sleep disorders were also with lower KC density in N3 sleep.

[a] Section 6.3.2 was not included in the manuscript under review.

Table 6-2: Mean (95% CI) of sleep quality metrics in sleep disorder group. The values represent the difference between the populations without insomnia and without OSA.

	Sleep disorder			**p-values**	
	Insomnia	OSA	COMISA	COMISA vs OSA	COMISA vs insomnia
Wake after sleep onset, min	7.1 * (2.5, 11.7)	8.0 (5.3, 10.7)	19.5 (14.6, 24.4)	< 0.001	<0.001
Total sleep time, min	-12.4 (-19.3, -5.6)	-7.9 (-11.9, -3.9)	-17.9 (-25.2, -10.5)	0.007	0.122
Sleep latency , min	2.8 (-0.4, 5.9)	1.1 (-0.7, 2.9)	3.9 (0.1, 4.9)	0.094	0.603
Arousal Index, event/hours	0.5 (-0.6, 1.5)	8.5 (7.9, 9.2)	9.0 (7.8,10.1)	0.074	<0.001
NREM delta [a], μv^2/Hz	-0.02 (-0.08,0.04)	-0.04 (-0.07, -0.01)	-0.02 (-0.08,0.03)	0.713	0.692
NREM theta [a], μv^2/Hz	0.05 (0.0, 0.11)	-0.04 (-0.07, -0.01)	-0.001 (-0.07, 0.05)	0.873	0.643
NREM alpha [a], μv^2/Hz	0.08 (0.02, 0.14)	-0.03 (-0.06, -0.01)	-0.001 (-0.07, 0.06)	0.877	0.722
NREM sigma [a], μv^2/Hz	0.05 (-0.02, 0.11)	0.02 (0.0, 0.07)	-0.03 (-0.08, 0.04)	0.130	0.146
NREM beta [a], μv^2/Hz	0.03 (-0.03, 0.09)	0.10 (0.07, 0.14)	0.07 (0.00, 0.14)	0.286	0.347
Spectral entropy	0.01 (-0.04, 0.06)	0.04 (0.01, 0.07)	0.1 (0.04, 0.16)	0.023	0.006
KC density in N2, events/min	-0.05 (-0.13,0.02)	0.01 (-0.03, 0.05)	0.03 (-0.5, 0.11)	0.562	0.097
KC density in N3, events/min	-0.24 (-0.41, -0.08)	-0.12 (-0.21, -0.02)	-0.19 (-0.37, -0.02)	0.195	0.825

[a] Values were log-transformed

6.3.3 All-cause mortality

The median (IQR) follow-up period for all-cause mortality was 11.8 (10.4, 15.9) years, over which there was a total 1210 deaths (21.1% of the analysis sample). The crude mortality rates were 17.3, 19.3, 24.9 and 30.4 events per 1000 person-years for the reference, insomnia-alone, OSA-alone and COMISA groups, respectively. Kaplan-Meier curves are shown in Figure 6-1A and suggest that

participants with COMISA had a lower survival probability than those with insomnia-alone ($p = 0.001$) and OSA-alone ($p = 0.047$) alone.

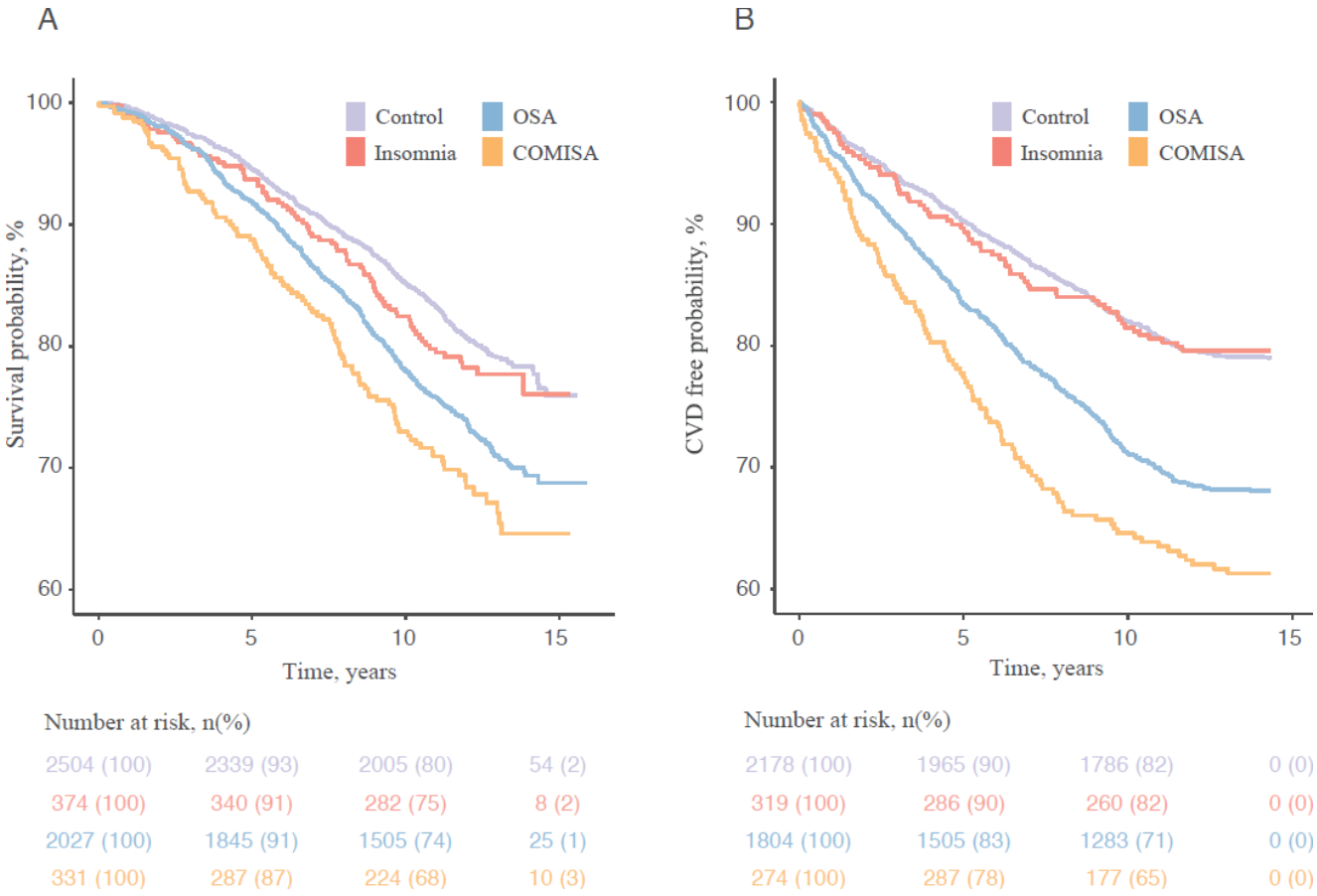

Figure 6-1: Unadjusted Kaplan-Meier's curve across sleep disorder categories for (A) All-cause mortality and (B) cardiovascular disease incidence. For both outcomes, log-rank test p-values<0.001. OSA = Obstructive sleep apnoea; COMISA = co-morbid insomnia and sleep apnoea.

COMISA was associated with an 80% increase in all-cause mortality risk in the unadjusted model, which was higher than either OSA-alone (46% increase) or insomnia-alone (12% increase; Table 6-3). After adjusting for all pre-specified covariates (Table 6-3), COMISA was associated with an increase in all-cause mortality compared to the reference group, and OSA-alone group ($p = 0.014$), but not the insomnia-alone group ($p = 0.50$). In the fully adjusted model, the interaction of sleep disorder category and sex on mortality approached significance ($p = 0.051$), while the interaction of sleep disorder category and total sleep time was not significant ($p = 0.57$). Sex-stratified and age-stratified models are reported in the Supplement (Table C1 and Table C2). The association of COMISA and mortality was strongest in older males (> 70

years). Excessive daytime sleepiness was present in 25.9% of participant with OSA-alone. However, symptomatic OSA was not associated with all-cause mortality (HR, 95%CI; 1.02 (0.83, 1.26)). Association between tertiles of % of time spent with less than 90% of oxygen saturation and insomnia with all-cause mortality is shown in Figure C1. A moderate to severe degree of hypoxemia (second and third tertiles) was associated with all-cause mortality when insomnia symptoms were also present. However, moderate hypoxemia alone was not associated with all-cause mortality

Table 6-3: Adjusted associations between sleep disorder groups and all-cause mortality.

	N	N event	Insomnia	OSA	COMISA
Model 1	5236	1210	1.12 (0.88, 1.43)	1.46 (1.29, 1.65)	1.80 (1.46, 2.22)
Model 2	5189	1198	1.28 (1.00, 1.62)	1.01 (0.89, 1.15)	1.37 (1.10, 1.70)
Model 3	4822	1152	1.20 (0.94, 1.54)	1.00 (0.88, 1.14)	1.32 (1.06, 1.64)
Model 4	4815	1150	1.19 (0.92, 1.53)	1.01 (0.88, 1.15)	1.32 (1.06, 1.64)

Quoted values are Hazard ratio (and 95% CI) against the reference group. OSA = Obstructive sleep apnoea; COMISA = co-morbid insomnia and sleep apnoea.
Model 1: Unadjusted ; Model 2: Age, BMI, sex, race, smoking status, total sleep time; Model 3: Model 2 AND obstructive pulmonary disease, cardiovascular disease, hypertension, diabetes, and lipid medication intake; Model 4: Model 3 AND benzodiazepines and tricyclic anti-depressants and sleep medication intake more than 5 times a month.

The association between COMISA and all-cause mortality did not change after excluding 96 (1.8% of sample) participants who died within the first two years of follow-up (Table C3). As expected, increasing the insomnia symptom threshold from at least 5 times a month to at least 15 times a month, substantially decreased the prevalence of both insomnia-alone (3.2%) and COMISA (2.6%). However, according to this definition, participants with COMISA had a 49% (HR, 95% CI; 1.49 (1.08, 2.07)) increase in all-cause mortality in the fully adjusted model (Table C4). Further adjustment for marital status and educational status did not change the association between COMISA and all-cause mortality (N = 4466, N event = 1126; 1.29 (1.03, 1.61)).

6.3.4 Cardiovascular disease

The median (IQR) follow-up period for detecting incident cardiovascular events was 11.4 (7.6, 12.7) years. Of the 4575 participants in the sample, 1243 unique events were recorded during follow-up. The unadjusted cardiovascular event incidence rate was 17.6, 17.3, 29.0 and 38.2 events per 1000 person-years for the reference, insomnia-alone, OSA-alone and COMISA groups, respectively. CVD-free probability for unadjusted Kaplan-Meier curves are shown in Figure 6-1. In the fully adjusted model, participants with COMISA had a 34% increase in cardiovascular event risk compared to those with no insomnia/OSA, but no difference in compared to those with insomnia-alone ($p = 0.37$) or OSA-alone ($p = 0.14$) (Table 6-4). Compared to the reference group, neither insomnia-alone, nor OSA-alone were associated with higher risk of cardiovascular event incidence in the fully adjusted model. The interactions between sleep disorder category and both sex ($p = 0.53$), and total sleep time ($p = 0.64$) were not significant in the fully adjusted model. Results of stratified models between sleep disorder group, age, and sex with cardiovascular events are reported in Table C5 and Table C6. Similar to the mortality analysis, the strongest association of COMISA and cardiovascular event risk was observed among older males (> 70 years). Associations between specific cardiovascular event types, sleep disorder group and sex are reported in Table C7.

Table 6-4: Adjusted associations between sleep disorder group and cardiovascular event.

	N	N event	Insomnia	OSA	COMISA
Model 1	4575	1243	0.98 (0.77, 1.27)	1.62 (1.44, 1.83)	2.12 (1.72, 2.61)
Model 2	4547	1239	1.06 (0.82, 1.37)	1.09 (0.96, 1.24)	1.48 (1.20, 1.83)
Model 3	4421	1216	1.00 (0.77, 1.30)	1.09 (0.96, 1.24)	1.38 (1.11, 1.70)
Model 4	4415	1213	0.99 (0.76, 1.29)	1.09 (0.96, 1.24)	1.38 (1.11, 1.71)

Quoted values are Hazards ratio (and 95% CI) against the reference group. OSA = Obstructive sleep apnoea; COMISA = co-morbid insomnia and sleep apnoea.
Model 1: Unadjusted; Model 2: Age, BMI, sex, race, smoking status, total sleep time; Model 3: Model 2 AND obstructive pulmonary disease, cardiovascular disease (list all), hypertension, diabetes, and lipid medication intake; Model 4: Model 3 AND benzodiazepines and tricyclic anti-depressants and sleep medication intake more than 5 times a months.

Of the 4575 participants included in the cardiovascular event analysis, 792 (17.3%) had prevalent cardiovascular disease at baseline. When these participants were removed, the number of incident cardiovascular events dropped from 1243 to 794. Subsequently, the association between COMISA and incident cardiovascular events remained significant (and of similar magnitude to the analysis retaining these participants) in Model 1 and 2, but not in Model 3 and 4 (Table C8). When the frequency of insomnia symptoms was increased to at least 15 times per month (Table C9), the association between COMISA and cardiovascular events remained significant, except for the fully adjusted model (HR, 95% CI 1.35 (0.97, 1.87)). Further adjustment for marital status and educational status did not change the association between COMISA and cardiovascular events (N = 4066, N event = 1174; 1.33 (1.07, 1.66)).

6.4 Discussion

The main findings of this Chapter are that co-morbid insomnia and sleep apnea (COMISA) may be associated with increased risk of sleep fragmentation, all-cause mortality, and cardiovascular events, compared to individuals without insomnia or OSA. Furthermore, no associations with all-cause mortality or cardiovascular events were observed for participants who had either OSA-alone or insomnia-alone.

Results are consistent with previous research demonstrating that COMISA is associated with worse physical and mental health compared to either insomnia, or OSA-alone (Bianchi et al., 2013, Lang et al., 2017b, Sweetman et al., 2019). This is the first study to demonstrate that COMISA is also associated with increased risk of all-cause mortality, compared to patients with no insomnia/OSA. The high prevalence, morbidity, and mortality risk associated with COMISA highlight the importance of developing more effective treatment approaches (Sweetman et al., 2017b, Sweetman et al., 2019, Zhang et al., 2019b).

Although no previous study has investigated the association of COMISA with all-cause mortality, reports from the Sleep Heart Health Study and other cohort studies have shown decreased quality of life (Baldwin et al., 2001), and increased cardiovascular event/all-cause mortality risk in specific sub-groups of people with insomnia or OSA (Gottlieb et al., 2010, Bertisch et al., 2018, Mazzotti et al., 2019). The findings are also consistent with previous SHHS reports investigating specific OSA sub-groups. For example, Butler reported increased all-cause mortality in participants with short-duration apnea events, which may be a marker of the low respiratory arousal threshold phenotype (Butler et al., 2019). Indeed, the 'low arousal threshold' OSA phenotype may share several common characteristics with the "hyper-arousal" model of chronic insomnia which postulates that insomnia is maintained by increased physiological and psychological arousal during the day/night (Bonnet and Arand, 2010). This could be consistent with our results suggesting that patients with COMISA have higher spectral entropy than controls. Furthermore, COMISA patients had fewer K-complexes and higher mean absolute beta-power during NREM sleep than controls. The potential additive effect of disrupted sleep to mortality risk in COMISA patient remains unclear, especially given that low (but not high) spectral entropy was associated with all-cause mortality in Chapter 4. Future research should investigate the potential contribution of a low arousal threshold, sleep disruption and "hyper-arousal–like" EEG on all-cause mortality risk in patients with COMISA.

The associations between all-cause mortality and cardiovascular event risk and COMISA highlight the need for further research to investigate the mechanisms underpinning this relationship. It is possible that the association of COMISA with mortality and cardiovascular event risk may result from the high prevalence of additional medical/psychiatric co-morbidities and increased morbidity in participants with COMISA (Yang et al., 2011, Bjornsdottir et al., 2012, Lang et al., 2017b), more complex interactions between the psychological and physiological mechanisms and manifestations of each disorder that may exacerbate the other and trigger worse physical and mental health, or other conditions/prodromal symptoms which were not identified or controlled within the available data. Given previous associations between all-cause mortality with both insomnia-alone, or OSA-alone (Punjabi et al., 2009, Vgontzas et al., 2010,

Kendzerska et al., 2014a), the most parsimonious explanation may be an additive effect of insomnia and OSA on all-cause mortality risk among participants with COMISA.

Although people with COMISA experience more substantial impairment to sleep, including reduced objective sleep duration (Bianchi et al., 2013), compared to those with either insomnia- or OSA-alone (Sweetman et al., 2017b), COMISA was associated with all-cause mortality and cardiovascular event risk independently of total sleep time. The interaction of total sleep time and sleep-disorder group on mortality was also investigated but no significant moderation effect was observed. However, previous studies investigating the association of insomnia with short sleep duration and mortality have reported mixed results (Bertisch et al., 2018, Lovato and Lack, 2019). Sleepiness associated with OSA could also influence and potentially extend sleep time to somewhat mask short sleep duration effects of insomnia in some participants. This potential confounding effect of COMISA may partly explain the increased risk of all-cause mortality and cardiovascular event incidence independently of total sleep time. The clinical presentation, manifestations and pathophysiology of OSA and response to treatment differs in women compared to men (Won et al., 2020). Earlier SHHS reports suggest that men with OSA are at higher risk of all-cause mortality (i,e, younger men) (Punjabi et al., 2009), incident coronary heart disease, heart failure (Gottlieb et al., 2010) and stroke (Redline et al., 2010). Similarly, in this Chapter, little evidence of increased all-cause mortality and cardiovascular event risk in women was observed. Consequently, when a fixed AHI threshold is used to define OSA, women with "OSA" are more likely to be undiagnosed, untreated, and experience treatment failure (Gagnadoux et al., 2016, Appleton et al., 2018). Thus, the lack of association could reflect under-diagnosis and/or misdiagnosis of OSA in women.

Patients with COMISA likely require tailored treatment approaches (Sweetman et al., 2017a, Sweetman et al., 2019). For example, patients with OSA and co-morbid insomnia have lower acceptance and poorer compliance with continuous positive airway pressure, compared to patients with OSA-alone (Sweetman et al., 2017b). Recent randomized controlled trials suggest that cognitive behavioral therapy for insomnia before commencing continuous positive airway pressure may increase adherence (Sweetman et al., 2019).

Given that no improvements in secondary cardiovascular event risks were observed following continuous positive airway pressure therapy for participants with OSA in randomized controlled trial data (McEvoy et al., 2016, Peker et al., 2016, Sánchez-de-la-Torre et al., 2020), a more specific investigation of the effects of combined therapy on markers of cardiovascular disease severity and risks may be warranted among OSA patients with co-occurring insomnia who may be at increased risk of cardiovascular events and mortality.

Several limitations of the current study warrant consideration. Firstly, although both nocturnal and daytime insomnia symptoms were used to classify insomnia, as recommended in diagnostic criteria, it was not possible to identify patients with 'chronic insomnia' persisting for ≥3 months. Furthermore, the frequency of insomnia symptoms used in this study (> 5/month) is lower than diagnostic criteria (> 3/week). However, given sensitivity analyses supported robust associations in the smaller group of patients with more frequent insomnia complaints it is possible that the association between COMISA and mortality may be even stronger if insomnia were defined according to *chronic* nocturnal and daytime symptoms. Secondly, socio-economic factors appear to be very strong predictors of all-cause mortality (Melaku et al., 2019), sleepiness (Adams et al., 2016), and insomnia (Talala et al., 2012). However, these covariates were either not available and/or not reliably collected. Thus, these results warrant confirmation in other cohorts, and ideally in RCTs, with more comprehensive socio-economic data, and with more clearly defined insomnia symptom chronicity. Finally, although diagnosed mental health condition data were not available, mental health symptoms was controlled for using anti-depressant medication use which will underestimate the burden of mental health conditions and residual confounding may be present. Future research should examine the association of COMISA and mortality, controlling for treated and untreated mental health symptoms and doctor-diagnosed mental health conditions.

Conclusion

In summary, this Chapter found that participants with co-morbid insomnia and sleep apnea may have decreased longevity and increased cardiovascular event risks compared to participants with no insomnia or OSA. It remains to be determined if these associations are causal and treatment with CBTi,

CPAP, or combination treatment can effectively decrease mortality and/or cardiovascular event risks in individuals with COMISA.

CHAPTER 7. CONCLUSION

This thesis described the development of automated methods and biomarkers of sleep fragmentation. These biomarkers were subsequently tested for clinical utility in several population groups relevant to sleep fragmentation, including a sample of individuals exposed to environmental noise and a large population sample including participants with sleep disorders. The work presented in this thesis suggests that EEG-based biomarkers designed to encapsulate core physiological and pathophysiological processes of sleep and sleep disorders are more informative than traditional manually scored sleep metrics and are likely to be more important and informative predictors of adverse health outcomes.

7.1 Summary of findings

A sophisticated KC detection algorithm was developed in this thesis and its performance was shown to be superior to existing algorithms (Chapter 2). While the higher F1-score (0.78 vs 0.6 in the literature) is high, the main strength and advantage of this new algorithm over previous approaches resides in the probabilistic scoring. The algorithm gives a probabilistic score to each KC, with a high probability indicative of larger and well-defined KCs. To the best of the author's knowledge, this is the first attempt to design an algorithm with intuitive probabilistic outcome scoring of KCs. Using this algorithm, dose-response relationships between environmental noise sound pressure level (and type) and KC response was investigated in Chapter 3. These results suggested that KCs occurred at sound pressure levels well below (as low as 33 dBA) the threshold needed to elicit arousals and awakenings. For the same sound pressure level, a noise stimulus was also twice as likely to evoke a KC, further suggesting that KCs are a more sensitive marker of sensory processing during sleep. Remarkably strong interactions between subjective noise-sensitivity and KC-response rates were also observed, with an almost two-fold reduction in KC-response occurrence in self-reported noise-sensitive participants. However, there were no corresponding changes in arousal rates. These findings support an important role of KCs in noise-sensory processing during sleep and the value of KC-responses as an objective marker of environmental noise effects on sleep.

Further work, substantially facilitated through the KC development and testing work presented in this thesis, is clearly needed to better understand the impact of more frequent KCs with noise exposure during sleep.

A novel marker of sleep quality/fragmentation based on the distribution of slow wave activity across the night was found to be predictive of all-cause mortality in the Sleep Heart Health Study data set (Chapter 4). Slow wave activity was quantified using a multi-taper based fast Fourier transform with signal to noise ratio advantages over conventional quantitative EEG methods. Quantifying the complexity of delta activity across the night using spectral entropy is particularly novel and sensible given that multiple biological processes are likely to be dependent on (such as the glymphatic system) and involved in the regulation of slow oscillations during sleep. Using this metric, it was found that delta wave fragmentation during sleep is associated with a 30% increase in all-cause mortality independent of OSA, total sleep time and traditional potential confounders. Although it is not possible to infer causal relationships from longitudinal associations alone, the finding of significant associations between EEG slow wave activity and mortality, in combination with biological plausibility of clearly important functions of slow wave activity during sleep, support the higher value of slow wave activity based metrics compared to traditional sleep metrics alone. These findings are also important in the context of respiratory sleep disorders such as OSA given that most treatments for OSA have focused on reducing the apnoea-hypopnea index and hypoxia with little regard for treatment effects on sleep quality *per se*. This suggests that metrics designed to more rigorously encapsulate underlying sleep physiology provide more robust estimators of sleep quality that are more strongly associated with adverse health outcomes compared to traditional methods.

Poor sleep quality and excessive daytime sleepiness were found to be associated with the odds ratio product, a novel EEG marker thought to reflect hyperarousal. While hyperarousal has been theorized to be a common physiological and psychological trait in participants with insomnia, most of the previously available evidence was derived from a small population study. Chapter 5 of this thesis showed that some parts of the population exhibit hyperarousal-like EEG with higher frequencies during the wake period, as measured with the

odds ratio product. High wake ORP was in turn significantly associated with a 30% likelihood of reporting poor sleep quality and a 20% decrease in excessive daytime sleepiness (consistent with an hyper-aroused state). Importantly, sleepiness was assessed as the propensity of dozing off/falling asleep (using the ESS) rather than feeling tired or fatigued. Previous reports suggest that insomnia symptoms are not associated with dozing off/falling asleep and are more strongly associated with feelings of daytime fatigue (Adams et al., 2016). Thus, it is perhaps not surprising that a high ORP during the wake period, indicative of a hyper-aroused state, was associated with insomnia symptoms. On the other hand, a low ORP_{wake}, which signifies the propensity to dose off while awake, was associated with excessive daytime sleepiness. Associations between specific PSQI measures of sleep quality and ORP-based metrics further suggest that these objective measures can successfully capture important information relevant to subjective sleep impairment. Conversely, traditional polysomnography markers of sleep quality were not predictive of excessive daytime sleepiness or poor sleep quality. This, together with findings from previous Chapters clearly suggests that more detailed analyses of specific EEG signal components compared to traditional polysomnography yields more clinically useful findings.

Co-morbid insomnia and OSA is a debilitating condition, and the work in Chapter 6 showed for the first time that patients with co-morbid insomnia and OSA are at higher risk of all-cause mortality, cardiovascular events and sleep fragmentation than patients with one of these disorders alone. These effects were more evident in men than in women. However, a traditional fixed AHI threshold to define OSA regardless of sex ignores known sex-dependent differences in OSA prevalence, risk factors and symptomatology. Thus, women may exhibit health impacts at a lower AHI compared to men, leading to underdiagnosis and/or misdiagnosis of OSA in women. Differential cut-offs and symptom impacts may well help to explain the lack of association in women. Further studies, incorporating more systematic approaches to test for sex-dependent and quite likely age-dependent differences in sleep problems and impacts, are clearly needed to better understand and define effectively treatable sleep disorders that negatively impact the community. The high prevalence,

morbidity, and mortality risk associated with COMISA highlight the importance of developing more effective diagnostic and treatment approaches for which better tailored and more sex-specific approaches remain needed.

7.2 Clinical and research implications

The algorithm and biomarkers developed in Chapter 2 and Chapter 4 have promising clinical utility. These metrics showed significant positive associations in the investigation of both the consequences of sleep fragmentation in the general population and the impact of environmental noise on sleep. Upon further validation of these metrics, ideally in combination with other potentially informative metrics, these new tools could be implemented within standard clinical sleep medicine software tools to quantify sleep disruption. Application of more sensitive and physiologically informative metrics compared to traditional methods are likely to help advance the understanding of underlying pathophysiological mechanisms and clinical outcome risks, and to better inform targeted treatment and management decisions for sleep problems. An interactive user-interface was also developed during this thesis (see section B.5 Open source software), to facilitate the independent use and validation of these new biomarkers by the wider sleep research community.

Another central finding of this thesis likely to influence future research is that the pattern of EEG slow wave distribution across the night follows a distinctive pattern of cyclical reduction across the night. This specific pattern is likely to be influenced by multiple biological processes, such as sleep pressure and circadian rhythms (Lazar et al., 2015). Thus, it is not surprising that deviance from this specific pattern is associated with clinical conditions (as shown with KCs in Chapter 2) or adverse health outcomes like all-cause mortality, as shown in Chapter 4. Previous research on other specific EEG-derived markers of sleep, such as quantitative EEG analysis or ORPs, has predominantly focused on averaged quantities, such as delta power, across sleep episodes and sleep stages with little regard for the more dynamic and cyclical nature of sleep. However, given the findings of relationships between spectral entropy of delta wave activity overnight and mortality, time-dependant markers appear likely to usefully complement these and potentially other sleep measurements.

The results in this thesis demonstrating added risk of adverse health outcomes in participants with multiple sleep disorders are clearly of clinical interest. Given that no improvements in secondary cardiovascular event risk have been demonstrated following CPAP therapy in patients with OSA in randomized controlled trials (McEvoy et al., 2016, Peker et al., 2016, Labarca et al., 2020, Sánchez-de-la-Torre et al., 2020), a more specific investigation of the effects of combined therapy (CBTi + CPAP) on markers of cardiovascular disease severity and risks may be warranted among OSA patients with co-occurring insomnia. Furthermore, given strong evidence to support poorer outcomes in patients with COMISA, patients undergoing clinical diagnosis for either insomnia or OSA alone should be considered for screening for COMISA.
Finally, findings regarding KCs and environmental noise are valuable towards advancing the understanding of environmental noise impacts on sleep, and for potential future evidence-based improvement of environmental noise measurements, regulations and management. Exposure-response curves are already used to inform political decision making designed to help mitigate the effects of environmental noise on sleep (Basner and McGuire, 2018). Given that measurable changes in KC response probabilities to noise occur a SPLs lower than for arousals and awakenings, potentially important noise effects on sleep could occur at lower SPLs than previously assumed in the design of current noise guidelines. Clearly further work is needed to better understand environmental noise impacts on sleep, daytime functioning and longer-term health.

7.3 Limitations

There are inevitably some limitations of this thesis work that warrant consideration and future work. The K-complex algorithm developed during this thesis was based on a small sample of 19 participants, with K-complexes scored by only one scorer, and on only one EEG channel. Therefore, while the impact of a small training dataset on algorithm performance was mitigated using uncertainty quantification, there would likely be value in training a similar algorithm using a larger dataset scored by multiple scorers and a consensus scoring approach. Furthermore, given evidence to support that K-complexes can be localised events within the brain (Mak-McCully et al., 2015), with somewhat

different morphology depending on recording site, further validation using scoring from different EEG locations would likely be useful.

The calculation of "complexity" of the delta wave activity overnight was only studied using spectral entropy. Spectral entropy was chosen based on previous literature suggesting potential clinical utility, but also because it does not require extraneous parameters to be calculated and optimized. However, there are multiple measures of complexity such as entropies, fractals or network analysis amongst several other potential approaches (Bradley and Kantz, 2015, Zou et al., 2019). Therefore, it is possible that one or more of these techniques may be better suited to explain and quantify delta-wave fluctuation overnight than the spectral entropy. Nonetheless, while the clinical utility of the delta-wave fluctuations marker was demonstrated using all-cause mortality, there remains a need to test associations between this and other potential markers and mortality, along with other adverse clinical outcomes, such as sleepiness and other health outcome risks.

The Sleep Heart Health study was used to study the association between the new entropy-based marker of sleep disruption with all-cause mortality. The same dataset was used to study the association between COMISA and all-cause mortality. Use of the SHHS dataset is clearly advantageous to compare with other research since this dataset has been used extensively across a range of previous studies. Furthermore, the SHHS is the only study available to date with multiple hard clinical outcomes, such as all-cause mortality and CV events. However, the results derived in this thesis based on the SHHS dataset remain to be validated in other cohorts with more diverse sleep disorders and clinical symptoms.

This thesis focused on KCs and more generalised slow oscillations during sleep and ignored higher frequency patterns such as sleep spindles and theta waves. Spindles, as well as slow oscillations-spindle coupling, have recently been shown to be sensitive markers of cognitive performance in a large population-based study (Djonlagic et al., 2021). Therefore, in addition to the K-complex and delta-wave based biomarkers developed in this thesis, the additive contribution of spindle-derived metrics warrants further study. Another limitation of this work is the focus on specific frequency components of the EEG. Scale-free component of neural activity (sometimes called "background brain activity" or

"1/f" activity) is a further key component of brain activity and has been recently hypothesized to be a biomarker of arousal level in human sleep (Lendner et al., 2020). In the same study, 1/f activity was higher during REM sleep episodes, further suggestive that important neuronal homeostatic and most likely synaptic reorganisation activity takes place during REM sleep. The findings in Chapter 4 that time spent in REM sleep was also a predictor of all-cause mortality, independent of known covariates further supports this concept. Therefore, further targeted metrics designed to capture key physiological features of both NREM and REM sleep, such as eye movements, theta waves, atonia and "background brain activity", clearly remain warranted to more comprehensively test for relationships between other markers of sleep homeostasis and clinical outcomes.

The impact of environmental noises on the KC response was only studied in a young and healthy population, which is not representative of the broader population habitually exposed to wind farm noise or traffic noise (Pedersen and Waye, 2004, Pedersen and Persson Waye, 2007). As these data were measured as part of a pilot study, the sample size was also relatively small. The effect of different noise types, ideally with a greater number of noise repetitions during the night would also be helpful to elucidate the potential contribution of frequency-specific components of noise in K-complex responses. The long- and short-term health consequences, and the potential impact on next day sleepiness of the absence vs presence of KCs were also not investigated and would benefit from further work.

Most of the data underpinning this thesis work was derived from single night home-based polysomnography studies, where first night effects and night-to-night variability are likely to influence study results. Given the increased cardiovascular risk for people with high irregularity in total sleep time (more than 2 hours difference across 7 days) (Huang et al., 2020); the night-to-night variability of the developed metrics, and their consequences, also remain to be studied. However, more affordable and portable monitoring sleep systems would be needed for such studies, which are prohibitively difficult and expensive with conventional technology.

7.4 Future work recommendations

Based on the work performed in this thesis, there are several recommendations for further research. The distribution of EEG events, such as slow waves, spindles, and KCs across the night should be studied more thoroughly. Given initial results, these micro-EEG elements appear to be distributed according to physiological processes likely to be at least partly related to mechanisms underpinning sleep homeostasis needed for good mental and physical functioning and health. Given the literature on coupling between different physiological signals (see section 1.2.8a), similar techniques, such as network physiology methods (Ivanov et al., 2016), could likely be usefully modified and applied to EEG and other physiological signals to further explore their distributions and relationships with outcomes relevant to mental and physical performance and health. This approach appears highly likely to reveal further important relationships between novel markers of sleep quality and health not currently detected using conventional sleep metrics. Novel approaches clearly require appropriate clinical validation and demonstration of clinical utility, but nevertheless, improved diagnostic and treatment advances would be expected to follow.

The effect of treatments, such as CBTi or CPAP for OSA, on novel EEG biomarkers should also be studied. Such studies would help to definitively establish causal relationships between markers of sleep disturbance and health impacts and help to reveal markers that are treatment responsive versus non-modifiable traits or irreversible signs of damage. Ultimately, sleep fragmentation, at least partly arising from the use of treatments such as CPAP, also warrants study with these techniques, particularly given that CPAP is inherently somewhat uncomfortable and restrictive and generates some noise. Given the higher all-cause mortality and cardiovascular event risk for people with multiple sleep disorder co-occurrence, more research is clearly needed to understand the potential of treatment combinations to improve sleep and reduce adverse health outcome risks.

Lastly, further research is needed to find appropriate methods for combining biomarkers. Sleep disorder pathogenesis is clearly complex and multi-faceted, and sleep fragmentation is only one of several aspects occurring in a range of sleep disorders. Different metrics are likely needed to capture different aspects

of sleep homeostasis and sleep physiological processes. A combination of multiple biomarkers is likely to support effective differential diagnosis, identification of specific underlying causal mechanisms, and to provide optimal guidance towards the most effective and cost-effective treatment approaches. Linear models used in sleep research epidemiology cannot adequately account for high-level interactions, which clearly restricts the number of possible inputs. Thus, machine learning-based methods are likely to be much better suited for such a task (Obermeyer and Emanuel, 2016). Ultimately, current methods for the diagnosis and management of sleep disorders typically involve long waiting times associated with limited access to sleep specialist resources, and lengthy diagnostic and trial and error treatment approaches, often with sub-optimal outcomes. Thus, substantial future advances in understanding the role of sleep in human health, and for sleep disorder diagnosis and management, are likely possible through further application of EEG and other signal processing methods for quantifying sleep quality in more depth than is currently possible through manual sleep scoring methods.

BIBLIOGRAPHY

Aasm 2014. *International Classification of Sleep Disorders (ICSD-3), Diagnostic and coding manual*, Westchester, IL.

Abasolo, D., Hornero, R., Espino, P., Alvarez, D. & Poza, J. (2006). Entropy analysis of the EEG background activity in Alzheimer's disease patients. *Physiol Meas,* 27**,** 241-53.

Adams, R. J., Appleton, S. L., Vakulin, A., Lang, C., Martin, S. A.*, et al.* (2016). Association of daytime sleepiness with obstructive sleep apnoea and comorbidities varies by sleepiness definition in a population cohort of men. *Respirology,* 21**,** 1314-21.

Afifi, L., Guilleminault, C. & Colrain, I. M. (2003). Sleep and respiratory stimulus specific dampening of cortical responsiveness in OSAS. *Respiratory Physiology & Neurobiology,* 136**,** 221-234.

Ageborg Morsing, J., Smith, M. G., Ogren, M., Thorsson, P., Pedersen, E.*, et al.* (2018). Wind Turbine Noise and Sleep: Pilot Studies on the Influence of Noise Characteristics. *Int J Environ Res Public Health,* 15.

Amzica, F. & Steriade, M. (1997a). Cellular substrates and laminar profile of sleep K-complex. *Neuroscience,* 82**,** 671-686.

Amzica, F. & Steriade, M. (1997b). The K-complex: its slow (<1-Hz) rhythmicity and relation to delta waves. *Neurology,* 49**,** 952-9.

Amzica, F. & Steriade, M. (2002). The functional significance of K-complexes. *Sleep Med Rev,* 6**,** 139-49.

Anttalainen, U., Grote, L., Fietze, I., Riha, R., Ryan, S.*, et al.* (2019). Insomnia symptoms combined with nocturnal hypoxia associate with cardiovascular comorbidity in the European sleep apnea cohort (ESADA). *Sleep and Breathing,* 23**,** 805-814.

Appleton, S., Gill, T., Taylor, A., Mcevoy, D., Shi, Z.*, et al.* (2018). Influence of Gender on Associations of Obstructive Sleep Apnea Symptoms with Chronic Conditions and Quality of Life. *Int J Environ Res Public Health,* 15.

Appleton, S. L., Vakulin, A., D'rozario, A., Vincent, A. D., Teare, A.*, et al.* (2019). Quantitative electroencephalography measures in rapid eye movement and nonrapid eye movement sleep are associated with apnea-hypopnea index and nocturnal hypoxemia in men. *Sleep,* 42**,** 1-9.

Appleton, S. L., Vakulin, A., Mcevoy, R. D., Vincent, A., Martin, S. A.*, et al.* (2015). Undiagnosed obstructive sleep apnea is independently associated with reductions in quality of life in middle-aged, but not elderly men of a population cohort. *Sleep Breath,* 19**,** 1309-16.

Azur, M. J., Stuart, E. A., Frangakis, C. & Leaf, P. J. (2011). Multiple imputation by chained equations: what is it and how does it work? *Int J Methods Psychiatr Res,* 20**,** 40-9.

Babadi, B. & Brown, E. N. (2014). A review of multitaper spectral analysis. *IEEE Trans Biomed Eng,* 61**,** 1555-64.

Babisch, W. (2011). Cardiovascular effects of noise. *Noise Health,* 13**,** 201-4.

Backhaus, J., Junghanns, K., Broocks, A., Riemann, D. & Hohagen, F. (2002). Test-retest reliability and validity of the Pittsburgh Sleep Quality Index in primary insomnia. *J Psychosom Res,* 53**,** 737-40.

Baglioni, C., Battagliese, G., Feige, B., Spiegelhalder, K., Nissen, C.*, et al.* (2011). Insomnia as a predictor of depression: a meta-analytic evaluation of longitudinal epidemiological studies. *J Affect Disord,* 135**,** 10-9.

Baldwin, C. M., Griffith, K. A., Nieto, F. J., O'connor, G. T., Walsleben, J. A., *et al.* (2001). The association of sleep-disordered breathing and sleep symptoms with quality of life in the Sleep Heart Health Study. *Sleep,* 24**,** 96-105.

Balkin, T. J., Rupp, T., Picchioni, D. & Wesensten, N. J. (2008). Sleep loss and sleepiness: current issues. *Chest,* 134**,** 653-660.

Bankman, I. N., Sigillito, V. G., Wise, R. A. & Smith, P. L. (1992). Feature-based detection of the K-complex wave in the human electroencephalogram using neural networks. *IEEE Trans Biomed Eng,* 39**,** 1305-10.

Bartsch, R. P., Schumann, A. Y., Kantelhardt, J. W., Penzel, T. & Ivanov, P. (2012). Phase transitions in physiologic coupling. *Proc Natl Acad Sci U S A,* 109**,** 10181-6.

Bashan, A., Bartsch, R. P., Kantelhardt, J. W., Havlin, S. & Ivanov, P. (2012). Network physiology reveals relations between network topology and physiological function. *Nat Commun,* 3**,** 702.

Basner, M., Babisch, W., Davis, A., Brink, M., Clark, C., *et al.* (2014). Auditory and non-auditory effects of noise on health. *The Lancet,* 383**,** 1325-1332.

Basner, M., Glatz, C., Griefahn, B., Penzel, T. & Samel, A. (2008). Aircraft noise: effects on macro- and microstructure of sleep. *Sleep Med,* 9**,** 382-7.

Basner, M. & Mcguire, S. (2018). WHO Environmental Noise Guidelines for the European Region: A Systematic Review on Environmental Noise and Effects on Sleep. *Int J Environ Res Public Health,* 15.

Bastien, C. & Campbell, K. (1992). The evoked K-complex: all-or-none phenomenon? *Sleep,* 15**,** 236-45.

Bastien, C. & Campbell, K. (1994). Effects of rate of tone-pip stimulation on the evoked K-Complex. *J Sleep Res,* 3**,** 65-72.

Bastuji, H., Perrin, F. & Garcia-Larrea, L. (2002). Semantic analysis of auditory input during sleep: studies with event related potentials. *Int J Psychophysiol,* 46**,** 243-55.

Bates, D., Machler, M., Bolker, B. M. & Walker, S. C. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software,* 67**,** 1-48.

Beebe, D. W. & Gozal, D. (2002). Obstructive sleep apnea and the prefrontal cortex: towards a comprehensive model linking nocturnal upper airway obstruction to daytime cognitive and behavioral deficits. *J Sleep Res,* 11**,** 1-16.

Begoli, E., Bhattacharya, T. & Kusnezov, D. (2019). The need for uncertainty quantification in machine-assisted medical decision making. *Nature Machine Intelligence,* 1**,** 20-23.

Belenky, G., Wesensten, N. J., Thorne, D. R., Thomas, M. L., Sing, H. C., *et al.* (2003). Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: a sleep dose-response study. *J Sleep Res,* 12**,** 1-12.

Ben Simon, E. & Walker, M. P. (2018). Sleep loss causes social withdrawal and loneliness. *Nat Commun,* 9**,** 3146.

Benjafield, A. V., Ayas, N. T., Eastwood, P. R., Heinzer, R., Ip, M. S. M., *et al.* (2019). Estimation of the global prevalence and burden of obstructive sleep apnoea: a literature-based analysis. *Lancet Respir Med,* 7**,** 687-698.

Berger, M., Hirotsu, C., Haba-Rubio, J., Betta, M., Bernardi, G., *et al.* (2020). Risk factors of excessive daytime sleepiness in a prospective population-based cohort. *J Sleep Res***,** e13069.

Bernert, R. A., Kim, J. S., Iwata, N. G. & Perlis, M. L. (2015). Sleep disturbances as an evidence-based suicide risk factor. *Curr Psychiatry Rep,* 17**,** 554.

Berry, R. B., Budhiraja, R., Gottlieb, D. J., Gozal, D., Iber, C., *et al.* (2012). Rules for Scoring Respiratory Events in Sleep: Update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events. *Journal of Clinical Sleep Medicine,* 8**,** 597-619.

Bersagliere, A. & Achermann, P. (2010). Slow oscillations in human non-rapid eye movement sleep electroencephalogram: effects of increased sleep pressure. *J Sleep Res,* 19**,** 228-37.

Bertisch, S. M., Pollock, B. D., Mittleman, M. A., Buysse, D. J., Bazzano, L. A.*, et al.* (2018). Insomnia with objective short sleep duration and risk of incident cardiovascular disease and all-cause mortality: Sleep Heart Health Study. *Sleep,* 41.

Bianchi, M. T., Williams, K. L., Mckinney, S. & Ellenbogen, J. M. (2013). The subjective-objective mismatch in sleep perception among those with insomnia and sleep apnea. *J Sleep Res,* 22**,** 557-68.

Bjornsdottir, E., Janson, C., Gislason, T., Sigurdsson, J. F., Pack, A. I.*, et al.* (2012). Insomnia in untreated sleep apnea patients compared to controls. *J Sleep Res,* 21**,** 131-8.

Bonnet, M. H. (1985). Effect of sleep disruption on sleep, performance, and mood. *Sleep,* 8**,** 11-9.

Bonnet, M. H. (1987). Sleep Restoration as a Function of Periodic Awakening, Movement, or Electroencephalographic Change. *Sleep,* 10**,** 364-373.

Bonnet, M. H. & Arand, D. L. (1997). Hyperarousal and insomnia. *Sleep Med Rev,* 1**,** 97-108.

Bonnet, M. H. & Arand, D. L. (1998). Heart rate variability in insomniacs and matched normal sleepers. *Psychosom Med,* 60**,** 610-5.

Bonnet, M. H. & Arand, D. L. (2010). Hyperarousal and insomnia: state of the science. *Sleep Med Rev,* 14**,** 9-15.

Bradley, E. & Kantz, H. (2015). Nonlinear time-series analysis revisited. *Chaos,* 25**,** 097610.

Brandenberger, G., Ehrhart, J., Piquard, F. & Simon, C. (2001). Inverse coupling between ultradian oscillations in delta wave activity and heart rate variability during sleep. *Clinical Neurophysiology,* 112**,** 992-996.

Bremer, G., Smith, J. R. & Karacan, I. (1970). Automatic Detection of the K-Complex in Sleep Electroencephalograms. *IEEE Transactions on Biomedical Engineering,* BME-17**,** 314-323.

Brindle, R. C., Duggan, K. A., Cribbet, M. R., Kline, C. E., Krafty, R. T.*, et al.* (2018). Cardiovascular Stress Reactivity and Carotid Intima-Media Thickness: The Buffering Role of Slow-Wave Sleep. *Psychosom Med,* 80**,** 301-306.

Butler, M. P., Emch, J. T., Rueschman, M., Sands, S. A., Shea, S. A.*, et al.* (2019). Apnea-Hypopnea Event Duration Predicts Mortality in Men and Women in the Sleep Heart Health Study. *Am J Respir Crit Care Med,* 199**,** 903-912.

Buuren, S. V. & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations inR. *Journal of Statistical Software,* 45.

Buysse, D. J., Hall, M. L., Strollo, P. J., Kamarck, T. W., Owens, J.*, et al.* (2008). Relationships Between the Pittsburgh Sleep Quality Index (PSQI), Epworth Sleepiness Scale (ESS), and Clinical/Polysomnographic Measures in a Community Sample. *Journal of Clinical Sleep Medicine,* 4**,** 563-571.

Buysse, D. J., Reynolds, C. F., 3rd, Monk, T. H., Berman, S. R. & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res,* 28**,** 193-213.

Campbell, K. B. & Colrain, I. M. (2002). Event-related potential measures of the inhibition of information processing: II. The sleep onset period. *International Journal of Psychophysiology,* 46**,** 197-214.

Cash, S. S., Halgren, E., Dehghani, N., Rossetti, A. O., Thesen, T.*, et al.* (2009). The human K-complex represents an isolated cortical down-state. *Science,* 324**,** 1084-7.

Catcheside, P. G., Chiong, S. C., Mercer, J., Saunders, N. A. & Mcevoy, R. D. (2002). Noninvasive cardiovascular markers of acoustically induced arousal from non-rapid-eye-movement sleep. *Sleep,* 25**,** 797-804.

Cervena, K., Dauvilliers, Y., Espa, F., Touchon, J., Matousek, M., *et al.* (2004). Effect of cognitive behavioural therapy for insomnia on sleep architecture and sleep EEG power spectra in psychophysiological insomnia. *J Sleep Res,* 13**,** 385-93.

Chambon, S., Galtier, M. N., Arnal, P. J., Wainrib, G. & Gramfort, A. (2018). A Deep Learning Architecture for Temporal Sleep Stage Classification Using Multivariate and Multimodal Time Series. *IEEE Trans Neural Syst Rehabil Eng,* 26**,** 758-769.

Chambon, S., Thorey, V., Arnal, P. J., Mignot, E. & Gramfort, A. (2019). DOSED: A deep learning approach to detect multiple sleep micro-events in EEG signal. *J Neurosci Methods,* 321**,** 64-78.

Chatburn, A., Coussens, S., Lushington, K., Kennedy, D., Baumert, M., *et al.* (2013). Sleep spindle activity and cognitive performance in healthy children. *Sleep,* 36**,** 237-43.

Chervin, R. D. (2000). Sleepiness, fatigue, tiredness, and lack of energy in obstructive sleep apnea. *Chest,* 118**,** 372-9.

Chervin, R. D., Aldrich, M. S., Pickett, R. & Guilleminault, C. (1997). Comparison of the results of the Epworth Sleepiness Scale and the Multiple Sleep Latency Test. *J Psychosom Res,* 42**,** 145-55.

Cho, Y. W., Kim, K. T., Moon, H. J., Korostyshevskiy, V. R., Motamedi, G. K., *et al.* (2018). Comorbid Insomnia With Obstructive Sleep Apnea: Clinical Characteristics and Risk Factors. *J Clin Sleep Med,* 14**,** 409-417.

Christodoulou, E., Ma, J., Collins, G. S., Steyerberg, E. W., Verbakel, J. Y., *et al.* (2019). A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *Journal of Clinical Epidemiology,* 110**,** 12-22.

Chung, C. C., Kang, J. H., Yuan, R. Y., Wu, D., Chen, C. C., *et al.* (2013). Multiscale entropy analysis of electroencephalography during sleep in patients with Parkinson disease. *Clin EEG Neurosci,* 44**,** 221-6.

Church, M. W., Johnson, L. C. & Seales, D. M. (1978). Evoked K-complexes and cardiovascular responses to spindle-synchronous and spindle-asynchronous stimulus clicks during NREM sleep. *Electroencephalogr Clin Neurophysiol,* 45**,** 443-53.

Colrain, I. M. (2005). The K-complex: a 7-decade history. *Sleep,* 28**,** 255-73.

Colrain, I. M., Crowley, K. E., Nicholas, C. L., Afifi, L., Baker, F. C., *et al.* (2010). Sleep evoked delta frequency responses show a linear decline in amplitude across the adult lifespan. *Neurobiol Aging,* 31**,** 874-83.

Colrain, I. M., Crowley, K. E., Nicholas, C. L., Padilla, M. & Baker, F. C. (2009). The impact of alcoholism on sleep evoked Delta frequency responses. *Biol Psychiatry,* 66**,** 177-84.

Colrain, I. M., Webster, K. E. & Hirst, G. (1999). The N550 component of the evoked K-complex: a modality non-specific response? *J Sleep Res,* 8**,** 273-80.

Crowley, K., Sullivan, E. V., Adalsteinsson, E., Pfefferbaum, A. & Colrain, I. M. (2005). Differentiating pathologic delta from healthy physiologic delta in patients with Alzheimer disease. *Sleep,* 28**,** 865-70.

Crowley, K., Trinder, J. & Colrain, I. M. (2004). Evoked K-complex generation: the impact of sleep spindles and age. *Clin Neurophysiol,* 115**,** 471-6.

Crowley, K., Trinder, J., Kim, Y., Carrington, M. & Colrain, I. M. (2002). The effects of normal aging on sleep spindle and K-complex production. *Clin Neurophysiol,* 113**,** 1615-22.

D'rozario, A. L., Cross, N. E., Vakulin, A., Bartlett, D. J., Wong, K. K. H., *et al.* (2017a). Quantitative electroencephalogram measures in adult obstructive sleep apnea - Potential biomarkers of neurobehavioural functioning. *Sleep Med Rev,* 36**,** 29-42.

D'rozario, A. L., Cross, N. E., Vakulin, A., Bartlett, D. J., Wong, K. K. H., *et al.* (2017b). Quantitative electroencephalogram measures in adult obstructive sleep apnea – Potential biomarkers of neurobehavioural functioning. *Sleep Medicine Reviews,* 36**,** 29-42.

Danker-Hopfe, H., Anderer, P., Zeitlhofer, J., Boeck, M., Dorn, H., *et al.* (2009). PSG : new scoring rules Interrater reliability for sleep scoring according to the Rechtschaffen & Kales and the new AASM standard. *Journal of Sleep Research***,** 74-84.

De Gennaro, L., Gorgoni, M., Reda, F., Lauri, G., Truglia, I., *et al.* (2017). The Fall of Sleep K-Complex in Alzheimer Disease. *Sci Rep,* 7**,** 39688.

De Vivo, L., Nagai, H., De Wispelaere, N., Spano, G. M., Marshall, W., *et al.* (2019). Evidence for sleep-dependent synaptic renormalization in mouse pups. *Sleep,* 42.

De Zambotti, M., Willoughby, A. R., Franzen, P. L., Clark, D. B., Baker, F. C., *et al.* (2016). K-Complexes: Interaction between the Central and Autonomic Nervous Systems during Sleep. *Sleep,* 39**,** 1129-37.

Dean, D. A., 2nd, Goldberger, A. L., Mueller, R., Kim, M., Rueschman, M., *et al.* (2016). Scaling Up Scientific Discovery in Sleep Medicine: The National Sleep Research Resource. *Sleep,* 39**,** 1151-64.

Devuyst, S., Dutoit, T., Stenuit, P. & Kerkhofs, M. (2010). Automatic K-complexes detection in sleep EEG recordings using likelihood thresholds. *Annu Int Conf IEEE Eng Med Biol Soc,* 2010**,** 4658-61.

Djonlagic, I., Mariani, S., Fitzpatrick, A. L., Van Der Klei, V., Johnson, D. A., *et al.* (2021). Macro and micro sleep architecture and cognitive performance in older adults. *Nat Hum Behav,* 5**,** 123-145.

Dres, M., Younes, M., Rittayamai, N., Kendzerska, T., Telias, I., *et al.* (2019). Sleep and Pathological Wakefulness at the Time of Liberation from Mechanical Ventilation (SLEEWE). A Prospective Multicenter Physiological Study. *Am J Respir Crit Care Med,* 199**,** 1106-1115.

Elmenhorst, E. M., Pennig, S., Rolny, V., Quehl, J., Mueller, U., *et al.* (2012). Examining nocturnal railway noise and aircraft noise in the field: sleep, psychomotor performance, and annoyance. *Sci Total Environ,* 424**,** 48-56.

Elwood, P. C., Bayer, A. J., Fish, M., Pickering, J., Mitchell, C., *et al.* (2011). Sleep disturbance and daytime sleepiness predict vascular dementia. *J Epidemiol Community Health,* 65**,** 820-4.

Erdamar, A., Duman, F. & Yetkin, S. (2012). A wavelet and teager energy operator based method for automatic detection of K-Complex in sleep EEG. *Expert Systems with Applications,* 39**,** 1284-1290.

Esser, S. K., Hill, S. L. & Tononi, G. (2007). Sleep homeostasis and cortical synchronization: I. Modeling the effects of synaptic strength on sleep slow waves. *Sleep,* 30**,** 1617-1630.

Evandt, J., Oftedal, B., Hjertager Krog, N., Nafstad, P., Schwarze, P. E., *et al.* (2017). A Population-Based Study on Nighttime Road Traffic Noise and Insomnia. *Sleep,* 40.

Ferri, R., Bruni, O., Miano, S., Smerieri, A., Spruyt, K., *et al.* (2005). Inter-rater reliability of sleep cyclic alternating pattern (CAP) scoring and validation of a new computer-assisted CAP scoring method. *Clin Neurophysiol,* 116**,** 696-707.

Forget, D., Morin, C. M. & Bastien, C. H. (2011). The role of the spontaneous and evoked k-complex in good-sleeper controls and in individuals with insomnia. *Sleep,* 34**,** 1251-60.

Franzen, P. L., Woodward, S. H., Bootzin, R. R., Germain, A. & Colrain, I. M. (2012). K-complexes are not preferentially evoked to combat sounds in combat-exposed Vietnam veterans with and without post-traumatic stress disorder. *Int J Psychophysiol,* 83**,** 393-8.

Freedman, R. R. (1986). EEG power spectra in sleep-onset insomnia. *Electroencephalogr Clin Neurophysiol,* 63**,** 408-13.

Freeman, D., Sheaves, B., Waite, F., Harvey, A. G. & Harrison, P. J. (2020). Sleep disturbance and psychiatric disorders. *The Lancet Psychiatry,* 7**,** 628-637.

Fultz, N. E., Bonmassar, G., Setsompop, K., Stickgold, R. A., Rosen, B. R.*, et al.* (2019). Coupled electrophysiological, hemodynamic, and cerebrospinal fluid oscillations in human sleep. *Science,* 366**,** 628-631.

Gagnadoux, F., Le Vaillant, M., Paris, A., Pigeanne, T., Leclair-Visonneau, L.*, et al.* (2016). Relationship Between OSA Clinical Phenotypes and CPAP Treatment Outcomes. *CHEST,* 149**,** 288-290.

Gardner, J. R., Pleiss, G., Bindel, D., Weinberger, K. Q. & Wilson, A. G. GPyTorch: Blackbox Matrix-Matrix Gaussian Process Inference with GPU Acceleration. Conference on Neural Information Processing Systems (NeurIPS 2018), 2019 Montréal, Canada.

Goldstein, C. A., Berry, R. B., Kent, D. T., Kristo, D. A., Seixas, A. A.*, et al.* (2020). Artificial intelligence in sleep medicine: an American Academy of Sleep Medicine position statement. *J Clin Sleep Med,* 16**,** 605-607.

Goodfellow, I., Bengio, Y. & Courville, A. 2016. *Deep Learning*, MIT Press.

Gorgoni, M., Reda, F., D'atri, A., Scarpelli, S., Ferrara, M.*, et al.* (2019). The heritability of the human K-complex: a twin study. *Sleep,* 42.

Goswami, B., Boers, N., Rheinwalt, A., Marwan, N., Heitzig, J.*, et al.* (2018). Abrupt transitions in time series with uncertainties. *Nat Commun,* 9**,** 48.

Gottlieb, D. J., Redline, S., Nieto, F. J., Baldwin, C. M., Newman, A. B.*, et al.* (2006). Association of usual sleep duration with hypertension: the Sleep Heart Health Study. *Sleep,* 29**,** 1009-14.

Gottlieb, D. J., Yenokyan, G., Newman, A. B., O'connor, G. T., Punjabi, N. M.*, et al.* (2010). Prospective study of obstructive sleep apnea and incident coronary heart disease and heart failure: the sleep heart health study. *Circulation,* 122**,** 352-60.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D.*, et al.* (2013). MEG and EEG data analysis with MNE-Python. *Front Neurosci,* 7**,** 267.

Grandner, M. A., Kripke, D. F., Yoon, I. Y. & Youngstedt, S. D. (2006). Criterion validity of the Pittsburgh Sleep Quality Index: Investigation in a non-clinical sample. *Sleep Biol Rhythms,* 4**,** 129-139.

Grant, J. F., Martin, S. A., Taylor, A. W., Wilson, D. H., Araujo, A.*, et al.* (2014). Cohort profile: The men androgen inflammation lifestyle environment and stress (MAILES) study. *Int J Epidemiol,* 43**,** 1040-53.

Griefahn, B., Brode, P., Marks, A. & Basner, M. (2008). Autonomic arousals related to traffic noise during sleep. *Sleep,* 31**,** 569-77.

Grote, L., Hedner, J. & Peter, J. H. (2000). Sleep-related breathing disorder is an independent risk factor for uncontrolled hypertension. *J Hypertens,* 18**,** 679-85.

Guilleminault, C., Harper, R., Hayes, B., Hirshkowitz, M., Ktonas, P.*, et al.* (1992). EEG Arousals: Scoring Rules and Examples. *Sleep,* 15**,** 173-184.

Gupta, M. A. & Knapp, K. (2014). Cardiovascular and psychiatric morbidity in obstructive sleep apnea (OSA) with insomnia (sleep apnea plus) versus obstructive sleep apnea without insomnia: a case-control study from a Nationally Representative US sample. *PLoS One,* 9**,** e90021.

Hahn, M. A., Heib, D., Schabus, M., Hoedlmoser, K. & Helfrich, R. F. (2020). Slow oscillation-spindle coupling predicts enhanced memory formation from childhood to adolescence. *Elife,* 9.

Halasz, P. (2016). The K-complex as a special reactive sleep slow wave - A theoretical update. *Sleep Med Rev,* 29**,** 34-40.

Halász, P. (2005). K-complex, a reactive EEG graphoelement of NREM sleep: An old chap in a new garment. *Sleep Medicine Reviews,* 9**,** 391-412.

Halasz, P., Bodizs, R., Parrino, L. & Terzano, M. (2014). Two features of sleep slow waves: homeostatic and reactive aspects--from long term to instant sleep homeostasis. *Sleep Med,* 15**,** 1184-95.

Hansen, K. L., Nguyen, P., Zajamšek, B., Catcheside, P. & Hansen, C. H. (2019). Prevalence of wind farm amplitude modulation at long-range residential locations. *Journal of Sound and Vibration,* 455**,** 136-149.

Hartmann, S. & Baumert, M. (2019). Automatic A-Phase Detection of Cyclic Alternating Patterns in Sleep Using Dynamic Temporal Information. *IEEE Trans Neural Syst Rehabil Eng,* 27**,** 1695-1703.

Hartmann, S., Bruni, O., Ferri, R., Redline, S. & Baumert, M. (2020a). Characterization of cyclic alternating pattern during sleep in older men and women using large population studies. *Sleep,* 43.

Hartmann, S., Bruni, O., Ferri, R., Redline, S. & Baumert, M. (2020b). Cyclic alternating pattern (CAP) in children with obstructive sleep apnea and its relationship with adenotonsillectomy, behavior, cognition, and quality-of-life. *Sleep*.

Harvey, A. G. (2002). A cognitive model of insomnia. *Behav Res Ther,* 40**,** 869-93.

Hayley, A. C., Williams, L. J., Venugopal, K., Kennedy, G. A., Berk, M.*, et al.* (2015). The relationships between insomnia, sleep apnoea and depression: findings from the American National Health and Nutrition Examination Survey, 2005-2008. *Aust N Z J Psychiatry,* 49**,** 156-70.

Heinzer, R., Vat, S., Marques-Vidal, P., Marti-Soler, H., Andries, D.*, et al.* (2015). Prevalence of sleep-disordered breathing in the general population: the HypnoLaus study. *The Lancet Respiratory Medicine,* 3**,** 310-318.

Helfrich, R. F., Lendner, J. D., Mander, B. A., Guillen, H., Paff, M.*, et al.* (2019). Bidirectional prefrontal-hippocampal dynamics organize information transfer during sleep in humans. *Nat Commun,* 10**,** 3572.

Hinton, G. (2018). Deep Learning-A Technology With the Potential to Transform Health Care. *JAMA,* 320**,** 1101-1102.

Hirotsu, C., Haba-Rubio, J., Andries, D., Tobback, N., Marques-Vidal, P.*, et al.* (2019). Effect of Three Hypopnea Scoring Criteria on OSA Prevalence and Associated Comorbidities in the General Population. *J Clin Sleep Med,* 15**,** 183-194.

Hogan, S. E., Delgado, G. M., Hall, M. H., Nimgaonkar, V. L., Germain, A.*, et al.* (2020). Slow-oscillation activity is reduced and high frequency activity is elevated in older adults with insomnia. *J Clin Sleep Med,* 16**,** 1445-1454.

Huang, T., Mariani, S. & Redline, S. (2020). Sleep Irregularity and Risk of Cardiovascular Events: The Multi-Ethnic Study of Atherosclerosis. *J Am Coll Cardiol,* 75**,** 991-999.

Hygge, S., Evans, G. W. & Bullinger, M. (2002). A prospective study of some effects of aircraft noise on cognitive performance in schoolchildren. *Psychol Sci,* 13**,** 469-74.

Iber, C., Ancoli-Israel, S., Chesson, A. L. & Quan, S. 2007. *AASM - Manual for the Scoring of Sleep and Associated Events*.

Inouye, T., Shinosaki, K., Sakamoto, H., Toi, S., Ukai, S.*, et al.* (1991). Quantification of EEG irregularity by use of the entropy of the power spectrum. *Electroencephalography and Clinical Neurophysiology,* 79**,** 204-210.

Ioffe, S. & Szegedy, C. (2015). Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift.

Ivanov, P. C., Liu, K. K. L. & Bartsch, R. P. (2016). Focus on the emerging new fields of Network Physiology and Network Medicine. *New J Phys,* 18.

Jahnke, K., Von Wegner, F., Morzelewski, A., Borisov, S., Maischein, M., *et al.* (2012). To wake or not to wake? The two-sided nature of the human K-complex. *NeuroImage,* 59**,** 1631-1638.

Jakovljevic, B., Belojevic, G., Paunovic, K. & Stojanov, V. (2006). Road traffic noise and sleep disturbances in an urban population: cross-sectional study. *Croat Med J,* 47**,** 125-33.

Jalali, L., Bigelow, P., Nezhad-Ahmadi, M. R., Gohari, M., Williams, D., *et al.* (2016). Before-after field study of effects of wind turbine noise on polysomnographic sleep parameters. *Noise Health,* 18**,** 194-205.

Javaheri, S. & Redline, S. (2012). Sleep, slow-wave sleep, and blood pressure. *Curr Hypertens Rep,* 14**,** 442-8.

Javaheri, S., Zhao, Y. Y., Punjabi, N. M., Quan, S. F., Gottlieb, D. J., *et al.* (2018). Slow-Wave Sleep Is Associated With Incident Hypertension: The Sleep Heart Health Study. *Sleep,* 41.

Johns, M. W. (1991). A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep,* 14**,** 540-5.

Johns, M. W. (1992). Reliability and factor analysis of the Epworth Sleepiness Scale. *Sleep,* 15**,** 376-81.

Johns, M. W. (1993). Daytime sleepiness, snoring, and obstructive sleep apnea. The Epworth Sleepiness Scale. *Chest,* 103**,** 30-6.

Johns, M. W. (1994). Sleepiness in different situations measured by the Epworth Sleepiness Scale. *Sleep,* 17**,** 703-10.

Johnson, N. L., Kotz, S. & Balakrishnan, N. 1995. *Chapter 25: Beta Distributions. Continuous Univariate Distributions Vol. 2 (2nd ed.)* Wiley.

Ju, Y. S., Ooms, S. J., Sutphen, C., Macauley, S. L., Zangrilli, M. A., *et al.* (2017). Slow wave sleep disruption increases cerebrospinal fluid amyloid-beta levels. *Brain,* 140**,** 2104-2111.

Kales, A. & Rechtschaffen, A. 1968. *A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects.*, Allan Rechtschaffen and Anthony Kales.

Kaminski, J., Brzezicka, A., Gola, M. & Wrobel, A. (2012). beta band oscillations engagement in human alertness process. *Int J Psychophysiol,* 85**,** 125-8.

Kannathal, N., Choo, M. L., Acharya, U. R. & Sadasivan, P. K. (2005). Entropies for detection of epilepsy in EEG. *Comput Methods Programs Biomed,* 80**,** 187-94.

Keith, S. E., Daigle, G. A. & Stinson, M. R. (2018). Wind turbine low frequency and infrasound propagation and sound pressure level calculations at dwellings. *J Acoust Soc Am,* 144**,** 981.

Kendzerska, T., Gershon, A. S., Hawker, G., Leung, R. S. & Tomlinson, G. (2014a). Obstructive sleep apnea and risk of cardiovascular events and all-cause mortality: a decade-long historical cohort study. *PLoS Med,* 11**,** e1001599.

Kendzerska, T., Leung, R. S., Atzema, C. L., Chandy, G., Meteb, M., *et al.* (2020). Cardiovascular consequences of obstructive sleep apnea in women: a historical cohort study. *Sleep Med,* 68**,** 71-79.

Kendzerska, T. B., Smith, P. M., Brignardello-Petersen, R., Leung, R. S. & Tomlinson, G. A. (2014b). Evaluation of the measurement properties of the Epworth sleepiness scale: a systematic review. *Sleep Med Rev,* 18**,** 321-31.

Kim, J. S., Azarbarzin, A., Wang, R., Djonlagic, I. E., Punjabi, N. M., *et al.* (2020). Association of novel measures of sleep disturbances with blood pressure: the Multi-Ethnic Study of Atherosclerosis. *Thorax,* 75**,** 57-63.

Kliuchko, M., Heinonen-Guzejev, M., Vuust, P., Tervaniemi, M. & Brattico, E. (2016). A window into the brain mechanisms associated with noise sensitivity. *Sci Rep,* 6**,** 39236.

Kohavi, R. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. International Joint Conference on Artiffcial Intelligence. , 1995.

Krakow, B., Melendrez, D., Ferreira, E., Clark, J., Warner, T. D., *et al.* (2001). Prevalence of insomnia symptoms in patients with sleep-disordered breathing. *Chest,* 120**,** 1923-9.

Kripke, D. F., Garfinkel, L., Wingard, D. L., Klauber, M. R., & Marler, M. R. (2002). Mortality associated with sleep duration and insomnia. *Archives of General Psychiatry,* 59**,** 131-136.

Krystal, A. D. & Edinger, J. D. (2008). Measuring sleep quality. *Sleep Med,* 9 Suppl 1**,** S10-7.

Krystal, A. D. & Edinger, J. D. (2010). Sleep EEG predictors and correlates of the response to cognitive behavioral therapy for insomnia. *Sleep,* 33**,** 669-77.

Krystal, A. D., Edinger, J. D., Wohlgemuth, W. K. & Marsh, G. R. (2002). NREM sleep EEG frequency spectral correlates of sleep complaints in primary insomnia subtypes. *Sleep,* 25**,** 630-40.

Kumar, M., Pachori, R. & Acharya, U. (2017). Automated Diagnosis of Myocardial Infarction ECG Signals Using Sample Entropy in Flexible Analytic Wavelet Transform Framework. *Entropy,* 19.

Kupfer, D. J. & Reynolds, C. F., 3rd (1997). Management of insomnia. *N Engl J Med,* 336**,** 341-6.

Labarca, G., Dreyse, J., Drake, L., Jorquera, J. & Barbe, F. (2020). Efficacy of continuous positive airway pressure (CPAP) in the prevention of cardiovascular events in patients with obstructive sleep apnea: Systematic review and meta-analysis. *Sleep Med Rev,* 52**,** 101312.

Lajnef, T., Chaibi, S., Eichenlaub, J. B., Ruby, P. M., Aguera, P. E., *et al.* (2015). Sleep spindle and K-complex detection using tunable Q-factor wavelet transform and morphological component analysis. *Front Hum Neurosci,* 9**,** 414.

Lang, C. J., Appleton, S. L., Vakulin, A., Mcevoy, R. D., Vincent, A. D., *et al.* (2017a). Associations of Undiagnosed Obstructive Sleep Apnea and Excessive Daytime Sleepiness With Depression: An Australian Population Study. *J Clin Sleep Med,* 13**,** 575-582.

Lang, C. J., Appleton, S. L., Vakulin, A., Mcevoy, R. D., Wittert, G. A., *et al.* (2017b). Co-morbid OSA and insomnia increases depression prevalence and severity in men. *Respirology,* 22**,** 1407-1415.

Lasa, L., Ayuso-Mateos, J. L., Vazquez-Barquero, J. L., Diez-Manrique, F. J. & Dowrick, C. F. (2000). The use of the Beck Depression Inventory to screen for depression in the general population: a preliminary analysis. *Journal of Affective Disorders,* 57**,** 261-265.

Lavie, P., Herer, P. & Hoffstein, V. (2000). Obstructive sleep apnoea syndrome as a risk factor for hypertension: population study. *BMJ,* 320**,** 479-82.

Lazar, A. S., Lazar, Z. I. & Dijk, D. J. (2015). Circadian regulation of slow waves in human sleep: Topographical aspects. *Neuroimage,* 116**,** 123-34.

Leary, E. B., Watson, K. T., Ancoli-Israel, S., Redline, S., Yaffe, K., *et al.* (2020). Association of Rapid Eye Movement Sleep With Mortality in Middle-aged and Older Adults. *JAMA Neurol.*

Leblanc, M., Merette, C., Savard, J., Ivers, H., Baillargeon, L., *et al.* (2009). Incidence and risk factors of insomnia in a population-based sample. *Sleep,* 32**,** 1027-37.

Lecci, S., Cataldi, J., Betta, M., Bernardi, G., Heinzer, R., *et al.* (2020). EEG changes associated with subjective under- and overestimation of sleep duration. *Sleep.*

Lechat, B., Hansen, K., Catcheside, P. & Zajamsek, B. (2020). Beyond K-complex binary scoring during sleep: probabilistic classification using deep learning. *Sleep,* 43.

Lee, G., Gommers, R., Waselewski, F., Wohlfahrt, K. & O'leary, A. (2019). PyWavelets: A Python package for wavelet analysis. *Journal of Open Source Software,* 4.

Lee, Y. S., Wingfield, A., Min, N. E., Kotloff, E., Grossman, M., *et al.* (2018). Differences in Hearing Acuity among "Normal-Hearing" Young Adults Modulate the Neural Basis for Speech Comprehension. *eNeuro,* 5.

Leger, D., Debellemaniere, E., Rabat, A., Bayon, V., Benchenane, K., *et al.* (2018). Slow-wave sleep: From the cell to the clinic. *Sleep Med Rev,* 41**,** 113-132.

Leibig, C., Allken, V., Ayhan, M. S., Berens, P. & Wahl, S. (2017). Leveraging uncertainty information from deep neural networks for disease detection. *Sci Rep,* 7**,** 17816.

Lendner, J. D., Helfrich, R. F., Mander, B. A., Romundstad, L., Lin, J. J., *et al.* (2020). An electrophysiological marker of arousal level in humans. *Elife,* 9.

Lercher, P., Evans, G. W. & Meis, M. (2003). Ambient noise and cognitive processes among primary schoolchildren. *Environment and Behavior,* 35**,** 725-735.

Lewinsohn, P. M., Seeley, J. R., Roberts, R. E. & Allen, N. B. (1997). Center for epidemiologic studies depression scale (CES-D) as a screening instrument for depression among community-residing older adults. *Psychology and Aging,* 12**,** 277-287.

Li, Y., Vgontzas, A. N., Fernandez-Mendoza, J., Bixler, E. O., Sun, Y., *et al.* (2015). Insomnia with physiological hyperarousal is associated with hypertension. *Hypertension,* 65**,** 644-50.

Lim, D. C., Mazzotti, D. R., Sutherland, K., Mindel, J. W., Kim, J., *et al.* (2020). Reinventing polysomnography in the age of precision medicine. *Sleep Med Rev,* 52**,** 101313.

Liu, S., Pan, J., Lei, Q., He, L., Zhong, B., *et al.* (2020). Spontaneous K-Complexes may be biomarkers of the progression of amnestic mild cognitive impairment. *Sleep Med,* 67**,** 99-109.

Lovato, N. & Lack, L. (2019). Insomnia and mortality: A meta-analysis. *Sleep Medicine reviews,* 43**,** 71-83.

Lunsford-Avery, J. R., Edinger, J. D. & Krystal, A. D. (2021). Optimizing computation of overnight decline in delta power: Evidence for slower rate of decline in delta power in insomnia patients. *Clin Neurophysiol,* 132**,** 545-553.

Lyznicki, J. M., Doege, T. C., Davis, R. M. & Williams, M. A. (1998). Sleepiness, driving, and motor vehicle crashes. Council on Scientific Affairs, American Medical Association. *JAMA,* 279**,** 1908-13.

Ma, Y., Shi, W., Peng, C. K. & Yang, A. C. (2018). Nonlinear dynamical analysis of sleep electroencephalography using fractal and entropy approaches. *Sleep Med Rev,* 37**,** 85-93.

Magalang, U. J., Chen, N. H., Cistulli, P. A., Fedson, A. C., Gislason, T., *et al.* (2013). Agreement in the scoring of respiratory events and sleep among international sleep centers. *Sleep,* 36**,** 591-6.

Maingret, N., Girardeau, G., Todorova, R., Goutierre, M. & Zugaro, M. (2016). Hippocampo-cortical coupling mediates memory consolidation during sleep. *Nat Neurosci,* 19**,** 959-64.

Mak-Mccully, R. A., Deiss, S. R., Rosen, B. Q., Jung, K. Y., Sejnowski, T. J., *et al.* (2014). Synchronization of isolated downstates (K-complexes) may be caused by cortically-induced disruption of thalamic spindling. *PLoS Comput Biol,* 10**,** e1003855.

Mak-Mccully, R. A., Rosen, B. Q., Rolland, M., Regis, J., Bartolomei, F., *et al.* (2015). Distribution, Amplitude, Incidence, Co-Occurrence, and Propagation of Human K-Complexes in Focal Transcortical Recordings. *eNeuro,* 2.

Mallat, S. (1989). A theory for multiresolution signal decomposition: the waveletrepresentation. *Transactions on pattern analysis and machine intelligence,* 11**,** 674-693.

Marks, A. & Griefahn, B. (2007). Associations between noise sensitivity and sleep, subjectively evaluated sleep quality, annoyance, and performance after exposure to nocturnal traffic noise. *Noise Health,* 9**,** 1-7.

Marks, A., Griefahn, B. & Basner, M. (2008). Event-related awakenings caused by nocturnal transportation noise. *Noise Control Engineering Journal,* 56**,** 52-62.

Marshall, L., Helgadottir, H., Molle, M. & Born, J. (2006). Boosting slow oscillations during sleep potentiates memory. *Nature,* 444**,** 610-3.

Marshall, N. S., Wong, K. K. H., Cullen, S. R. J., Knuiman, M. W. & Grunstein, R. R. (2014). Sleep Apnea and 20-Year Follow-Up for All-Cause Mortality, Stroke, and Cancer Incidence and Mortality in the Busselton Health Study Cohort. *Journal of Clinical Sleep Medicine,* 10**,** 355-362.

Martin, S. E., Wraith, P. K., Deary, I. J. & Douglas, N. J. (1997). The effect of nonvisible sleep fragmentation on daytime function. *Am J Respir Crit Care Med,* 155**,** 1596-601.

Massimini, M., Huber, R., Ferrarelli, F., Hill, S. & Tononi, G. (2004). The sleep slow oscillation as a traveling wave. *J Neurosci,* 24**,** 6862-70.

Mazzotti, D. R., Keenan, B. T., Lim, D. C., Gottlieb, D. J., Kim, J.*, et al.* (2019). Symptom Subtypes of Obstructive Sleep Apnea Predict Incidence of Cardiovascular Outcomes. *Am J Respir Crit Care Med,* 200**,** 493-506.

Mcevoy, R. D., Antic, N. A., Heeley, E., Luo, Y., Ou, Q.*, et al.* (2016). CPAP for Prevention of Cardiovascular Events in Obstructive Sleep Apnea. *N Engl J Med,* 375**,** 919-31.

Mcguire, S., Muller, U., Elmenhorst, E. M. & Basner, M. (2016). Inter-individual Differences in the Effects of Aircraft Noise on Sleep Fragmentation. *Sleep,* 39**,** 1107-10.

Melaku, Y. A., Gill, T. K., Appleton, S. L., Hill, C., Boyd, M. A.*, et al.* (2019). Sociodemographic, lifestyle and metabolic predictors of all-cause mortality in a cohort of community-dwelling population: an 18-year follow-up of the North West Adelaide Health Study. *BMJ Open,* 9**,** e030079.

Mendonca, F., Mostafa, S. S., Morgado-Dias, F., Ravelo-Garcia, A. G. & Penzel, T. (2019). A Review of Approaches for Sleep Quality Analysis. *IEEE Access,* 7**,** 24527-24546.

Michaud, D. S., Feder, K., Keith, S. E., Voicescu, S. A., Marro, L.*, et al.* (2016). Effects of Wind Turbine Noise on Self-Reported and Objective Measures of Sleep. *Sleep,* 39**,** 97-109.

Michaud, D. S., Keith, S. E., Guay, M., Voicescu, S., Denning, A.*, et al.* (2021). Sleep actigraphy time-synchronized with wind turbine output. *Sleep.*

Micic, G., Zajamsek, B., Lack, L., Hansen, K., Doolan, C.*, et al.* (2018). A Review of the Potential Impacts of Wind Farm Noise on Sleep. *Acoustics Australia,* 46**,** 87-97.

Monstad, P. & Guilleminault, C. (1999). Cardiovascular changes associated with spontaneous and evoked K-complexes. *Neurosci Lett,* 263**,** 211-3.

Montplaisir, J., Boucher, S., Gosselin, A., Poirier, G. & Lavigne, G. (1996). Persistence of repetitive EEG arousals (K-alpha complexes) in RLS patients treated with L-DOPA. *Sleep,* 19**,** 196-199.

Morin, C. M., Belleville, G., Belanger, L. & Ivers, H. (2011). The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. *Sleep,* 34**,** 601-8.

Morin, C. M. & Benca, R. (2012). Chronic insomnia. *Lancet,* 379**,** 1129-41.

Mostafa, S. S., Mendonca, F., Ravelo-Garcia, A. G. & Morgado-Dias, F. (2019). A Systematic Review of Detecting Sleep Apnea Using Deep Learning. *Sensors (Basel),* 19.

Muehlroth, B. E., Sander, M. C., Fandakova, Y., Grandy, T. H., Rasch, B.*, et al.* (2019). Precise Slow Oscillation-Spindle Coupling Promotes Memory Consolidation in Younger and Older Adults. *Sci Rep,* 9**,** 1940.

Munzel, T., Schmidt, F. P., Steven, S., Herzog, J., Daiber, A.*, et al.* (2018). Environmental Noise and the Cardiovascular System. *J Am Coll Cardiol,* 71**,** 688-697.

Nair, V. & Hinton, G. E. (2010). Rectified Linear Units Improve Restricted Boltzmann Machines *International Conference on Machine Learning***,** 6421113.

Natsky A., V. A., Chai-Coetzer Cl., Lack L., Mcevoy Rd., Lovato N., Sweetman A., Gordon Cj., Adams Rj., Kaambwa B. (2020). Economic evaluation of cognitive behavioural therapy for insomnia (CBT-I) for improving health outcomes in adult populations: a systematic review *Sleep Medicine Reviews,* in press.

Newson, R. B. (2010). Comparing the predictive powers of survival models using Harrell's C or Somers' D. *Stata Journal,* 10**,** 339-358.

Nguyen, C. D., Wellman, A., Jordan, A. S. & Eckert, D. J. (2016). Mild Airflow Limitation during N2 Sleep Increases K-complex Frequency and Slows Electroencephalographic Activity. *Sleep,* 39**,** 541-50.

Nguyen, D.-P., Hansen, K. & Zajamsek, B. (2019). Human perception of wind farm vibration. *Journal of Low Frequency Noise, Vibration and Active Control,* 39**,** 17-27.

Nicholas, C. L., Trinder, J. & Colrain, I. M. (2002). Increased production of evoked and spontaneous K-complexes following a night of fragmented sleep. *Sleep,* 25**,** 882-7.

Nicholas, C. L., Trinder, J., Crowley, K. E. & Colrain, I. M. (2006). The impact of slow wave sleep proximity on evoked K-complex generation. *Neurosci Lett,* 404**,** 127-31.

Nieto, F. J., Young, T. B., Lind, B. K., Shahar, E., Samet, J. M.*, et al.* (2000). Association of sleep-disordered breathing, sleep apnea, and hypertension in a large community-based study. Sleep Heart Health Study. *JAMA,* 283**,** 1829-36.

Nir, Y., Staba, R. J., Andrillon, T., Vyazovskiy, V. V., Cirelli, C.*, et al.* (2011). Regional slow waves and spindles in human sleep. *Neuron,* 70**,** 153-69.

Nofzinger, E. A., Buysse, D. J., Germain, A., Price, J. C., Miewald, J. M.*, et al.* (2004). Functional neuroimaging evidence for hyperarousal in insomnia. *Am J Psychiatry,* 161**,** 2126-8.

O'reilly, C., Gosselin, N., Carrier, J. & Nielsen, T. (2014). Montreal Archive of Sleep Studies: an open-access resource for instrument benchmarking and exploratory research. *J Sleep Res,* 23**,** 628-635.

Obermeyer, Z. & Emanuel, E. J. (2016). Predicting the Future - Big Data, Machine Learning, and Clinical Medicine. *N Engl J Med,* 375**,** 1216-9.

Ohayon, M., Wickwire, E. M., Hirshkowitz, M., Albert, S. M., Avidan, A.*, et al.* (2017). National Sleep Foundation's sleep quality recommendations: first report. *Sleep Health,* 3**,** 6-19.

Ohayon, M. M., & Reynolds, C. F. (2009). Epidemiological and clinical relevance of insomnia diagnosis algorithms according to the DSM-IV and the International Classification of Sleep Disorders (ICSD). *Sleep Medicine,* 10.

Ohayon, M. M. & Reynolds, C. F. (2009). Epidemiological and clinical relevance of insomnia diagnosis algorithms according to the DSM-IV and the International Classification of Sleep Disorders (ICSD). *Sleep Medicine,* 10**,** 952-960.

Parekh, A., Mullins, A. E., Kam, K., Varga, A. W., Rapoport, D. M.*, et al.* (2019). Slow-wave activity surrounding stage N2 K-complexes and daytime function measured by psychomotor vigilance test in obstructive sleep apnea. *Sleep,* 42.

Parekh, A., Selesnick, I. W., Rapoport, D. M. & Ayappa, I. (2015). Detection of K-complexes and sleep spindles (DETOKS) using sparse optimization. *J Neurosci Methods,* 251**,** 37-46.

Park, J., Lindberg, C. R. & Vernon, F. L. (1987). Multitaper spectral analysis of high-frequency seismograms. *Journal of Geophysical Research,* 92.

Parrino, L., Ferri, R., Bruni, O. & Terzano, M. G. (2012). Cyclic alternating pattern (CAP): the marker of sleep instability. *Sleep Med Rev,* 16**,** 27-45.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J.*, et al.* PyTorch: An Imperative Style, High-Performance Deep Learning Library. Advances in Neural Information Processing Systems 32, 2019.

Patel, S. R., Zhu, X., Storfer-Isser, A., Mehra, R., Jenny, N. S.*, et al.* (2009). Sleep duration and biomarkers of inflammation. *Sleep,* 32**,** 200-4.

Pedersen, E. & Persson Waye, K. (2007). Wind turbine noise, annoyance and self-reported health and well-being in different living environments. *Occup Environ Med,* 64**,** 480-6.

Pedersen, E. & Waye, K. P. (2004). Perception and annoyance due to wind turbine noise--a dose-response relationship. *J Acoust Soc Am,* 116**,** 3460-70.

Peker, Y., Glantz, H., Eulenburg, C., Wegscheider, K., Herlitz, J.*, et al.* (2016). Effect of Positive Airway Pressure on Cardiovascular Outcomes in Coronary Artery Disease Patients with Nonsleepy Obstructive Sleep Apnea. The RICCADSA Randomized Controlled Trial. *Am J Respir Crit Care Med,* 194**,** 613-20.

Penzel, T., Kantelhardt, J. W., Bartsch, R. P., Riedl, M., Kraemer, J. F.*, et al.* (2016). Modulations of Heart Rate, ECG, and Cardio-Respiratory Coupling Observed in Polysomnography. *Front Physiol,* 7**,** 460.

Peppard, P. E., Szklo-Coxe, M., Hla, K. M. & Young, T. (2006). Longitudinal association of sleep-related breathing disorder and depression. *Arch Intern Med,* 166**,** 1709-15.

Peppard, P. E., Young, T., Barnet, J. H., Palta, M., Hagen, E. W.*, et al.* (2013). Increased prevalence of sleep-disordered breathing in adults. *American journal of epidemiology,* 177**,** 1006-1014.

Perlis, M. L., Smith, M. T., Andrews, P. J., Orff, H. & Giles, D. E. (2001). Beta/Gamma EEG activity in patients with primary and secondary insomnia and good sleeper controls. *Sleep,* 24**,** 110-7.

Phan, H., Andreotti, F., Cooray, N., Chen, O. Y. & De Vos, M. (2019). SeqSleepNet: End-to-End Hierarchical Recurrent Neural Network for Sequence-to-Sequence Automatic Sleep Staging. *IEEE Trans Neural Syst Rehabil Eng,* 27**,** 400-410.

Plotly Technologies Inc. 2015. Collaborative data science. Montreal, QC: Plotly Technologies Inc.

Poulsen, A. H., Raaschou-Nielsen, O., Pena, A., Hahmann, A. N., Nordsborg, R. B.*, et al.* (2019). Impact of Long-Term Exposure to Wind Turbine Noise on Redemption of Sleep Medication and Antidepressants: A Nationwide Cohort Study. *Environ Health Perspect,* 127**,** 37005.

Prerau, M. J., Brown, R. E., Bianchi, M. T., Ellenbogen, J. M. & Purdon, P. L. (2017). Sleep Neurophysiological Dynamics Through the Lens of Multitaper Spectral Analysis. *Physiology (Bethesda),* 32**,** 60-92.

Punjabi, N. M., Caffo, B. S., Goodwin, J. L., Gottlieb, D. J., Newman, A. B.*, et al.* (2009). Sleep-disordered breathing and mortality: a prospective cohort study. *PLoS Med,* 6**,** e1000132.

Quan, S. F., Howard, B. V., Iber, C., Kiley, J. P., Nieto, F. J.*, et al.* (1997). The Sleep Heart Health Study: design, rationale, and methods. *Sleep,* 20**,** 1077-85.

R Core Team 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.

Rasch, B., Buchel, C., Gais, S. & Born, J. (2007). Odor cues during slow-wave sleep prompt declarative memory consolidation. *Science,* 315**,** 1426-9.

Rasmussen, C. E. (2006). Gaussian processes for machine learning. *IJNS, International journal of neural systems,* 14**,** 69-106.

Rechtschaffen, A., Bergmann, B. M., Everson, C. A., Kushida, C. A. & Gilliland, M. A. (2002). Sleep Deprivation in the Rat: X. Integration and Discussion of the Findings. *Sleep,* 25**,** 68-87.

Rechtschaffen, A. & Kales, A. 1968. A manual of standardized terminology: techniques and scoring system for sleep stages of human subjects. Los Angeles: UCLA Brain Information Service/Brain Research Institute.

Redline, S., Dean, D., 3rd & Sanders, M. H. (2013). Entering the era of "big data": getting our metrics right. *Sleep,* 36**,** 465-9.

Redline, S., Sanders, M. H., Lind, B. K., Quan, S. F., Iber, C., *et al.* (1998). Methods for obtaining and analyzing unattended polysomnography data for a multicenter study. Sleep Heart Health Research Group. *Sleep,* 21**,** 759-67.

Redline, S., Strauss, M. E., Adams, N., Winters, M., Roebuck, T., *et al.* (1997). Neuropsychological function in mild sleep-disordered breathing. *Sleep,* 20**,** 160-7.

Redline, S., Tishler, P. V., Schluchter, M., Aylor, J., Clark, K., *et al.* (1999). Risk factors for sleep-disordered breathing in children. Associations with obesity, race, and respiratory problems. *Am J Respir Crit Care Med,* 159**,** 1527-32.

Redline, S., Yenokyan, G., Gottlieb, D. J., Shahar, E., O'connor, G. T., *et al.* (2010). Obstructive Sleep Apnea–Hypopnea and Incident Stroke. *American Journal of Respiratory and Critical Care Medicine,* 182**,** 269-277.

Richard, C. & Lengelle, R. (1998). Joint time and time-frequency optimal detection of K-complexes in sleep EEG. *Comput Biomed Res,* 31**,** 209-29.

Riedner, B. A., Vyazovskiy, V. V., Huber, R., Massimini, M., Esser, S., *et al.* (2007). Sleep homeostasis and cortical synchronization: III. A high-density EEG study of sleep slow waves in humans. *Sleep,* 30**,** 1643-1657.

Riemann, D., Spiegelhalder, K., Feige, B., Voderholzer, U., Berger, M., *et al.* (2010). The hyperarousal model of insomnia: a review of the concept and its evidence. *Sleep Med Rev,* 14**,** 19-31.

Rodenbeck, A., Huether, G., Ruther, E. & Hajak, G. (2002). Interactions between evening and nocturnal cortisol secretion and sleep parameters in patients with severe chronic primary insomnia. *Neuroscience Letters,* 324**,** 159-163.

Rosenberg, R. S. & Van Hout, S. (2013). The American Academy of Sleep Medicine inter-scorer reliability program: sleep stage scoring. *J Clin Sleep Med,* 9**,** 81-7.

Rosenthal, L. D. & Dolan, D. C. (2008). The Epworth sleepiness scale in the identification of obstructive sleep apnea. *J Nerv Ment Dis,* 196**,** 429-31.

Roy, Y., Banville, H., Albuquerque, I., Gramfort, A., Falk, T. H., *et al.* (2019). Deep learning-based electroencephalography analysis: a systematic review. *J Neural Eng,* 16**,** 051001.

Rubin, D. B. 1987. *Multiple Imputation for Nonresponse in Surveys,* New York.

Rudzik, F., Thiesse, L., Pieren, R., Wunderli, J. M., Brink, M., *et al.* (2018). Sleep spindle characteristics and arousability from nighttime transportation noise exposure in healthy young and older individuals. *Sleep,* 41**,** 1-14.

Ruehland, W. R., Churchward, T. J., Schachter, L. M., Lakey, T., Tarquinio, N., *et al.* (2015). Polysomnography using abbreviated signal montages: impact on sleep and cortical arousal scoring. *Sleep Med,* 16**,** 173-80.

Ruehland, W. R., O'donoghue, F. J., Pierce, R. J., Thornton, A. T., Singh, P., *et al.* (2011). The 2007 AASM recommendations for EEG electrode placement in polysomnography: impact on sleep and cortical arousal scoring. *Sleep,* 34**,** 73-81.

Sánchez-De-La-Torre, M., Sánchez-De-La-Torre, A., Bertran, S., Abad, J., Duran-Cantolla, J., *et al.* (2020). Effect of obstructive sleep apnoea and its treatment with continuous positive airway pressure on the prevalence of cardiovascular events in patients with acute coronary syndrome (ISAACC study): a randomised controlled trial. *The Lancet Respiratory Medicine,* 8**,** 359-367.

Scammell, T. E., Arrigoni, E. & Lipton, J. O. (2017). Neural Circuitry of Wakefulness and Sleep. *Neuron,* 93**,** 747-765.

Schaffer, B., Schlittmeier, S. J., Pieren, R., Heutschi, K., Brink, M., *et al.* (2016). Short-term annoyance reactions to stationary and time-varying wind turbine and road traffic noise: A laboratory study. *J Acoust Soc Am,* 139**,** 2949.

Schoenfeld, D. (1982). Partial Residuals for the Proportional Hazards Regression-Model. *Biometrika,* 69**,** 239-241.

Scott, H., Lechat, B., Lovato, N. & Lack, L. (2020). Correspondence between physiological and behavioural responses to vibratory stimuli during the sleep onset period: A quantitative electroencephalography analysis. *J Sleep Res***,** e13232.

Sforza, E., Pichot, V., Martin, M. S., Barthelemy, J. C. & Roche, F. (2015). Prevalence and determinants of subjective sleepiness in healthy elderly with unrecognized obstructive sleep apnea. *Sleep Med,* 16**,** 981-6.

Shahar, E., Whitney, C. W., Redline, S., Lee, E. T., Newman, A. B.*, et al.* (2001). Sleep-disordered breathing and cardiovascular disease: cross-sectional results of the Sleep Heart Health Study. *Am J Respir Crit Care Med,* 163**,** 19-25.

Silvani, A. & Dampney, R. A. (2013). Central control of cardiovascular function during sleep. *Am J Physiol Heart Circ Physiol,* 305**,** H1683-92.

Sivertsen, B., Bjornsdottir, E., Overland, S., Bjorvatn, B. & Salo, P. (2013). The joint contribution of insomnia and obstructive sleep apnoea on sickness absence. *J Sleep Res,* 22**,** 223-30.

Slater, G. & Steier, J. (2012). Excessive daytime sleepiness in sleep disorders. *J Thorac Dis,* 4**,** 608-16.

Smith, M. G., Ogren, M., Thorsson, P., Hussain-Alkhateeb, L., Pedersen, E.*, et al.* (2020). A laboratory study on the effects of wind turbine noise on sleep: Results of the polysomnographic WiTNES study. *Sleep.*

Smith, M. G., Ögren, M., Thorsson, P. & Pedersen, E. 2016. Physiological effects of wind turbine noise on sleep. *International Congress of Acoustic.* Buenos Aires.

Smith, M. G., Younes, M., Aeschbach, D., Müller, U. & Basner, M. (2019). 0131 Planes, Trains And Automobiles: Traffic Noise And Its Impact On Sleep Depth Measured By Odds Ratio Product. *Sleep,* 42**,** A54-A54.

Snelson, E. & Ghahramani, Z. (2006). Sparse Gaussian Processes using Pseudo-inputs. *Neural Information Processing Systems***,** 1-24.

Snyder, A. C. & Foxe, J. J. (2010). Anticipatory attentional suppression of visual features indexed by oscillatory alpha-band power increases: a high-density electrical mapping study. *J Neurosci,* 30**,** 4024-32.

Srivastava, N., Hinton, G., Krizhevsky, A. & Salakhutdinov, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. Journal of Machine Learning Research, 2014. 1929-1958.

Stepanski, E., Glinn, M., Zorick, F., Roehrs, T. & Roth, T. (1994). Heart-Rate Changes in Chronic Insomnia. *Stress Medicine,* 10**,** 261-266.

Stephansen, J. B., Olesen, A. N., Olsen, M., Ambati, A., Leary, E. B.*, et al.* (2018). Neural network analysis of sleep stages enables efficient diagnosis of narcolepsy. *Nat Commun,* 9**,** 5229.

Steriade, M. (2006). Grouping of brain rhythms in corticothalamic systems. *Neuroscience,* 137**,** 1087-106.

Steriade, M., Nunez, A. & Amzica, F. (1993). A novel slow ($<$ 1 Hz) oscillation of neocortical neurons in vivo: Depolarizing and hyperpolarizing components. *Journal of Neuroscience,* 13**,** 3252-3265.

Steriade, M., Timofeev, I. & Grenier, F. (2001). Natural waking and sleep states: a view from inside neocortical neurons. *J Neurophysiol,* 85**,** 1969-85.

Stickgold, R. (2005). Sleep-dependent memory consolidation. *Nature,* 437**,** 1272-8.

Supratak, A., Dong, H., Wu, C. & Guo, Y. (2017). DeepSleepNet: A Model for Automatic Sleep Stage Scoring Based on Raw Single-Channel EEG. *IEEE Trans Neural Syst Rehabil Eng,* 25**,** 1998-2008.

Sweetman, A., Lack, L. & Bastien, C. (2019). Co-Morbid Insomnia and Sleep Apnea (COMISA): Prevalence, Consequences, Methodological Considerations, and Recent Randomized Controlled Trials. *Brain Sci,* 9.

Sweetman, A., Lack, L., Lambert, S., Gradisar, M. & Harris, J. (2017a). Does comorbid obstructive sleep apnea impair the effectiveness of cognitive and behavioral therapy for insomnia? *Sleep Med,* 39**,** 38-46.

Sweetman, A., Lack, L., Mcevoy, R. D., Antic, N. A., Smith, S.*, et al.* (2020). Cognitive behavioural therapy for insomnia reduces sleep apnoea severity: a randomised controlled trial. *ERJ Open Res,* 6.

Sweetman, A., Lack, L. C., Catcheside, P. G., Antic, N. A., Chai-Coetzer, C.*, et al.* (2017b). Developing a successful treatment for co-morbid insomnia and sleep apnoea. *Sleep Medicine Reviews,* 33**,** 28-38.

Taillard, J., Sagaspe, P., Berthomier, C., Brandewinder, M., Amieva, H.*, et al.* (2019). Non-REM Sleep Characteristics Predict Early Cognitive Impairment in an Aging Population. *Front Neurol,* 10**,** 197.

Talala, K. M., Martelin, T. P., Haukkala, A. H., Harkanen, T. T. & Prattala, R. S. (2012). Socio-economic differences in self-reported insomnia and stress in Finland from 1979 to 2002: a population-based repeated cross-sectional survey. *BMC Public Health,* 12**,** 650.

Tasali, E., Leproult, R., Ehrmann, D. A. & Van Cauter, E. (2008). Slow-wave sleep and the risk of type 2 diabetes in humans. *Proc Natl Acad Sci U S A,* 105**,** 1044-9.

Tasbakan, M. S., Gunduz, C., Pirildar, S. & Basoglu, O. K. (2018). Quality of life in obstructive sleep apnea is related to female gender and comorbid insomnia. *Sleep Breath,* 22**,** 1013-1020.

Teran-Santos, J., Jimenez-Gomez, A. & Cordero-Guevara, J. (1999). The association between sleep apnea and the risk of traffic accidents. Cooperative Group Burgos-Santander. *N Engl J Med,* 340**,** 847-51.

Terzano, M. G., Parrino, L., Sherieri, A., Chervin, R., Chokroverty, S.*, et al.* (2001). Atlas, rules, and recording techniques for the scoring of cyclic alternating pattern (CAP) in human sleep. *Sleep Medicine,* 2**,** 537-553.

Thomas, R. J., Chen, S., Eden, U. T. & Prerau, M. J. (2020). Quantifying statistical uncertainty in metrics of sleep disordered breathing. *Sleep Med,* 65**,** 161-169.

Thomas, R. J., Mietus, J. E., Peng, C. K. & Goldberger, A. L. (2005). An electrocardiogram-based technique to assess cardiopulmonary coupling during sleep. *Sleep,* 28**,** 1151-61.

Thomas, R. J., Wood, C. & Bianchi, M. T. (2018). Cardiopulmonary coupling spectrogram as an ambulatory clinical biomarker of sleep stability and quality in health, sleep apnea, and insomnia. *Sleep,* 41.

Tononi, G. & Cirelli, C. (2006). Sleep function and synaptic homeostasis. *Sleep Med Rev,* 10**,** 49-62.

Tononi, G. & Cirelli, C. (2014). Sleep and the price of plasticity: from synaptic and cellular homeostasis to memory consolidation and integration. *Neuron,* 81**,** 12-34.

Torsvall, L. & Akerstedt, T. (1988). Extreme sleepiness: quantification of EOG and spectral EEG parameters. *Int J Neurosci,* 38**,** 435-41.

Tsapanou, A., Gu, Y., Manly, J., Schupf, N., Tang, M. X.*, et al.* (2015). Daytime Sleepiness and Sleep Inadequacy as Risk Factors for Dementia. *Dement Geriatr Cogn Dis Extra,* 5**,** 286-95.

Tsinalis, O., Matthews, P. M., Guo, Y. & Zafeiriou, S. P. (2016). Automatic Sleep Stage Scoring with Single-Channel EEG Using Convolutional Neural Networks. *CoRR.*

Vgontzas, A. N., Liao, D., Pejovic, S., Calhoun, S., Karataraki, M.*, et al.* (2010). Insomnia with short sleep duration and mortality: the Penn State cohort. *Sleep,* 33**,** 1159-64.

Von Stein, A. & Sarnthein, J. (2000). Different frequencies for different scales of cortical integration: from local gamma to long range alpha/theta synchronization. *International Journal of Psychophysiology,* 38**,** 301-313.

Vozoris, N. T. (2012). Sleep apnea-plus: prevalence, risk factors, and association with cardiovascular diseases using United States population-level data. *Sleep Med,* 13**,** 637-44.

Vyazovskiy, V. V., Olcese, U., Hanlon, E. C., Nir, Y., Cirelli, C.*, et al.* (2011). Local sleep in awake rats. *Nature,* 472**,** 443-7.

Wang, M., Tao, D. & Huet, B. (2014). Data Augmentation for EEG-Based Emotion Recognition with Deep Convolutional Neural Networks Fang. *MultiMedia Modeling,* 281**,** 521-522.

Watson, N. F. (2016). Health Care Savings: The Economic Value of Diagnostic and Therapeutic Care for Obstructive Sleep Apnea. *J Clin Sleep Med,* 12**,** 1075-7.

Weinstein, N. D. Individual Differences in Reactions to Noise: A Longitudinal Study in a College Dormitory. Journal of Applied Psychology, 1978. 458-466.

Werth, E., Dijk, D. J., Achermann, P. & Borbely, A. A. (1996). Dynamics of the sleep EEG after an early evening nap: experimental data and simulations. *Am J Physiol,* 271**,** R501-10.

Willoughby, A. R., De Zambotti, M., Baker, F. C. & Colrain, I. M. (2020). Evoked K-complexes and altered interaction between the central and autonomic nervous systems during sleep in alcohol use disorder. *Alcohol,* 84**,** 1-7.

Wilson, A. G., Hu, Z., Salakhutdinov, R. & Xing, E. P. (2015). Deep Kernel Learning.

Wilson, A. G., Hu, Z., Salakhutdinov, R. & Xing, E. P. Stochastic Variational Deep Kernel Learning. Conference on Neural Information Processing Systems, 2016 Barcelona.

Won, C. H. J., Reid, M., Sofer, T., Azarbarzin, A., Purcell, S.*, et al.* (2020). Sex differences in obstructive sleep apnea phenotypes, the multi-ethnic study of atherosclerosis. *Sleep,* 43.

World Health Organization 2011. *Burden of disease from environmental noise: Quantification of healthy life years lost in Europe,* Copenhagen, WHO, Regional Office for Europe.

Wulff, K., Gatti, S., Wettstein, J. G. & Foster, R. G. (2010). Sleep and circadian rhythm disruption in psychiatric and neurodegenerative disease. *Nat Rev Neurosci,* 11**,** 589-99.

Yang, C. M., Liao, Y. S., Lin, C. M., Chou, S. L. & Wang, E. N. (2011). Psychological and behavioral factors in patients with comorbid obstructive sleep apnea and insomnia. *J Psychosom Res,* 70**,** 355-61.

Younes, M. (2017). The case for using digital EEG analysis in clinical sleep medicine. *Sleep Science and Practice,* 1.

Younes, M. & Giannouli, E. (2020). Mechanism of excessive wake time when associated with obstructive sleep apnea or periodic limb movements. *J Clin Sleep Med,* 16**,** 389-399.

Younes, M. & Hanly, P. J. (2016). Immediate postarousal sleep dynamics: an important determinant of sleep stability in obstructive sleep apnea. *J Appl Physiol (1985),* 120**,** 801-8.

Younes, M., Ostrowski, M., Soiferman, M., Younes, H., Younes, M.*, et al.* (2015). Odds ratio product of sleep EEG as a continuous measure of sleep state. *Sleep,* 38**,** 641-54.

Younes, M., Schweitzer, P. K., Griffin, K. S., Balshaw, R. & Walsh, J. K. (2020). Comparing two measures of sleep depth/intensity. *Sleep.*

Young, T., Finn, L., Peppard, P. E., Szklo-Coxe, M., Austin, D., *et al.* (2008). Sleep disordered breathing and mortality: eighteen-year follow-up of the Wisconsin sleep cohort. *Sleep,* 31**,** 1071-8.

Zare Sakhvidi, M. J., Zare Sakhvidi, F., Mehrparvar, A. H., Foraster, M. & Dadvand, P. (2018). Association between noise exposure and diabetes: A systematic review and meta-analysis. *Environ Res,* 166**,** 647-657.

Zeman, A., Britton, T., Douglas, N., Hansen, A., Hicks, J., *et al.* (2004). Narcolepsy and excessive daytime sleepiness. *BMJ,* 329**,** 724-8.

Zhang, D., Ding, H., Liu, Y., Zhou, C., Ding, H., *et al.* (2009). Neurodevelopment in newborns: a sample entropy analysis of electroencephalogram. *Physiol Meas,* 30**,** 491-504.

Zhang, H., Cisse, M., Dauphin, Y. N. & Lopez-Paz, D. 2018. MixUp: Beyond empirical risk minimization. *6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings.*

Zhang, J. J., Jin, X. T., Li, R. H., Gao, Y., Li, J. M., *et al.* (2019a). Influence of rapid eye movement sleep on all-cause mortality: a community-based cohort study. *Aging-Us,* 11**,** 1580-1588.

Zhang, Y., Ren, R., Lei, F., Zhou, J., Zhang, J., *et al.* (2019b). Worldwide and regional prevalence rates of co-occurrence of insomnia and insomnia symptoms with obstructive sleep apnea: A systematic review and meta-analysis. *Sleep Med Rev,* 45**,** 1-17.

Zhao, W., Van Someren, E. J. W., Li, C., Chen, X., Gui, W., *et al.* (2021). EEG spectral analysis in insomnia disorder: A systematic review and meta-analysis. *Sleep Med Rev,* 59**,** 101457.

Zhou, R., Cai, R. & Tong, G. (2013). Applications of Entropy in Finance: A Review. *Entropy,* 15**,** 4909-4931.

Zou, Y., Donner, R. V., Marwan, N., Donges, J. F. & Kurths, J. (2019). Complex network approaches to nonlinear time series analysis. *Physics Reports,* 787**,** 1-97.

APPENDIX A

Supplementary material: K-complexes are a sensitive marker of noise–related sensory processing during sleep: A pilot study

Table A1: Mean and SD of the number of noise instances played, stratified by sleep stages and sound pressure levels, for noise sensitive and non-sensitive participants.

	All	Non-sensitive	Sensitive	p-value
N	21	13	8	
Sleep stage 2				
23 dBA	30.48 (12.59)	29.77 (13.55)	31.62 (11.64)	0.752
33 dBA	26.38 (10.70)	25.23 (11.47)	28.25 (9.77)	0.544
36 dBA	28.05 (11.48)	26.85 (12.66)	30.00 (9.74)	0.555
39 dBA	26.38 (10.38)	25.31 (11.01)	28.12 (9.70)	0.559
42 dBA	26.95 (10.92)	25.38 (12.16)	29.50 (8.68)	0.416
45 dBA	27.29 (10.36)	26.15 (11.76)	29.12 (7.95)	0.537
48 dBA	26.52 (10.39)	25.46 (11.57)	28.25 (8.56)	0.564
Sleep stage 3				
23 dBA	24.86 (7.56)	23.77 (8.19)	26.62 (6.55)	0.415
33 dBA	20.62 (5.31)	19.77 (5.82)	22.00 (4.34)	0.363
36 dBA	19.14 (7.14)	18.85 (8.15)	19.62 (5.58)	0.815
39 dBA	21.33 (6.22)	20.46 (7.29)	22.75 (3.99)	0.427
42 dBA	20.43 (6.04)	20.00 (6.84)	21.12 (4.79)	0.689
45 dBA	20.76 (5.14)	20.77 (5.10)	20.75 (5.55)	0.994
48 dBA	20.62 (5.57)	20.54 (6.08)	20.75 (5.04)	0.935

Table A2: Mean and SD of the percentage of noise instances that evoked a K-complex, an arousal or an awakening in sleep stage 2.

	All	Noise non-sensitive	Noise sensitive	p-value
N	21	13	8	
KC, %				
23 dBA	7 (7)	8 (9)	5 (5)	0.344
33 dBA	9 (9)	10 (9)	6 (8)	0.253
36 dBA	12 (12)	13 (14)	9 (7)	0.383
39 dBA	15 (11)	14 (10)	17 (13)	0.571
42 dBA	17 (11)	20 (9)	11 (13)	0.055
45 dBA	20 (13)	22 (11)	17 (16)	0.388
48 dBA	22 (17)	28 (18)	14 (11)	0.06
Arousals, %				
23 dBA	1 (2)	1 (1)	1 (3)	0.673
33 dBA	1 (2)	1 (2)	1 (2)	0.8
36 dBA	1 (2)	1 (2)	1 (2)	0.502
39 dBA	2 (3)	2 (3)	3 (3)	0.687
42 dBA	3 (4)	3 (5)	3 (3)	0.782
45 dBA	3 (3)	4 (4)	2 (3)	0.109
48 dBA	5 (6)	5 (5)	5 (8)	0.905
Awakenings, %				
23 dBA	1 (3)	2 (3)	0 (1)	0.354
33 dBA	1 (2)	2 (3)	0 (1)	0.157
36 dBA	1 (2)	1 (2)	0 (0)	0.105
39 dBA	2 (3)	2 (2)	2 (4)	0.79
42 dBA	4 (5)	4 (7)	3 (3)	0.762
45 dBA	2 (3)	2 (3)	2 (3)	0.886
48 dBA	3 (4)	4 (5)	3 (3)	0.439

Table A3: Marginal probability of evoked K-complexes for different sound pressure levels and scoring thresholds (50, 60, 70, 80, 90, 95%) for automated K-complex detection.

	Sound pressure level, dBA						
	23	33	36	39	42	45	48
N2 sleep							
50	6 (4, 9)	10 (7, 14) #	12 (9, 17) #	14 (10, 19) #	18 (13, 24) #	21 (15, 27) #	23 (17, 30) #
60	5 (4, 8)	9 (6, 14) #	11 (8, 16) #	13 (9, 18) #	16 (12, 22) #	20 (15, 26) #	22 (16, 29) #
70	5 (3, 7)	9 (6, 13) #	10 (7, 17) #	12 (8, 17) #	15 (11, 21) #	19 (14, 26) #	21 (15, 28) #
80	4 (3, 7)	8 (5, 12) *	9 (6, 13) #	12 (8, 16) #	14 (10, 20) #	17 (13, 24) #	19 (14, 26) #
90	4 (2, 6)	6 (4, 9)	7 (4, 10) #	10 (7, 14) #	12 (8, 17) #	15 (11, 21) #	17 (12, 23) #
95	3 (2, 5)	5 (3, 7)	5 (3, 9)	9 (6, 13) #	10 (7, 15) #	14 (9, 20) #	15 (10, 21) #
N3 sleep							
50	22 (18, 27)	21 (17, 26)	25 (20, 31)	27 (23, 33)	22 (18, 27)	29 (24, 35)	32 (27, 38)
60	19 (16, 24)	19 (15, 23)	23 (18, 28)	26 (21, 32) *	20 (16, 26)	27 (22, 33) #	29 (24, 35) #
70	18 (14, 22)	18 (14, 22)	21 (16, 26)	24 (19, 29) *	18 (14, 23)	26 (21, 31) #	28 (23, 34) #
80	17 (13, 21)	16 (13, 21)	18 (14, 24)	21 (16, 26)	17 (13, 22)	23 (19, 29) #	26 (21, 31) #
90	13 (10, 18)	14 (10, 18)	14 (10, 19)	19 (15, 24) *	16 (12, 21)	21 (16, 27) #	21 (17, 27) #
95	11 (8, 15)	11 (8, 15)	13 (9, 17)	17 (13, 22)	14 (10, 18)	20 (16, 26) #	20 (16, 26) #

* $p < 0.05$; # $p<0.01$; 23 dBA (no noise control) is the comparator for p-values calculation

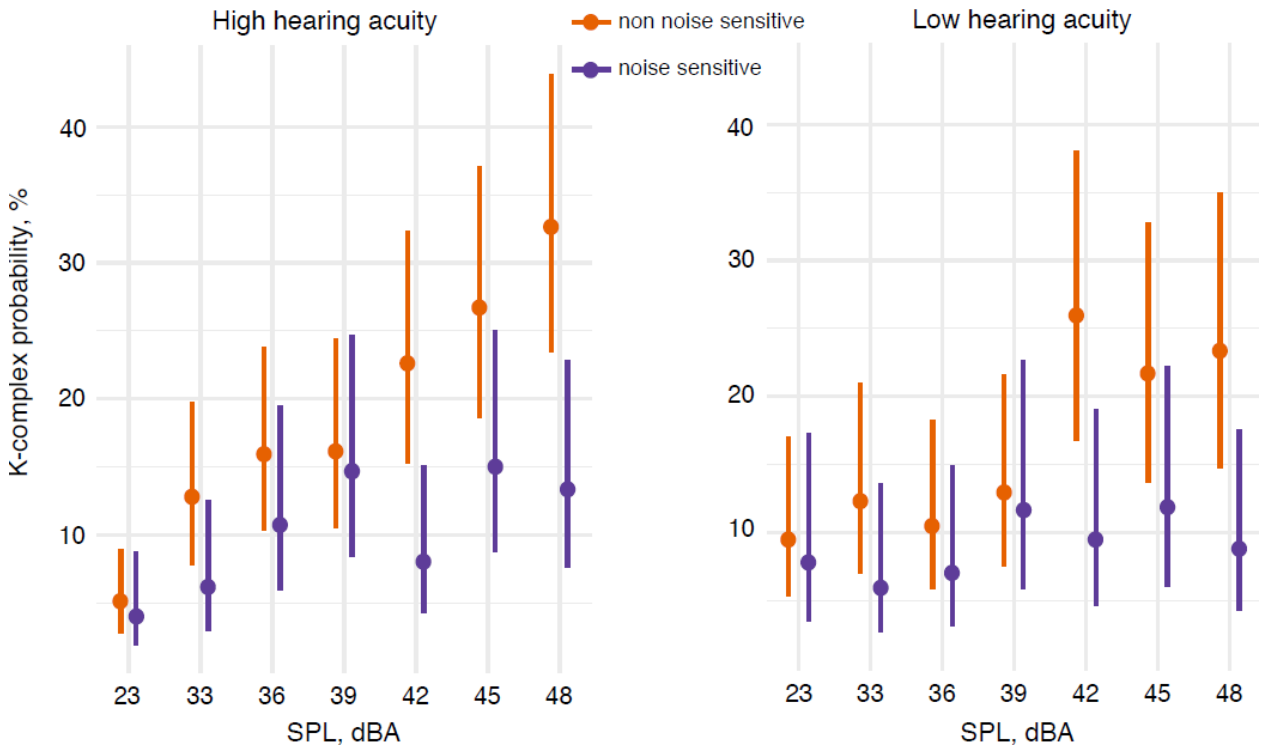

Figure A1: Marginal probability of the association between the KC-response in N2 sleep and SPL with high and low hearing acuity and noise sensitivity. Participants with hearing thresholds ≤ 3.9 dB HL and a Weinstein noise sensitivity score < 54 were considered to have high hearing acuity and to be non-noise-sensitive, respectively.

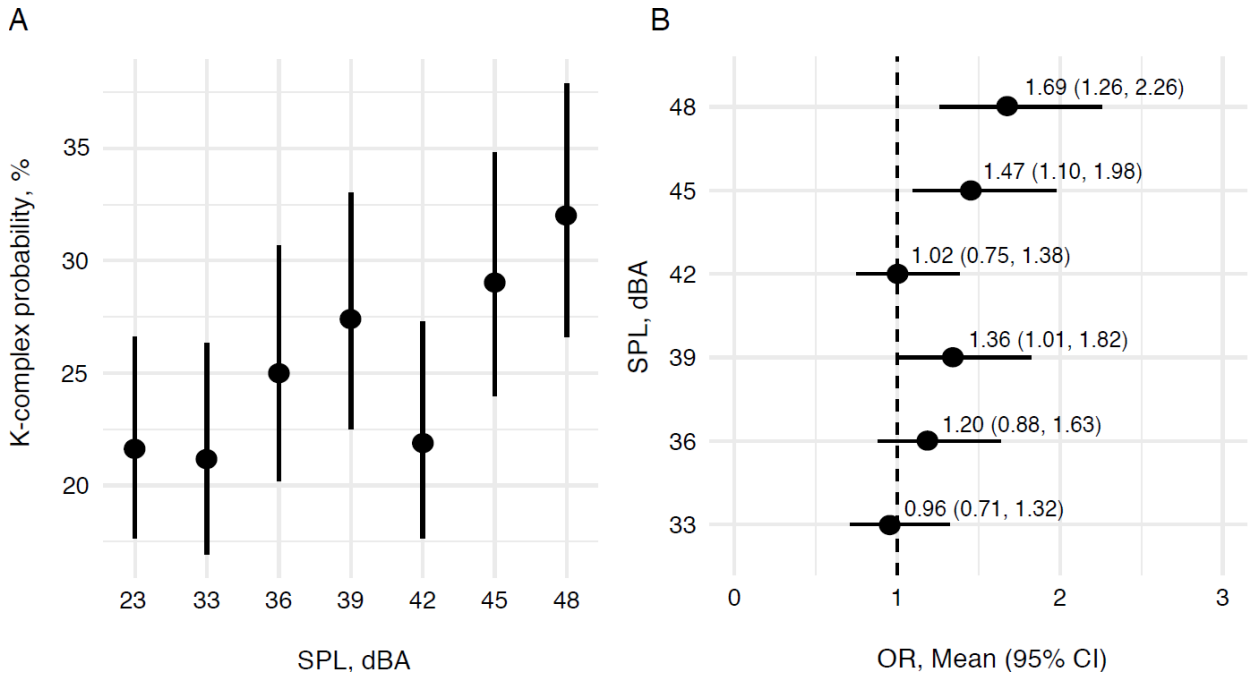

Figure A2: Association between noise SPL and the K-complex response in N3 sleep. A) Probability of occurrence of a K-complex at a given noise level. B) Odds ratio (95% CI) of evoking a K-complex at a given SPL compared to background noise.

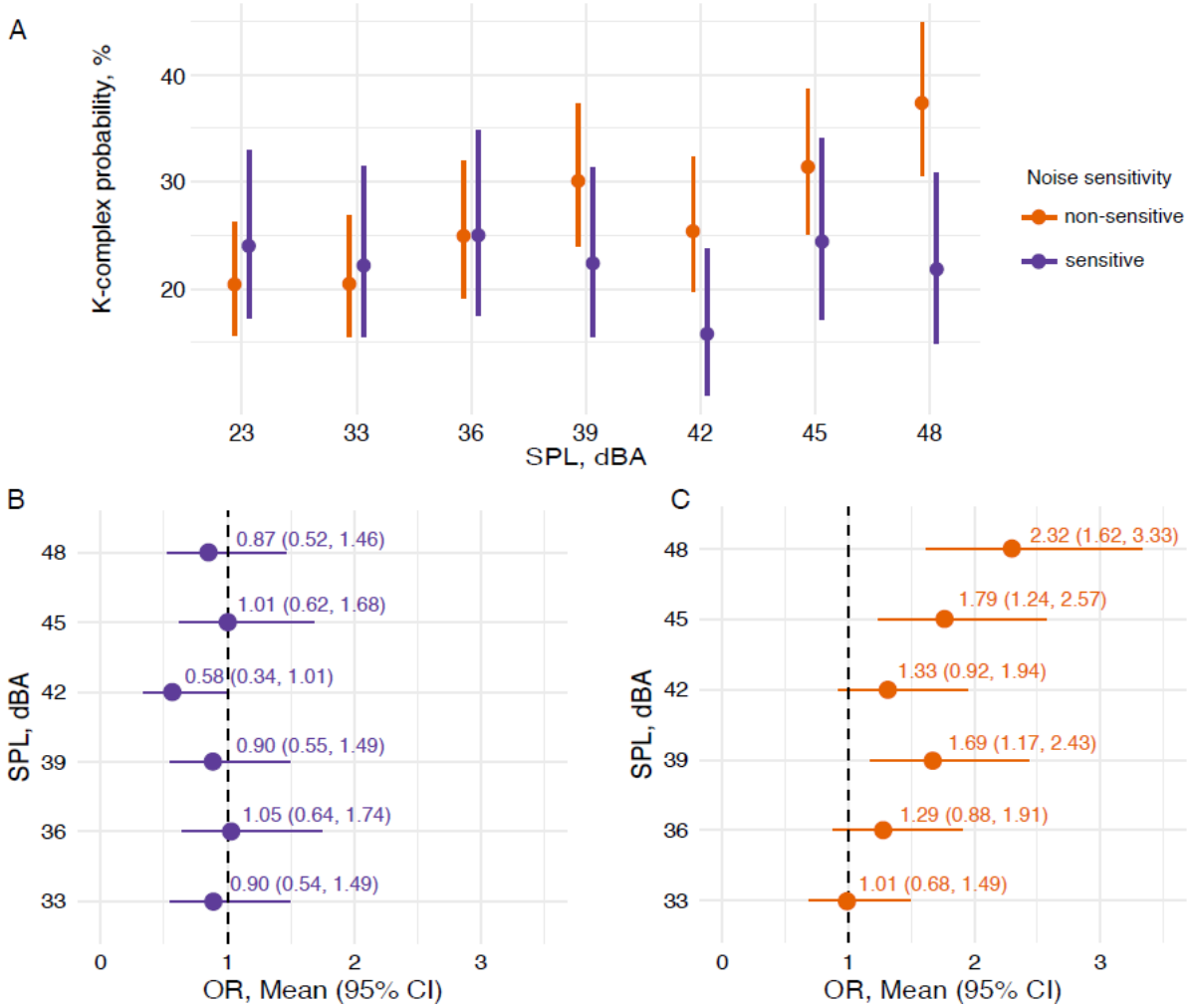

Figure A3: Association between noise SPL and the K-complex response in N 3 sleep for noise non-sensitive (orange) and sensitive (purple) group. A) Probability of occurrence of a K-complex at a given noise level. B), C) Odds ratio (95% CI) of evoking a K-complex at a given SPL compared to background noise.

APPENDIX B

Supplementary material: A novel EEG derived measure of fragmentation of delta wave activity during sleep predicts all-cause mortality

B.1 Methodology

EEG power spectral analysis

The multi-taper technique (Prerau et al., 2017) was used, which minimizes the uncertainty in the spectral estimate across frequencies by multiplying the original signal with multiple orthogonal windows called tapers. This approach reduces windowing artifacts compared to traditional quantitative EEG. The EEG power spectral analysis was primarily performed on the C3 channel with a signal quality score, recorded by the human-expert scorer, of $\geq$3 indicating that at least 50% of the EEG signals were artefact-free. The C4 channel was used in sensitivity analysis. EEG data were re-sampled to 128 Hz, and band-pass filtered between 0.3 and 35 Hz. Processed raw data were segmented into five sec non-overlapping segments on which the multi-taper power spectral analysis was performed with a 1 Hz resolution using four tapers, the results from which were then ensemble averaged. Technical details regarding the multi-taper method can be found in Prerau et al. (2017). The absolute power was calculated for each 5-second window in delta, theta, alpha, sigma and beta frequency bands (0.5 - 4.5, 4.5 - 8, 8 - 12, 12 - 15, and 15 - 32 Hz, respectively).

Spectral entropy

The variation of absolute power in the delta frequency band was captured using a density function. The average power of wake and sleep stage 1 (but not REM) was set to zero, since slow wave activity predominantly occurs in sleep stage 2 and 3 and to help reduce movement artefacts in wake and light transitional sleep. A weighted moving average Gaussian window (120 points, standard deviation = 10) was applied to a density function to reduce stochastic noise. The shape of the final delta wave density function was quantified using spectral entropy, an information measure that determines the degree of uniformity of the distribution (Inouye et al., 1991). The spectral entropy was calculated as the Shannon entropy (using Eq B1)

$$PSE = -\sum_{f=0}^{fs/2} p_f \ln p_f \quad \text{Eq. B1}$$

Where PSE represents the power spectral entropy, and p_f the probability density function of the power spectral density $P(f)$.

$$p_f = \frac{P(f)}{\sum_f P(f)} \quad \text{Eq. B2}$$

$$P(f) = \frac{1}{N}|X(f)|^2 \quad \text{Eq. B3}$$

Where $X(f)$ is the Fourier transform of the delta wave density function. The implementation of the spectral entropy in this thesis was based on the one available in the package "mne-features" and "Entropy". The implementation is also available on Github (see Open source software section).

B.2 Sleep fragmentation and all-cause mortality

Kaplan-Meier survival curves (Figure B1) show a lower survival probability for participants within the lowest and highest tertile of spectral entropy (compared to the middle tertile).

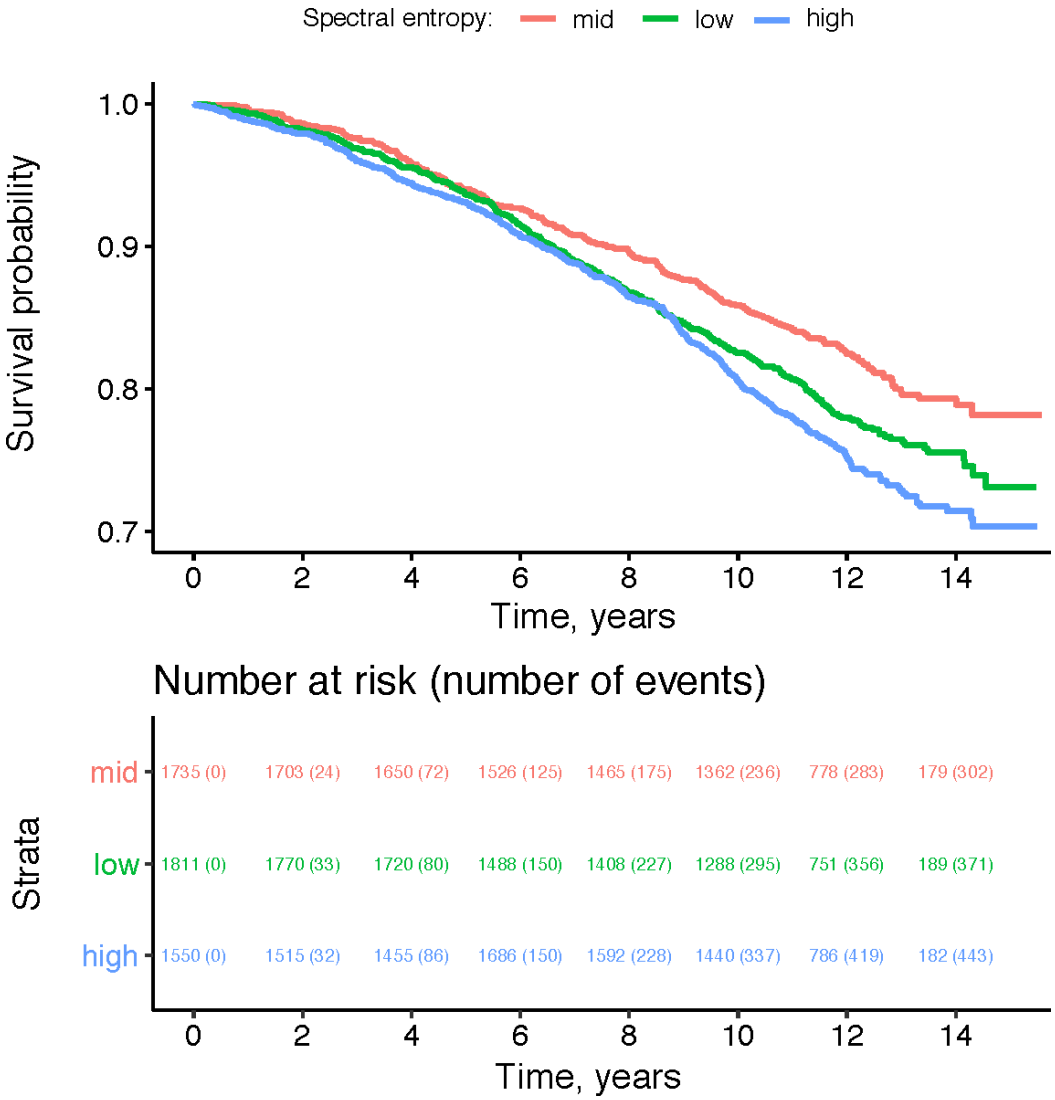

Figure B1: Kaplan-Meier survival curves for tertiles of the spectral entropy.

Four sensitivity analysis were conducted. First, participants with CVD at baseline (Sensitivity analysis 1; S1) were excluded. Second, participants who died in the first three years were removed (S2). Third, the EEG quality was elevated to at least 75% of artefact-free EEG signals (S3). Finally (S4), the analysis was repeated on C4 EEG channel, to help test if the shape of the delta-density function on C3 and C4 were similar. Hazard ratio and 95%CI of the associations between fragmented delta wave activity and all-cause mortality are shown in Table B1.

Table B1: Summary statistics of model 1 (spectral entropy) sensitivity analysis.

		S1	S2	S3	S4
N		4912	4255	3519	5245
No		965	750	804	1155
HR (95% CI)	5 vs 50th	1.28 (1.09, 1.49)	1.33 (1.12, 1.59)	1.28 (1.06, 1.54)	1.31 (1.15, 1.49)
	95 vs 50th	1.02 (0.88, 1.17)	1.01 (0.88, 1.17)	1.10 (0.94, 1.29)	1.03 (0.91, 1.17)
D-index		0.58	0.56	0.59	0.60
C-index		0.79	0.78	0.80	0.80
LR test	χ^2	9.31	9.35	8.05	15.7
	p-values	0.025	0.025	0.045	0.001

N, population number; No, number of outcomes; LR, likelihood-ratio
S1, removing participants with $<$ 3 years follow up
S2, removing participants with CVD at baseline
S3, participants with at least $\geq$ 75% of artefact-free EEG
S4, full model with spectral entropy calculated on C4

Stratified analyses

Hazard ratios for sex-stratified and age-stratified associations between low spectral entropy (5th vs 50th percentiles) and all-cause mortality risk are summarized in Table B2. The interaction term between age (continuous) and spectral entropy was not significant ($p\text{-}value = 0.36$), whereas the interaction term between sex and spectral entropy was significant ($p\text{-}value = 0.02$).

Table B2: Association between spectral entropy and all-cause mortality risk (5^{th} vs 50^{th}) for sex- and age-stratified models.

		Age		
		$<$ 70 y	$>$ 70 y	All
Sex	Male	0.78 (0.46, 1.28)	1.50 (1.15, 1.95)	1.24 (1.00, 1.53)
	Female	1.91 (1.10, 3.31)	1.27 (1.03, 1.58)	1.41 (1.17, 1.70)
	All	1.23 (0.91, 1.67)	1.34 (1.14, 1.58)	

B.3 Sleep fragmentation and sleep disordered breathing

AHI was not associated with mortality-risk (Figure B2) and there was no interaction between spectral entropy and AHI ($\chi^2 = 1.9$, $p = 0.16$). However, the percentage of total sleep time with oxygen saturation less than 90% was associated with an increase in mortality-risk (18.2% vs 0.2 %; 1.24 (1.13, 1.36))

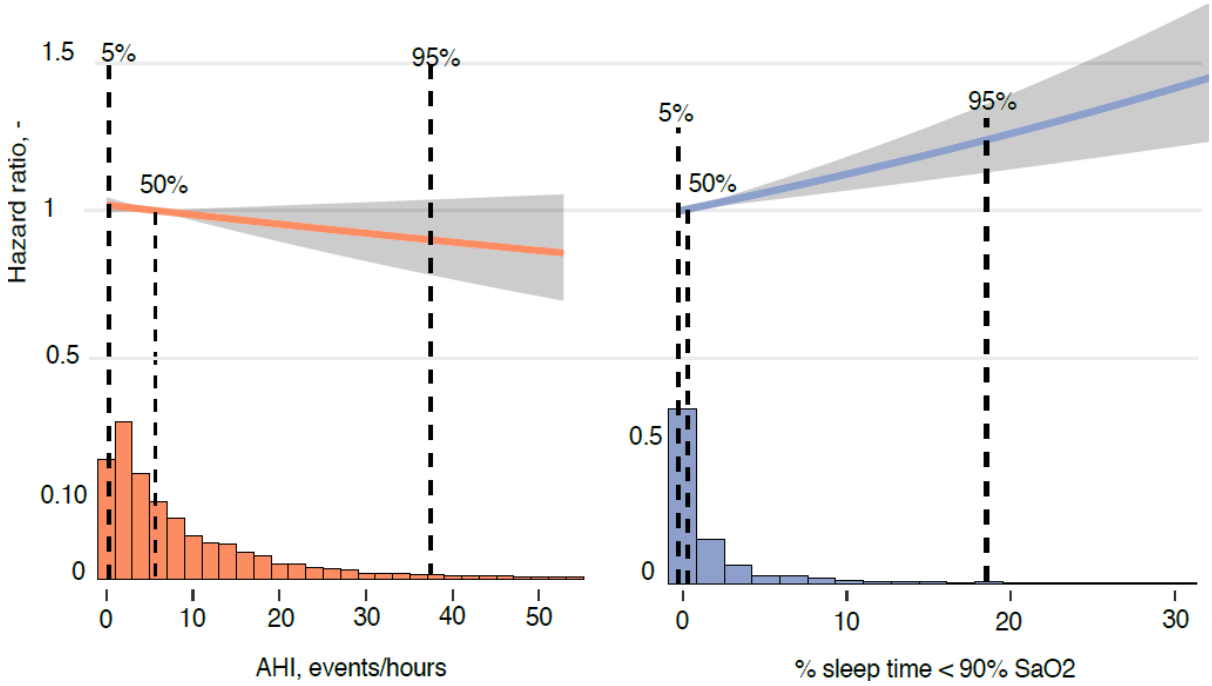

Figure B2: Association between (A) AHI and (B) percent of total sleep time with oxygen saturation less than 90% of oxygen saturation and all-cause mortality risk.

Stratified analyses and interactions

The interaction effect (using quartile of % of sleep time spent with oxygen saturation less than 90%) is shown in Figure B3.

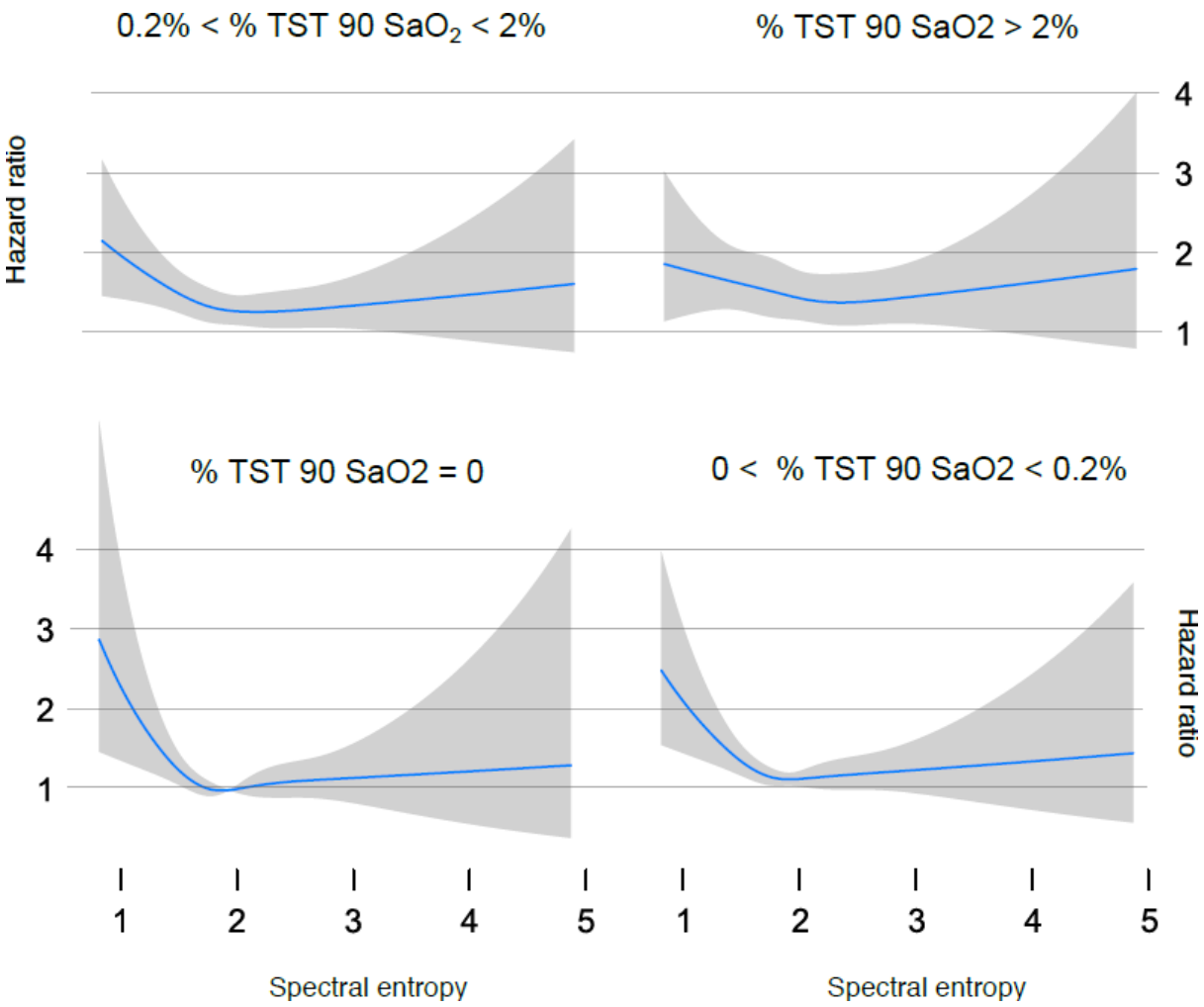

Figure B3: Association between spectral entropy and all-cause mortality risk for quartile-based (with interaction term) subgroup analysis of percent of total sleep time spent with oxygen saturation less than 90% models.

B.4 Traditional polysomnography markers of sleep quality and all-cause mortality

Non-linear associations were tested but none were found for any for traditional polysomnography markers of sleep quality, wake after sleep onset and arousal index. The association between arousal index (75^{th} percentile against 25^{th} percentile, HR, 95% CI, 1.04 (0.98, 1.10), p-value = 0.23) and wake after sleep onset (1.04 (1.00, 1.09), p-value = 0.07) with all-cause mortality were not significant (Figure B4).

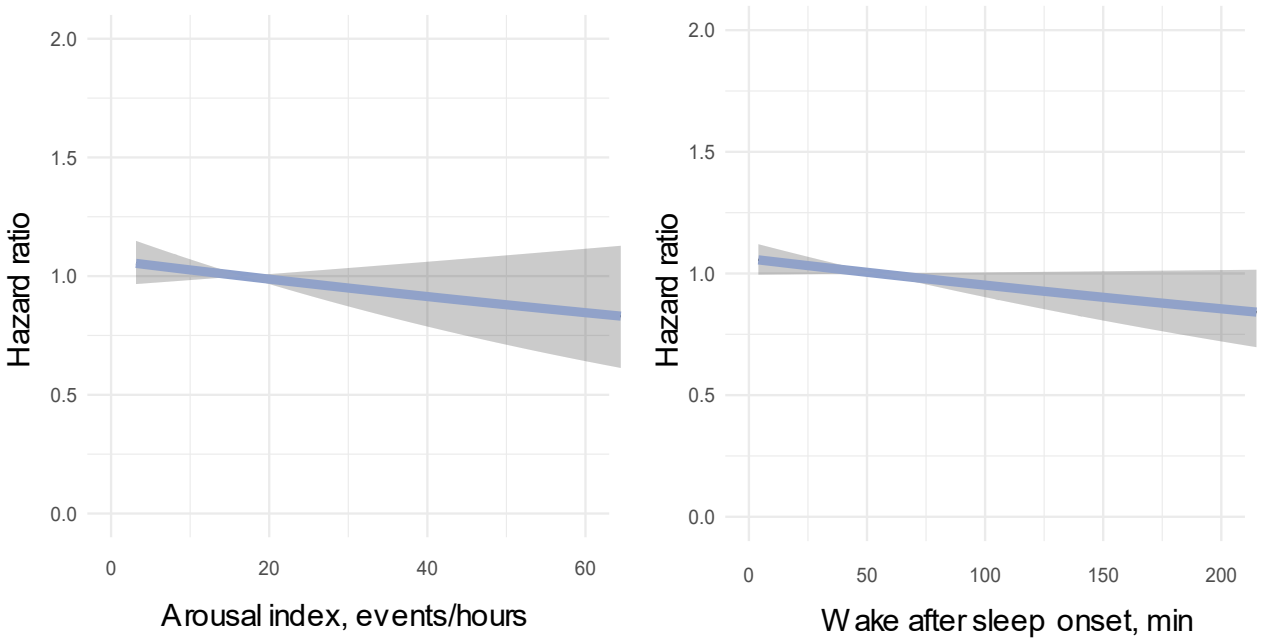

Figure B4: Association between the arousal index and wake after sleep onset and all-cause mortality.

Quantitative EEG and all-cause mortality

The association between the absolute power of each bands and all-cause mortality is shown in Table B3 and compares the 75th percentile of the population to the 25th using the hazard ratio and 95% confidence interval.

Table B3: Association between mean absolute power of typical frequency bands in NREM sleep and all-cause mortality risk (75th vs 25th).

Band power	HR, 95% CI
Delta	1.00 (0.94, 1.06)
Theta	1.00 (0.95, 1.02)
Alpha	0.96 (0.90, 1.02)
Sigma	1.00 (0.99, 1.01)
Beta	1.00 (0.99, 1.01)

B.5 Open source software

The code used in this thesis has been made freely available under a common license rule at https://github.com/Adelaide-Institute-for-Sleep-Health/ upon

publication of Chapter 4. Figure B5 is an overview of the user interface/dashboard developed to facilitate uptake of the tools by sleep practitioners.

Figure B5: Software layout. Top, summary metrics of a given dataset, including total sleep time (TST), total recording time (TRT), sleep efficiency and information's about missing files. Bottom, analysis summary quantitative EEG is described in chapter 4.

APPENDIX C

Supplementary material: Co-morbid insomnia and obstructive sleep apnea is associated with all-cause mortality.

C.1 All-cause mortality

Sex-stratified models (Table C1) showed significantly higher risk of all-cause mortality among males in co-morbid insomnia and sleep apnoea (COMISA) (1.57 (1.14, 2.16)) and insomnia-alone participants (1.76 (1.12, 2.78)) compared to the reference group. This effect was strongest in males aged over 70 years (Table C2). However, limited statistical power associated with the low number of events in the younger age category (N = 219) may limit interpretation of these comparisons. No significant associations of sleep disorder category and mortality were found in women after adjustment of all covariates (Table C1), irrespective of age category (Table C2). These exploratory analyses should be interpreted with caution, due to the reduction in sample size and statistical power to detect significant associations.

For Table D1 to D6, the quoted values are Hazards ratio (and 95% CI) against the reference group (no insomnia/OSA). Model 1 is unadjusted. Model 2 is adjusted for demographics and anthropometrics (age, BMI, race, sex), behavioural (smoking status) and polysomnography-measured total sleep time. Model 3 is additionally adjusted for pre-existing cardio-metabolic conditions, including diabetes, CVD, hypertension, lipid lowering medications and COPD. Finally, Model 4 is additionally adjusted for use of benzodiazepines, tricyclic anti-depressants and a binary variable constructed representing participants taking sleep medication more than 5 times a month.

Figure C1: Association between tertiles (T1, T2 and T3) of % of time spent with less than 90% of oxygen saturation (TST90) and insomnia (Ins) with all-cause mortality.

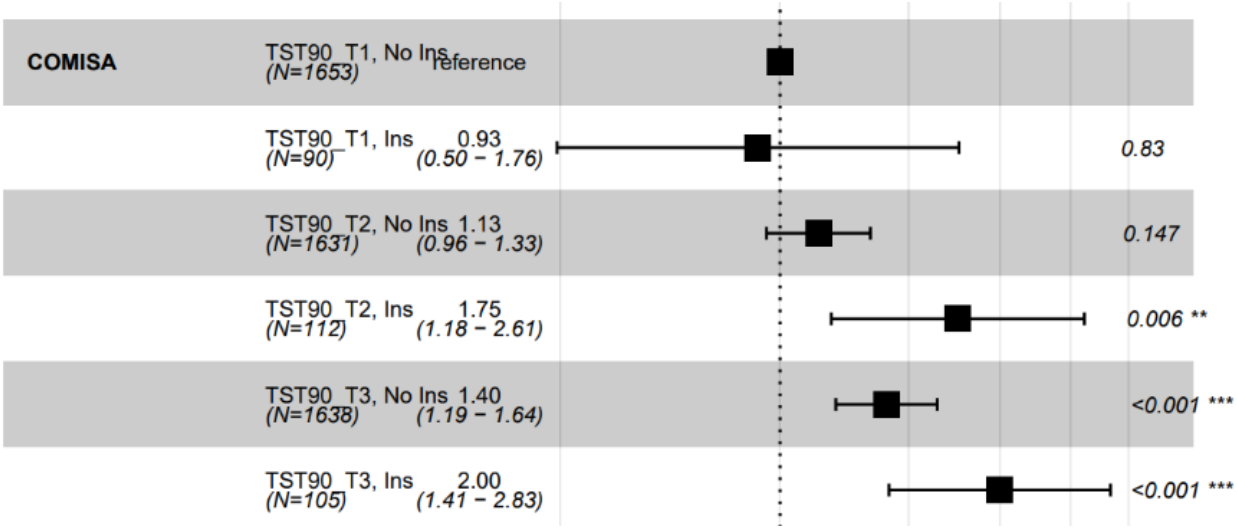

Table C1: Adjusted associations between sleep disorder group and all-cause mortality stratified by sex.

	N	N event	Insomnia	OSA	COMISA
Men					
Model 1	2435	589	1.45 (0.92, 2.28)	1.36 (1.13, 1.63)	1.58 (1.16, 2.15)
Model 2	2412	584	1.66 (1.05, 2.61)	1.12 (0.93, 1.35)	1.53 (1.12, 2.10)
Model 3	2210	558	1.76 (1.12, 2.78)	1.13 (0.93, 1.37)	1.57 (1.14, 2.16)
Model 4	2207	557	1.76 (1.12, 2.78)	1.13 (0.93, 1.37)	1.57 (1.14, 2.16)
Women					
Model 1	2705	525	1.00 (0.74, 1.37)	1.47 (1.22, 1.79)	1.81 (1.30, 2.52)
Model 2	2681	518	1.07 (0.79, 1.46)	0.91 (0.75, 1.11)	1.15 (0.82, 1.61)
Model 3	2517	499	0.94 (0.68, 1.30)	0.89 (0.73, 1.09)	1.03 (0.73, 1.44)
Model 4	2513	498	0.91 (0.66, 1.28)	0.90 (0.74, 1.11)	1.02 (0.72, 1.44)

Table C2: Fully adjusted associations between sleep disorder group and all-cause mortality stratified by sex and age.

	N	N event	Insomnia	OSA	COMISA
Men					
$<$ 70 yo	1565	219	1.19 (0.51, 2.80)	1.13 (0.84, 1.54)	1.45 (0.90, 2.33)
$>$ 70 yo	643	339	2.15 (1.23, 3.72)	1.17 (0.91, 1.51)	1.58 (1.01, 2.49)
Women					
$<$ 70 yo	1694	137	0.71 (0.38, 1.31)	0.91 (0.60, 1.38)	0.98 (0.57, 2.05)
$>$ 70 yo	822	361	1.02 (0.69, 1.52)	0.90 (0.72, 1.15)	1.08 (0.73, 1.61)

Table C3: Adjusted associations between sleep disorder group and all-cause mortality, excluding participants that died within the first two years (N = 96).

	N	N event	Insomnia	OSA	COMISA
Model 1	5140	1114	1.07 (0.84, 1.39)	1.47 (1.29, 1.66)	1.73 (1.38, 2.17)
Model 2	5093	1102	1.24 (0.96, 1.60)	1.02 (0.89, 1.17)	1.34 (1.06, 1.68)
Model 3	4727	1057	1.16 (0.89, 1.51)	1.02 (0.89, 1.17)	1.29 (1.02, 1.63)
Model 4	4720	1055	1.14 (0.87, 1.50)	1.02 (0.89, 1.17)	1.29 (1.02, 1.63)

Table C4: Adjusted associations between sleep disorder group and all-cause mortality. Insomnia was defined as difficulties initiating and maintaining sleep at least 15 times a month and symptoms of daytime impairment.

	N	N event	Insomnia	OSA	COMISA
Model 1	5236	1210	1.06 (0.74, 1.50)	1.47 (1.31, 1.65)	1.77 (1.29, 2.42)
Model 2	5189	1198	1.20 (0.84, 1.71)	1.01 (0.90, 1.14)	1.51 (1.10, 2.07)
Model 3	4822	1152	1.07 (0.73, 1.55)	1.00 (0.88, 1.14)	1.51 (1.10, 2.09)
Model 4	4815	1150	1.05 (0.72, 1.52)	1.01 (0.89, 1.14)	1.49 (1.08, 2.07)

C.2 Cardiovascular disease

Sex-stratified associations of sleep disorder category and cardiovascular events are displayed in Table C5. The strongest association of COMISA and cardiovascular events occurred in older men (>70 years old; Table C6). Furthermore, older men (> 70 years old) with OSA-alone showed a 36% increase in cardiovascular event risk versus older men with no insomnia or OSA (reference) (1.36 (1.05, 1.77)). Conversely, no significant associations were found in women with insomnia-alone, OSA-alone or COMISA once adjusted for all pre-specified covariates, irrespective of age category.

Table C5: Adjusted associations between sleep disorder group and cardiovascular disease stratified by sex.

	N	N event	Insomnia	OSA	COMISA
Men					
Model 1	2123	702	1.28 (0.82, 2.00)	1.41 (1.20, 1.67)	1.70 (1.28, 2.26)
Model 2	2111	700	1.24 (0.79, 1.94)	1.17 (0.99, 1.39)	1.48 (1.11, 1.98)
Model 3	2051	687	1.25 (0.79, 1.96)	1.17 (0.99, 1.40)	1.31 (0.98, 1.77)
Model 4	2049	685	1.22 (0.77, 1.86)	1.17 (0.99, 1.39)	1.29 (0.96, 1.73)
Women					
Model 1	2452	541	0.99 (0.72, 1.35)	1.54 (1.28, 1.86)	2.40 (1.76, 3.26)
Model 2	2436	539	0.98 (0.72, 1.34)	0.98 (0.81, 1.19)	1.46 (1.07, 2.00)
Model 3	2367	528	0.86 (0.62, 1.18)	0.96 (0.79, 1.18)	1.24 (0.90, 1.71)
Model 4	2366	527	0.84 (0.61, 1.18)	0.97 (0.80, 1.19)	1.26 (0.91, 1.73)

Table C6: Fully adjusted associations between sleep disorder group and cardiovascular event stratified by sex and age.

	N	N event	Insomnia	OSA	COMISA
Men					
< 70 yo	1328	313	0.98 (0.48, 2.02)	1.05 (0.82, 1.36)	1.14 (0.75, 1.72)
> 70 yo	643	346	1.69 (0.91, 3.10)	1.36 (1.05, 1.77)	1.60 (1.02, 2.51)
Women					
<70 yo	1539	162	0.79 (0.45, 1.40)	1.12 (0.77, 1.63)	1.32 (0.67, 2.6)
> 70 yo	824	365	0.95 (0.63, 1.43)	0.95 (0.75, 1.20)	1.33 (0.92, 1.94)

C.3 Associations between sleep disorder group and cardiovascular event subtypes

Associations between cardiovascular event subtypes and sleep disorders suggest that COMISA and OSA-alone are associated with each cardiovascular event subtypes in the unadjusted models (CHD, heart failure and stroke; Table C7). In the fully adjusted models, COMISA was associated with a 30% increase in likelihood of heart failure compared to the reference group (HR 95% CI, 1.30 (0.96, 1.77)) and OSA-alone was associated with a 22% (1.22 (1.04,1.43)) increase in CHD events compared to the reference group.

Table C7: Associations between sleep disorder group and cardiovascular event subtypes.

		N	N events	Insomnia	OSA	COMISA
Stroke						
All	Model 1	4575	271	1.10 (0.67, 1.90)	1.5 (1.15, 1.90)	1.9 (1.23, 3.0)
	Model 2	4434	266	1.00 (0.60, 1.70)	1.11 (0.85, 1.5)	1.28 (0.81, 2.0)
Men	Model 1	2123	114	1.26 (0.38, 4.11)	1.76 (1.15, 2.70)	1.73 (0.82, 3.15)
	Model 2	2056	111	1.13 (0.34, 3.74)	1.35 (0.87, 2.10)	1.29 (0.60, 2.77)
Women	Model 1	2452	157	1.02 (0.58, 1.81)	1.48 (1.04, 2.10)	2.37 (1.34, 4.18)
	Model 2	2378	155	0.97 (0.54, 1.72)	0.93 (0.65, 1.34)	1.18 (0.66, 2.13)
Coronary heart disease						
All	Model 1	4575	765	0.91 (0.65, 1.27)	1.63 (1.40, 1.90)	1.84 (1.40, 2.42)
	Model 2	4434	751	0.85 (0.61, 1.19)	1.22 (1.04, 1.43)	1.23 (0.93, 1.63)
Men	Model 1	2123	501	1.23 (0.72, 2.09)	1.38 (1.13, 1.67)	1.63 (1.16, 2.28)
	Model 2	2056	493	1.17 (0.69, 2.00)	1.15 (0.94, 1.41)	1.27 (0.90, 1.80)
Women	Model 1	2452	264	0.95 (0.62, 1.46)	1.26 (0.96, 1.66)	1.60 (0.98, 2.58)
	Model 2	2378	258	0.85 (0.55, 1.33)	0.79 (0.59, 1.06)	0.86 (0.53, 1.39)
Heart failure						
All	Model 1	4575	568	1.32 (0.94, 1.87)	1.71 (1.42, 2.05)	2.35 (1.74, 3.17)
	Model 2	4434	561	1.09 (0.77, 1.55)	1.08 (0.89, 1.30)	1.30 (0.96, 1.77)
Men	Model 1	2123	302	1.39 (0.70, 2.77)	1.60 (1.23, 2.07)	1.91 (1.24, 2.93)
	Model 2	2056	299	1.22 (0.61, 2.44)	1.09 (0.83, 1.42)	1.24 (0.80, 1.92)
Women	Model 1	2452	266	1.37 (0.91, 2.06)	1.64 (1.25, 2.15)	2.77 (1.82, 4.2)
	Model 2	2378	262	1.13 (0.74, 1.72)	0.92 (0.69, 1.22)	1.24 (0.81, 1.90)

Model 1: Unadjusted; Model 2: age, BMI, race, smoking status, cardiovascular disease (list all), hypertension, diabetes. COMISA = co-morbid insomnia and sleep apnoea, OSA = obstructive sleep apnoea. Quoted values are Hazards ratio (and 95% CI) against the reference group (no insomnia/OSA). COMISA = co-morbid insomnia and sleep apnoea, OSA = obstructive sleep apnoea.

For Table C8 to C9, the confounders are the same as the one enumerated for Table C1 to C6.

Table C8: Sensitivity analysis on the adjusted associations between sleep disorder group and incident cardiovascular events. Participants with baseline cardiovascular disease (N = 792) were removed from this analysis.

	N	N event	Insomnia	OSA	COMISA
Model 1	3769	794	1.10 (0.82, 1.48)	1.46 (1.26, 1.70)	1.93 (1.46, 2.6)
Model 2	3747	792	1.28 (0.96, 1.73)	0.99 (0.85, 1.16)	1.34 (1.01, 1.78)
Model 3	3638	775	1.19 (0.88, 1.61)	0.97 (0.83, 1.14)	1.24 (0.93, 1.64)
Model 4	3634	774	1.17 (0.86, 1.61)	0.97 (0.83, 1.15)	1.25 (0.94, 1.67)

Table C9: Sensitivity analysis on the adjusted associations between sleep disorder group and incident cardiovascular events. Insomnia was defined as difficulty in initiating and maintaining sleep at least 15 times a month and symptoms of daytime impairment.

	N	N event	Insomnia	OSA	COMISA
Model 1	4575	1243	1.09 (0.77, 1.55)	1.68 (1.50, 1.89)	1.96 (1.43, 2.70)
Model 2	4547	1239	1.13 (0.80, 1.60)	1.13 (1.00, 1.27)	1.41 (1.02, 1.94)
Model 3	4421	1216	1.03 (0.75, 1.48)	1.11 (0.98, 1.26)	1.36 (0.99, 1.88)
Model 4	4415	1213	1.01 (0.70, 1.45)	1.12 (0.99, 1.27)	1.35 (0.97, 1.87)