Measures of Sleep Duration and Quality in Sri Lanka

Aaron Schokman

Faculty of Medicine

The University of Sydney

A thesis submitted in fulfilment of the requirements for the degree of Master of Philosophy

This is to certify that to the best of my knowledge; the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

Signature:

Name: Aaron Shane Schokman

Abstract

A link between poor sleep and abnormal health outcomes has been established.The majority of this research is conducted in Western, high-income country (HIC) settings which warrants the question of how relevant the findings and the derived consensus statements are to low-middle income countries (LMICs). Sleep duration and quality are known to be affected by cultural, social, environmental and geographical influences which vary significantly between LMIC and their HIC counterparts. This thesis provides a thorough examination of objective and subjective measurements of sleep in a LMIC setting. Objective and subjective measures of sleep were obtained as part of an actigraphic sub-study nested within the Colombo twin and singleton study, a cohort study of twins and singletons randomly selected from Colombo, Sri Lanka. Results comparing actigraphic and self-reported sleep duration indicated that Sri Lankans have short sleep duration; averaging 6.4h (SD 1.5) self-reported and 6.0h (SD 0.9) actigraphically. Poor sleep quality was prevalent with an average WASO of 49 min, and sleep efficiency <85%. Bias was observed, with self-report consistently over-reporting sleep on average by 27.6 min (95% CI: -0.68, -0.24) compared to objective measures, but wide individual variation in disagreement, ranging from over-reporting by 3.34h to under-reporting by 2.42h. Agreement between subjective and objective measurements of sleep quality was also assessed. Objective indices of sleep efficiency, sleep onset latency and wake after sleep onset did not agree with each other and were unable to be combined into more useful composite indices as a 'global' measure of an individual's satisfaction with their sleep. Given the high cardiometabolic morbidity in Sri Lanka and poor measurement agreement observed, this warrants further investigation and supports the need for culturally appropriate, reliable, and valid assessment for analytic epidemiology in non-Western settings.

Acknowledgements

I would like to start off by thanking my supervisor Professor Nick Glozier for his guidance and mentoring over the course of this degree. My ability to think logically, critically analyse and convey it in writing has improved markedly and is a testament to your expertise and skill. I'd also like to extend my thanks to Yu Sun Bin, whom always went above and beyond what was necessary whether it be providing feedback on writing or offering support, Guido Simonelli for his guidance in submitting my first journal article and Molly Goldrick and Mark Bourke for proofreading the final edit of this thesis. I would also like to acknowledge all those whom worked on the CoTASS project, without your efforts, I would never have been offered the amazing opportunity to research the country my father migrated from.

On a personal note, I'd like to take the opportunity to convey my heartfelt thanks to my grandfather Felix Muga and aunt Barbara Ratynayke for their unwavering support, confidence and for just being there in my darkest of days. I'd also like to thank my parents for their continuing support over the years.

Table of Contents

Table of commonly used abbreviations

CI	**Confidence Interval**
HIC	**High - Income Countries**
LMIC	**Low – Middle Income Countries**
LOA	**Limit of Agreement**
PSQI	**Pittsburg Sleep Quality Index**
REM	**Rapid Eye Movement**
SE	**Sleep Efficiency**
SOL	**Sleep Onset Latency**
SWS	**Slow Wave Sleep**
TST	**Total Sleep Time**
TiB	**Time in Bed**
WASO	**Wake After Sleep Onset**

Overview of the Thesis

This thesis is concerned with the global perspective of sleep. The first chapter is a brief overview of current research within the field of sleep and gives context to novel research presented in chapters 2 & 3.

The first chapter is broken into key research in the field of sleep, summarised in section 1.1 focuses on the importance of sleep duration and quality. In section 1.2, the way sleep is measured, the different and evolving methodology involved, and the accuracy and reliability of these methods are reviewed. Global perspectives on sleep and the variation of sleep by level of economic development (i.e. High-Income Country vs. Low-Middle Income Country) is reviewed in section 1.3.

Chapter 2 of this thesis is published as 'Agreement Between Subjective and Objective Measures of Sleep Duration in a Low-Middle Income Country Setting' in the special edition of Sleep Health Journal titled 'Global perspectives of Sleep'. A copy of the article has been added to the appendix of this thesis (see appendix A) as Chapter 2 contains minor edits to the original article. A self-contained introduction, methods, results, discussion, conclusion, and appendix are included from pages 14 to 34 and should be treated as such. I was the primary author of this publication where I analysed, annotated and coded raw actigraphy data and self-reported measures of sleep collected in the Colombo Twin and Singleton Study (CoTASS) 2. I did not play a role in the design or collection of data associated with the COTASS 2 study and as such the efforts of those associated with the CoTASS 2 study must be acknowledged: Athula Sumathipala, Sisira Siribaddana, Mathew Hotopf, Peter McGuffin, Nick Glozier, Harriet Ball, Yulia Kovas, Fruhling Rijsdijk, Lalani Yatawara, Carmine Pariante, Helena Zavos, Chesmal Siriwardhana, Gayani Pannala, Kaushalya Jayaweera, Anushka Adikari and Dinesha Gunewardane.

Chapter 3 focuses on novel research into sleep quality and the creation of a composite variable based on objective indices of sleep quality. An examination of agreement between objective indices of sleep quality against subjective measures is also included.

A conclusion and future work in the field are discussed in chapter 4.

Chapter 1 - Introduction

1.1 – Sleep Through the Ages

The importance of sleep has always been recognised and respected throughout history. Many historical accounts exist of the unique spiritual and divine role that sleep held in society, with both Ancient Greeks and Romans consulting dream oracles [1]. Sleep has also been traditionally associated with healing, with Shakespeare recognises the healing properties of sleep in Henry IV – “O Sleep! O gentle sleep! Nature’s soft nurse”. Only recently has the way society has treated sleep changed; an activity that is viewed as frivolous and a waste of time. Our sleep patterns have also changed. Roger Ekirch in his book, ‘At Day's Close: Night in Times Past’ (2006), provides more than 500 historical references ranging from Homers Odyssey to medieval literature, evidence that our ancestors engaged in segmented sleeping patterns [2]. One common example is a biphasic sleep schedule, the practice of sleeping in two distinctly separate periods over 24 hours. This is in contrast to current recommendations of sleep, where a monophasic sleep schedule has been universally adopted (i.e. 8 hours of recommended sleep a night). These changes come at a time of rapid technological advancement, where light pollution, ease of travel and the conversion into a ‘24-hour seven day a week’ economy have slowly reduced the quality of sleep the average person obtains [3].

1.1.1 – The Sleep-Health Link

Sleep is one of few basic needs necessary for survival. It is a finely tuned system involving complex, interwoven and cascading mechanisms with distinct stages generating unique brain activity. While some fundamental questions surrounding the how and why we sleep remain elusive, interest in the field has exploded over the last 20 years.

There is compelling evidence that sleep and health are intrinsically linked. A strong U-shaped association has been shown to exist between sleep duration and optimal physical health across many different cohorts[4]. Individuals with short sleep duration (i.e. those that sleep less than what is needed for optimal health) have been shown to be at increased risk of cardiovascular disease [5], type 2 diabetes [6], hypertension, respiratory disorders and have distinct changes in metabolic, endocrine and immune system functionality[7]. Similarly, sleeping more than what is necessary for optimal health has also been associated with adverse health outcomes such as cognitive dysfunction, obesity and an overall higher rate of mortality. This link extends to mental health, with abnormal sleep implicated as either the cause or symptom in almost every mental disorder. Similarly, studies have shown sleep interventions providing relief to those suffering from psychosis, depression and PTSD symptoms [8-10]. Our knowledge of sleep, or lack thereof, as an important marker in neurodegenerative disease is also expanding[11]. Older individuals ‘at risk' of developing dementia were found to have longer and more frequent naps throughout the day, which were also associated with poorer cognitive functioning, as well as higher levels of depressive symptoms.[12]

1.1.2 – Sleep Duration

Sleep duration is the most frequently investigated sleep characteristic in relation to health [13]. The amount of sleep required for optimal health changes across our lifespan. According to recommendations made by the National Sleep Foundation (USA) in 2015, we require between 14-17 hours of sleep as newborns, 8-10h as teenagers, 7-9h as adults, which drops to 7-8 h for those 60+ years [14]. This parabolic model [15] (i.e. where too much or too little sleep predicts adverse health outcomes) of sleep is representative of a tightly controlled physiological system and viewed by Cappuccio 2010 [16] as a testament to its importance.

1.1.3 – Sleep Quality

While sleep duration is an important predictor of health, it is but one factor that affects the complex system of sleep (Figure 1) [17]. Sleep quality has similarly been observed to be a predictor of health, wellness and vitality [18-20]. Interestingly, no official consensus amongst clinicians or the scientific community exists explicitly defining it, with several studies beginning with a statement of what the authors interpret ‘sleep quality’ to be[18, 21, 22]. Overall, most definitions revolve around a similar premise, that sleep quality is a reflection of an individual's satisfaction with his or her sleep [18].

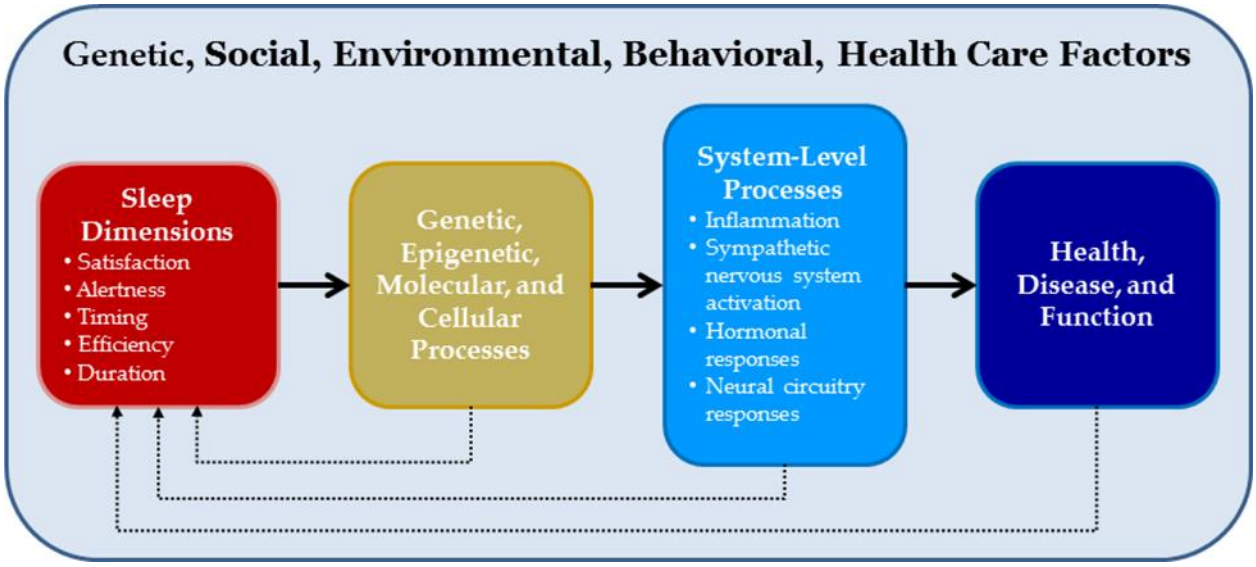


Figure 1: From Sleep Health: Can We Define It? Does It Matter? [17]. Buysees conceptual model of the relationship between sleep and health.

1.1.4 – Sleep as a Multidimensional Construct

Traditionally, studies that have examined the relationship between sleep and health have done so through the lens that specific sleep characteristics are themselves individual qualities[23]. Sleep duration is a key example, where adequate or optimal sleep has been reduced to a single, measurable value. Sleep as a whole is influenced by internal (i.e. circadian rhythm[24], diet[25], exercise[26]) and external environmental factors (i.e. artificial light[27], geography[28] and culture[29]) and thus multidimensional by its very nature. In order to properly frame the sleep-health link and make recommendations for optimal health, we must move past using individual qualities of sleep as independent measurements and instead move towards a multivariable, holistic approach to truly encapsulate sleep as a predictor of adverse health outcomes[17].

1.1.5 – The Economic Cost of Sleep Deprivation

As discussed in 1.1, there has been a major societal shift in our perception of sleep. There is a disconnect between our knowledge of the increased risk of adverse health outcomes as a result of poor sleep and how society conducts itself, be it in the frame of business and social norms. Unsurprisingly, the consequences of sleep deprivation have been framed from an economic standpoint, where on an annual basis, the United States alone loses roughly 1.23 million workdays due to insufficient sleep. Concerning direct cost to the economy per annum, this equates to a $411 billion loss to the US economy. This is not contained to the USA alone, with sleep deprivation costing the Japanese, German and British economy $138 billion, $60 billion and $50 billion respectively[30].

1.2 – Measuring Sleep

1.2.1 – Measuring Sleep Duration and Quality

Given the intrinsic links between health and sleep, the need for reliable, inexpensive and simple diagnostic tools to measure sleep has never been greater. Unsurprisingly given a lack of consensus on a definition, no ‘gold standard’ exists to measure sleep quality. For the most part, subjective measurements are used where individuals are asked to rate their sleep and determine how restful their sleep felt the next morning [31]. Laboratory studies have deconstructed sleep quality into objectively measured indices derived from polysomnography (PSG) [22].

1.2.2 – Polysomnography

A polysomnography (PSG) is a diagnostic tool used to assess sleep and identify sleep disorders. It is the current gold standard in sleep observations [32], consisting of an electroencephalogram (EEG), electromyogram (EMG), electrooculography (EOG) and oximeter [33]. Specifically, the PSG is able to measure objective indices of sleep that have been shown to affect sleep quality including in-depth analysis of sleep architecture and temporal analysis spent in each stage of sleep [34]. These objective indices include:

- Sleep Onset Latency (SOL): A measure of how long it takes for an individual to fall asleep once they start trying to fall asleep. An example of abnormal SOL can be observed in insomnia sufferers who have trouble ‘entering' sleep.
- Wake After Sleep Onset (WASO): The accumulated amount of time an individual spends awake during the night after the onset of sleep. Long WASO

typically indicates disrupted sleep, with individuals waking up multiple times a night.

- Sleep efficiency: The percentage of time an individual spends sleeping, according to their EEG, while in bed. A low sleep efficiency can result from a long sleep latency and long sleep offset to lights on time and is usually used to get an overall sense of how well a patient has slept.

While providing a wealth of information on sleep, the PSG has limitations that prevent its widespread use as a practical diagnostic tool. These include high costs associated with infrastructure (i.e. equipment and laboratory space), requiring expertise to set-up, run and interpret recorded data [35], and potential lack of validity in that it does not provide insights into sleep habits at home. There is also a high burden on participants, as the study is usually conducted onsite at a sleep laboratory and involves numerous wires and probes attached to an individual for the course of the study. This can lead to a 'first-night effect' phenomenon, where recorded sleep is not an accurate representation of an individual's regular sleep [36].

1.2.3 – Actigraphy

Actigraphy is another common method used to measure sleep objectively. Generally worn on the wrist, these devices measures activity (i.e. movement). Specific actigraphy devices can also measure light intensity, including exposure to red, green and blue light, which can assist in analysing circadian rhythm. Using specialised software and sleep/wake algorithms, objective measurements of sleep can be determined, similar to those obtained via PSG such as sleep efficiency, WASO and SOL. Actigraphy has also been validated against PSG and other methods in several

studies with varying degrees of agreement [37-41]. Overall, actigraphy provides a reasonable estimate of sleep time, and WASO, however, has been shown to provide a poor estimate of SOL [36, 37].

Actigraphy has several benefits over PSG. It is far less expensive and cumbersome and can provide an aggregate measurement of sleep over a period. While less expensive than PSG, each device can still cost thousands of dollars and requires an accompanying sleep diary to be completed to assist with interpretation as per standardised practice. [42]. There are also numerous actigraphy watches on the market that vary in scoring algorithms, sensitivity and specificity and complicate efforts to standardise the widespread use of these devices in sleep studies[43, 44].

1.2.4 – Subjective Measurements

Subjective measures of sleep are self-reported questionnaires that individuals complete. This typically involves a sleep diary, which prompts users to keep a record over a set period recording when they got into and out of bed, how long they slept for, if their sleep was interrupted and how well they felt they slept. These typically are completed on a nightly basis over the course of a week or two.

Another commonly used subjective measure of sleep is the Pittsburgh Sleep Quality Index (PSQI). The questionnaire consists of several questions asking for self-reports of sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleeping medication and daytime dysfunction [45]. Few studies have attempted to validate subjective sleep measurements against objective measurements. Lauderdale et al. [46] reports significant discrepancies between subjective measurements of sleep duration, with other studies also observing bias,

with subjective measures consistently overestimating sleep duration when compared to objective measures [47, 48]. Conversely, several studies have observed agreement with objective measures and have argued in favour of using subjective measures of sleep as an estimate of sleep duration given its minimal cost, ease of access and simplicity[21, 49].

1.3 – Sleep in LMIC

A disconnect has emerged between our increasing knowledge of the sleep-health link and the devaluation of sleep in western culture. Sleep is commonly associated with weakness, surmounting as a trivial task with no purpose[50]. Leading entrepreneurs and politicians such as PepsiCo CEO Indra Nooyi and Donald Trump extenuate this idea, reporting thriving careers on the back of four hours of sleep [51]. What is interesting is that amidst western culture devaluating sleep, a majority of sleep research is conducted in western, high-income countries (HIC) settings. Sleep is known to be treated differently between countries and cultures with the Spanish 'siesta' and Japanese 'inemuri' as two notable examples [52]. Both cultures accept daytime napping as being socially acceptable; however, this differs significantly from most western countries, where daytime napping is frowned upon and even considered a marker of poor sleep quality in the literature[18, 45]. Many factors that affect sleep also differ specifically between countries with lower income levels, typically related to diet[25], exercise [53] and technological use [54], quality of housing[55], security at home [56] and cohabitation with extended families[57]. This begs the question of how relevant the current body of knowledge of sleep derived from HICs and current sleep recommendations are to low-middle income countries (LMICs).

Chapter 2 – Agreement Between Subjective and Objective Measures of Sleep Duration in a Low-Middle Income Country Setting

Author names and affiliations:
Aaron Schokman [a], Yu Sun Bin [a, b], Guido Simonelli [c], Jonathon Pye [d, e], Richard Morris [d], Athula Sumathipala [f, g], Sisira H Siribaddana [f, h], Matthew Hotopf [g], Fruhling Rijsdijk [g], Kaushalya Jayaweera [f], Nick Glozier [a, d, e.]

[a] *Central Clinical School, Sydney Medical School, University of Sydney, Australia*
[b] *Sleep Group, Charles Perkins Centre, University of Sydney, Australia*
[c] *Behavioural Biology Branch, Walter Reed Army Institute of Research, Silver Spring, Maryland, USA*
[d] *Brain and Mind Centre, University of Sydney, Australia*
[e] *Neurosleep, NHMRC Centre of Research Excellence, Australia*
[f] *Institute for Research & Development, Colombo, Sri Lanka*
[g] *Institute of Psychiatry, Psychology and Neurology, King's College London, London, United Kingdom*
[h] *Department of Medicine, Professorial Unit, Teaching Hospital, Anuradhapura, Sri Lanka*

Corresponding author: Professor Nick Glozier, Brain and Mind Centre, Faculty of Health and Medicine, University of Sydney, NSW 2050, Australia. E-mail address: nick.glozier@sydney.edu.au

Conflicts of Interest: None of the authors have any relevant conflicts of interest to report.

Disclaimer: Material has been reviewed by the Walter Reed Army Institute of Research. There is no objection to its presentation and publication. The opinions or assertions contained herein are the private views of the author and are not to be construed as official, or as reflecting the views of the Department of the Army or the Department of Defence. The investigators have adhered to the policies for protection of human subjects as prescribed in AR 70–25.

Key Words: Actigraphy; Self-Report; Validation; Criterion Cut-points; Sleep Duration; Sleep Quality; Low-Middle Income Country (LMIC); Public Health; Analytical Epidemiology

Tables: 3

Figures: 2

2.2 – Abbreviations

aTST = Objectively recorded Total Sleep Time

CI = Confidence Interval

HIC = High-Income Country

LMIC = Low-Middle Income Country

LOA = Limit of Agreement

PSG = Polysomnography

SE = Sleep Efficiency

srTST = Self-reported Total Sleep Time

TST = Total Sleep Time

WASO = Wake After Sleep Onset

2.3 – Introduction

Interest in sleep epidemiology and the link between sleep and optimal health has grown over the last two decades [5, 16, 58, 59]. The majority of this research is conducted in Western, high-income country (HIC) settings [1-13] which warrants the question of how relevant findings and the derived consensus statements are [18, 60] to low-middle income countries (LMICs). Sleep duration and quality are known to be affected by cultural, social, environmental and geographical influences [28, 52, 61] which vary greatly between LMIC and their HIC counterparts. This is compounded by the rapid demographic and epidemiological transitions occurring in developing countries [51, 62, 63] resulting from lifestyle and cultural changes, uptake in technology, and shifts to urban living, which may have substantial population-level effects on sleep [64, 65].

Evidence on sleep from a public health perspective in low and middle-income settings is lacking in part due to logistical and financial limitations. Polysomnography (PSG) is the current “gold standard” diagnostic tool used to study sleep physiology [32, 34]. However, its use in population-based sleep epidemiology is untenable, as the equipment is expensive and specialised training is required to conduct and analyse the recordings [38]. A viable alternative to PSG is the use of actigraphy, a wearable device worn on the wrist that quantifies objective sleep measures. Actigraphy has previously been validated against PSG measurements of sleep duration, quality and efficiency [37], but is better suited for smaller studies as the devices are currently relatively expensive, require multiple days of continuous recording and an accompanying sleep diary to be simultaneously filled out for best practice [42]. A third option is the use of subjective self-reported sleep questionnaires that ask individuals

about their sleep habits duration and perceived sleep quality. Even though this method of sleep measurement is relatively inexpensive and thus ideally suited to studies involving large samples, there is a lack of consensus on the validity of its use to assess sleep characteristics compared to “objective” measures [21, 66].

Population-based studies of sleep duration have at least two aims; (i) descriptive epidemiology of sleep parameters in a population, e.g. establishing trends and (ii) analytical epidemiology to ascertain potential risks of negative health outcomes borne by those defined as belonging to a category, e.g. the extremes of sleep duration. The cut-points defining these categories are often determined by consensus based upon HIC setting data [18, 67] however may not be applicable in populations with different sleep distributions and demographics. Furthermore, the criterion validity and agreement of self-reported measures of behaviour and symptoms against “gold standards” vary across cultures [60]; thus the results of validation studies from HIC populations may not be applicable to other cultural settings. Validation studies often compare the accuracy of two comparable measurements using pre-set cut points yet fail to explore if other, better-suited cut-points exist, e.g. identifying agreed short and long sleepers with the smallest margin of error [21, 37, 66]. Criterion cut-point methods can also identify meaningful thresholds within the reference method as well as the corresponding cut points within a second method of measurement concurrently, rather than using one as “gold standard”.

This study aims to describe sleep duration in a well-ascertained sample and validate self-reported measurements of sleep duration against objective measurements in Sri Lanka, a South Asian LMIC. More specifically, the study aims to:

1. Describe the sleep duration of Sri Lankan adults using self-report and actigraphic methods.
2. Validate self-reported “subjective” sleep duration against “objective” sleep diary informed actigraphy in this setting.
3. Demonstrate whether the application of a criterion cut-point algorithm could identify agreed values discriminating short and long subjective and actigraphic measured sleep duration for use in analytic epidemiology in resource-poor settings.

2.4 – Methods

This study was designed as an actigraphic sub-study nested within the Colombo Twin and Singleton Study (CoTASS) 2. The study received ethical approval from the Psychiatry, Nursing & Midwifery Research Ethics Subcommittee, King's College London, UK (ref: PNM/10/11-124), the Faculty of Medical Sciences University of Sri Jayewardenepura Ethical Review Committee (USJP ERC) (ref: 596/11) and the Research Integrity & Ethics Administration, The University of Sydney, Australia (ref: 2012/2181).

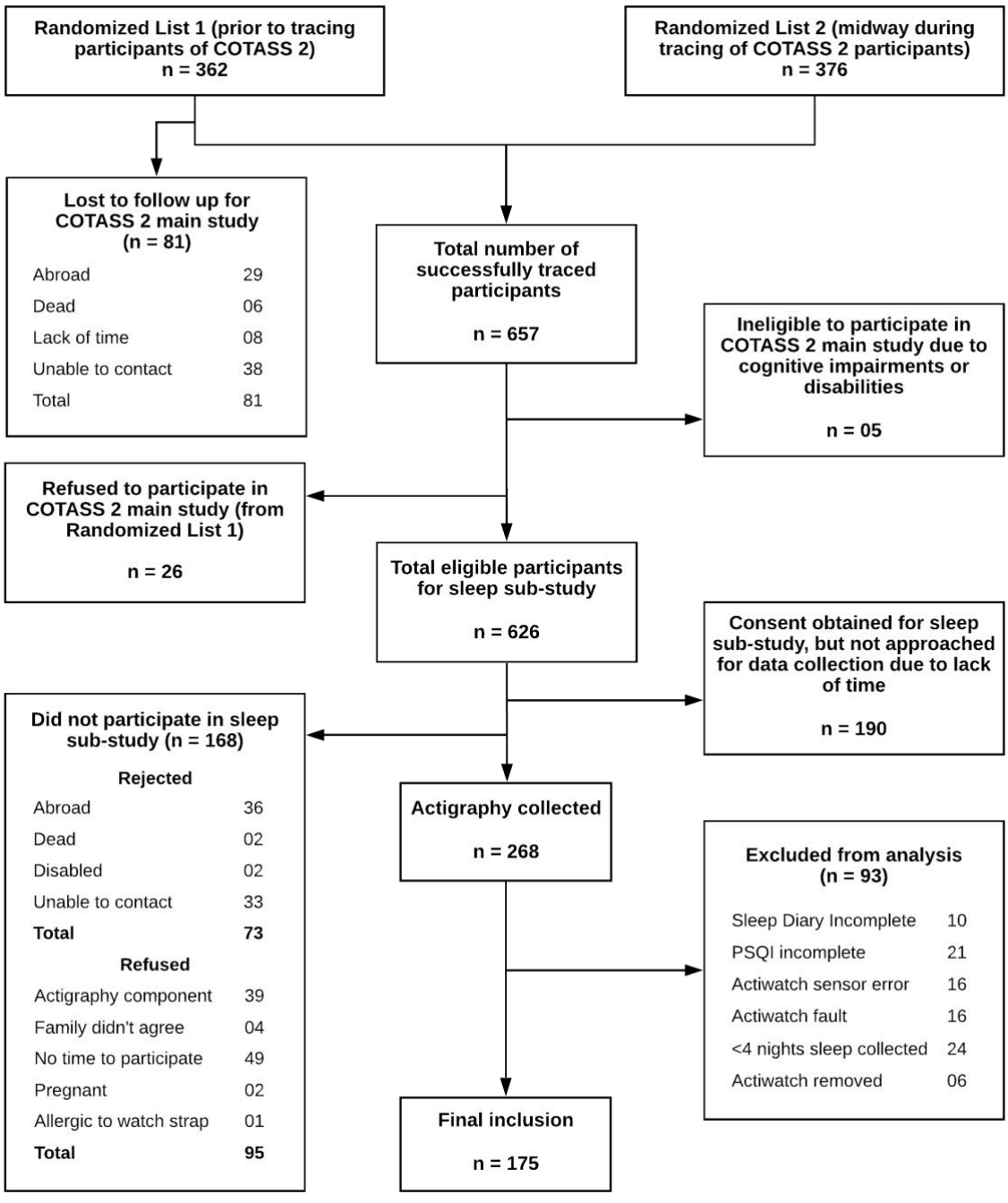
2.4.1 - Population-based Sample

CoTASS is a cohort study of twins randomly selected from the Colombo twin registry and a sample of singletons randomly selected from non-twin households, stratified by Grama Niladhari Divisions (GND, the smallest administrative unit, with approximately 4,000 people in each) from which the twins were ascertained, in Colombo, the capital of Sri Lanka. Baseline data collection took place between 2005 and 2007 as described in detail in Siribaddana et al. (2008) [68]. CoTASS 2 was a follow-up study that took place between 2012 and 2015 involving 3969 participants (2934 twins and 1035 singletons) and focused on genetic and environmental influence on mental health and cardiovascular disease. Trained researchers conducted face to face interviews and collected biometric samples and validated self-report questionnaires semantically translated into Sinhala [69], including the Pittsburgh Sleep Quality Index (PSQI), as described in further detail in Jayaweera et al. (2017) [70].

2.4.2 - Nested Actigraphy Sub-study

Invitations to participate in the sleep sub-study were given to 628 randomly selected participants traced from COTASS 2 cohort. Only twins and no singletons were sampled for the sleep sub-study. (Figure 2.1). Participants were required to wear an actigraphy wristwatch that recorded activity over a period of seven days while simultaneously completing a sleep diary. Of those invited, 95 individuals (15.2%) refused to participate, and a further 73 (11.7%) were uncontactable during recruitment (Figure 2.1). The self-reported sleep duration of the 168 not consenting to, or unable to be contacted for, the actigraphy sub-study was less (6.1h) than those in our analysed sample, (6.5h), $t(420) = 3.13$, $P = 0.002$, but did not differ on other health or demographic variables.

Figure 2.1: Flowchart describing twins participant recruitment from the twin cohort of the COTASS 2 sample of the sleep sub-study.



Four hundred and fifty-eight (73.2%) cohort participants accepted the invitation to participate. However, only 268 (58.5%) had their actigraphy recorded as a result of time restrictions on COTASS2 fieldwork assessments and delays on actiwatch importation. Of those with actigraphy, 16 (6.0%) were excluded from analysis due to the participant recorded as constantly moving even when asleep. Another 24 (9.0%) were excluded as fewer than 4 nights of data were recorded, and a further 6 (2.2%) due to extended periods of time where the device was removed. Two Actiwatch devices were found to be faulty, with several recordings made with the same devices recording extensive periods of maximum movement and zero light, and all 16 (6.0%) recordings associated with these two devices were excluded. A further 21 (7.8%) failed to complete the PSQI and 10 (3.7%) did not complete the accompanying sleep diary, leaving 175 (65.3%) included in the analysis. Of the 268 participants whom had actigraphy recorded, no demographic health and sleep duration differences were observed between those who were included and excluded from the final analysis (Appendix Table B1-2).

2.4.3 - Data Quality

Of the 3969 participants in COTASS 2, 3672 (94.9%) answered the Pittsburgh Sleep Quality Index (PSQI). Seven (0.2%) were removed as outliers having reported spending more than 15 hours in bed (>3 s.d) and a further 168 (4.6%) removed due to an interpretation error where total sleep time (TST) was reported longer than time spent in bed leaving 3497 (95.2%) participants with subjective measurements of sleep duration included in the final sample.

2.4.4 – Measures

"Objective" sleep duration

Actigraphic measurements were collected using wrist-worn actigraphy devices (Actiwatch Spectrum Pro, Phillips Respironics, USA, firmware: 01.01.2009) which have produced reliable sleep statistics [71]. Measurements were collected in 30-second epochs over a 7-day period, with a minimum of four days of valid recording to be included in the study. Participants were required to complete a daily sleep diary and indicated bedtime and rise time using the event button on the Actiwatch device. Sleep-wake detection algorithms were used by Actiware 6.0 software (Phillips Respironics), set to a medium sensitivity threshold and 10 minutes of immobility for sleep onset. [42, 72]. Manual scoring of actigraphy was conducted by one of the authors (AS) based on visual inspection, sleep diary entries and Actiwatch timestamps according to standardised guidelines [42]. Average scores of total sleep time (TST) were calculated, defined as the total amount of time scored as sleep during the main rest period, which was nights for most participants. To ensure internal scoring, validation and minimization of bias, a second researcher (JP) scored a random selection of 40 actigraphy recordings (20% of total sample) and agreement of these assessed as per standardised guidelines[42]. The resulting inter-scorer agreement fell within acceptable limits with no systematic bias observed (95% CI: -0.17, 0.01 and limit of agreement: -0.58, 0.42).

"Subjective" sleep duration

The PSQI is a standardised self-reported questionnaire that retrospectively assesses sleep of the prior 30 days [45]. Self-reported total sleep time (srTST) was calculated

using question 4 of the PSQI ‘during the past month, how many hours of actual sleep did you get at night?’.

Health and sociodemographic measures

Both health and sociodemographic information were collected during CoTASS 2. Height and weight were measured, and body mass index (BMI) was calculated for each. Depressive symptoms were assessed using the Beck Depression Inventory [73], and the presence of depression was defined as a score of over 21. Alcohol abuse was assessed using the Alcohol Use Disorders Identification Test [74]. A score greater than eight indicated alcohol abuse. Participants were asked about whether they currently smoked or not (‘smoker’ or ‘non-smoker’). Participants were also interviewed about their medical history, including hospital visits, surgery and targeted questions regarding specific illnesses and chronic diseases. A positive answer during the interview indicated the presence of chronic disease. Finally, employment status was similarly queried in the interview, and the information later dichotomized into two broad categories: “employed” (including part-time work) and “unemployed” (including students/retired).

2.4.5 - Statistical Analysis

Differences in self-reported sleep duration (PSQI Q4) across the CoTASS 2 sample population was assessed by stratifying by sex, age and employment status. Differences between binary and categorical variables was assessed using an independent sample t-test and a one-way ANOVA with Games-Howell posthoc analysis respectively.

We similarly compared actigraphy subsample to those of the entire CoTASS 2 population to assess representativeness. Binary variables were compared using chi-square analysis with variability assessed by odds ratio scores while continuous variables including self-reported sleep measurements were assessed using independent-sample t-tests.

Subjective and objective sleep measurements

Validation of subjective against objective measurements of sleep duration was conducted using an adapted version of the Bland Altman plot method [75]. Actigraphy was plotted on the x-axis (as opposed to the average of the two methods) as actigraphy is used as the gold standard reference method [76]. Mean difference, confidence intervals, and limit of agreement were used to examine agreement between methods. These were recalculated following stratification by age dichotomized at the mean age of sample and sleep efficiency dichotomized at 85% as recommended by the National Sleep Foundation [18]. As only two participants had a wake after sleep onset (WASO) of < 20 min, the next cut-points of <30 mins and ≥ 31 mins were used to stratify measurements [18].

2.4.6 - Criterion Validation of Short and Long Sleep Duration in this Sample

We employed a criterion method defining cut points for short and long sleep duration in this sample using data from both measurement approaches. We imposed a criterion on our algorithm that at least 10% of the total sample be assigned to the short and long sleep categories and that a minimum of 40% of the sample be assigned to "normal" duration, with no result produced if these conditions are not met (Eq. A2)

Each subject was assigned two scores 'x_i', created using converted z score of self-reported TST and objective measurement of TST. Three categories were created (C_1, C_2, C_3) that corresponded to short, normal and long sleep duration and defined by the parameter Θ that consisted of a lower and upper threshold.

$$\Theta = [\theta_{lower}, \theta_{upper}]$$

θ_{lower} and θ_{upper} refer to the defining cut-points between each category and once applied to 'x_i', allowed for the creation of two class assignments, $class_j$ and $class_k$ that correspond to self-report and objective measurements respectively. A confusion matrix '*m*' was created to distinguish between classes, with the sum of the diagonal directional elements a representation of error or "cost" (Eq. A3). This cost function (Eq. A4) was minimized using the default Nelder-Mead Simplex algorithm in scikit-learn (Python 3.6.1). The function was minimized over 400 starting seeds spread evenly across parameter space between z = 1.28 and -1.28. The top 5 results of this cost function (i.e. when agreement errors were the least) were applied separately to both self-report and actigraphy datasets, creating corresponding duration cut-points that identified objective and subjective short, average and long sleepers with agreement between these categories assessed using a frequency analysis table.

2.5 - Results

2.5.1 – Sleep duration of Sri Lankan adults

Differences in self-reported sleep duration between common binary and categorical variables is described in table 2.1 and 2.2. On average, females reported to have slept 21 minutes less than their male counterparts (95% CI: 0.25 – 0.45). The largest mean difference was observed between those classified as being depressed, with almost an hour difference noted (95% CI: 0.68 – 1.29). With regards to employment, students reported to have slept the most at 6.8h, followed by those employed at 6.1h and those unemployed at 6.0h ($F(2, 3490) = 6.87$, $p < 0.01$). Of interest, when further stratified by gender, no significant difference was observed between females that were employed and those that indicated that their role was a housewife.

Table 2.1: Differences between commonly observed characteristics associated with sleep duration within the CoTASS 2 cohort

Binary variables		n	Sleep Duration (h) [a] $\bar{x}$ (SD)	Mean difference (h) (95% CI)
Sex	**Male**	1452	6.26 (1.49)	0.35 (0.25 – 0.45)
	Female	2045	5.92 (1.50)	
Depressed	**Yes**	126	5.11 (1.72)	0.98 (0.68 – 1.29)
	No	3353	6.09 (1.49)	
Alcohol Abuse	**Yes**	422	6.05 (1.60)	0.01 (-0.15 – 0.17)*
	No	3071	6.06 (1.50)	
Current Smoker	**Yes**	439	6.28 (1.57)	0.25 (-0.41 – -0.09)
	No	3058	6.03 (1.50)	

* indicates no significant difference, BMI: body mass index, [a] self-reported answer to PSQI Q4.

Table 2.2: Mean sleep duration, standard deviations and correlations of categorical variables collected during the CoTASS 2 study.

Variable	n	Sleep Duration [a] (h) $\bar{x}$ (SD)	Mean Differences | *Xi* −*Xj* | (Effect sizes are indicated in parentheses)			
			1	**2**	**3**	**4**
Age						
1. ≤ 40 (y)	1632	6.38 (1.50)	--			
2. 40 – 70 (y)	1692	5.79 (1.45)	0.59** (0.40)	--		
3. ≥ 70 (y)	173	5.61 (1.64)	0.77** (0.49)	0.18	--	
Employment status						
1. Yes	1946	6.11 (1.47)	--			
2. No [b]	1512	5.98 (1.56)	0.13* (0.09)	--		
3. Student	35	6.77 (1.42)	0.79** (0.55)	0.67* (0.45)	--	
Role of females						
1. Employed	746	5.83 (1.44)	--			
2. At home [c]	1252	5.96 (1.51)	0.13	--		
3. Unemployed	47	6.10 (2.28)	0.27	0.13	--	
BMI						
1. Underweight	399	6.15 (1.45)	--			
2. Normal	1054	6.15 (1.58)	0.00	--		
3. Overweight	1141	6.04 (1.48)	0.11	0.11	--	--
4. Pre-Obese / Obese	634	5.89 (1.51)	0.26* (0.18)	0.26** (0.17)	0.15	--

* $p < .05$, ** $p < 0.01$, BMI: body mass index, [a] self-reported answer to PSQI Q4, [b] working at home or as a housewife was classified as unemployed, [c] worked in the capacity of a housewife.

2.5.2 – Actigraphy and CoTASS2 2 cohort comparison

The actigraphy subsample and overall CoTASS 2 sample characteristics are described in table 2.3. The actigraphy sample was younger, less likely to abuse alcohol and possibly less depressed than the CoTASS 2 sample. No significant difference between samples was observed in sex, current smoking status, or BMI. Rates of chronic disease were high in both samples, with more than 85% of both samples reporting to suffer from chronic physical illness (ex. dental disease).

Table 2.3: Comparison of demographics of sub-study and CoTASS 2 sample.

Binary variables (n, %)		**Actigraphy (n=175)**	**CoTASS (n=3322)** [a]	**Difference Odd Ratio (95% CI)**
Sex	Male	73 (41.7%)	1379 (41.5%)	0.99 (0.73 – 1.35)
	Female	102 (58.3%)	1943 (58.5%)	
Comorbid disease	Yes	155 (88.6%)	2865 (86.3%)	1.23 (0.99 – 1.98)
Employed [b]		101 (57.7%)	1845 (55.6%)	1.09 (0.80 – 1.48)
Alcohol abuse		11 (6.3%)	411 (12.4%)	0.48 (0.26 – 0.89)*
Depressed [c]		2 (1.2%)	172 (5.3%)	0.30 (0.07 – 1.22)
Current smoker		17 (9.7%)	422 (12.7%)	0.74 (0.44 – 1.23)
Continuous. variables		**($\bar{x}$, SD)**	**($\bar{x}$, SD)**	**Difference t (df), p**
Age (y)		38.7 (11.6)	43.3 (14.4)	4.2 (3495), p < 0.01
BMI (kg/m^2) [d]		23.5 (4.5)	23.8 (4.6)	0.9 (3226), p = 0.38*
Sleep measurements		**($\bar{x}$, SD)**		
srTST (h)		6.4 (1.5)	6.0 (1.5)	3.3 (3495), p < 0.01
aTST (h)		6.0 (0.9)	-	-
WASO (min)		48.6 (22.8)	-	-
Sleep Efficiency (%)		84.6 (5.9)	-	-

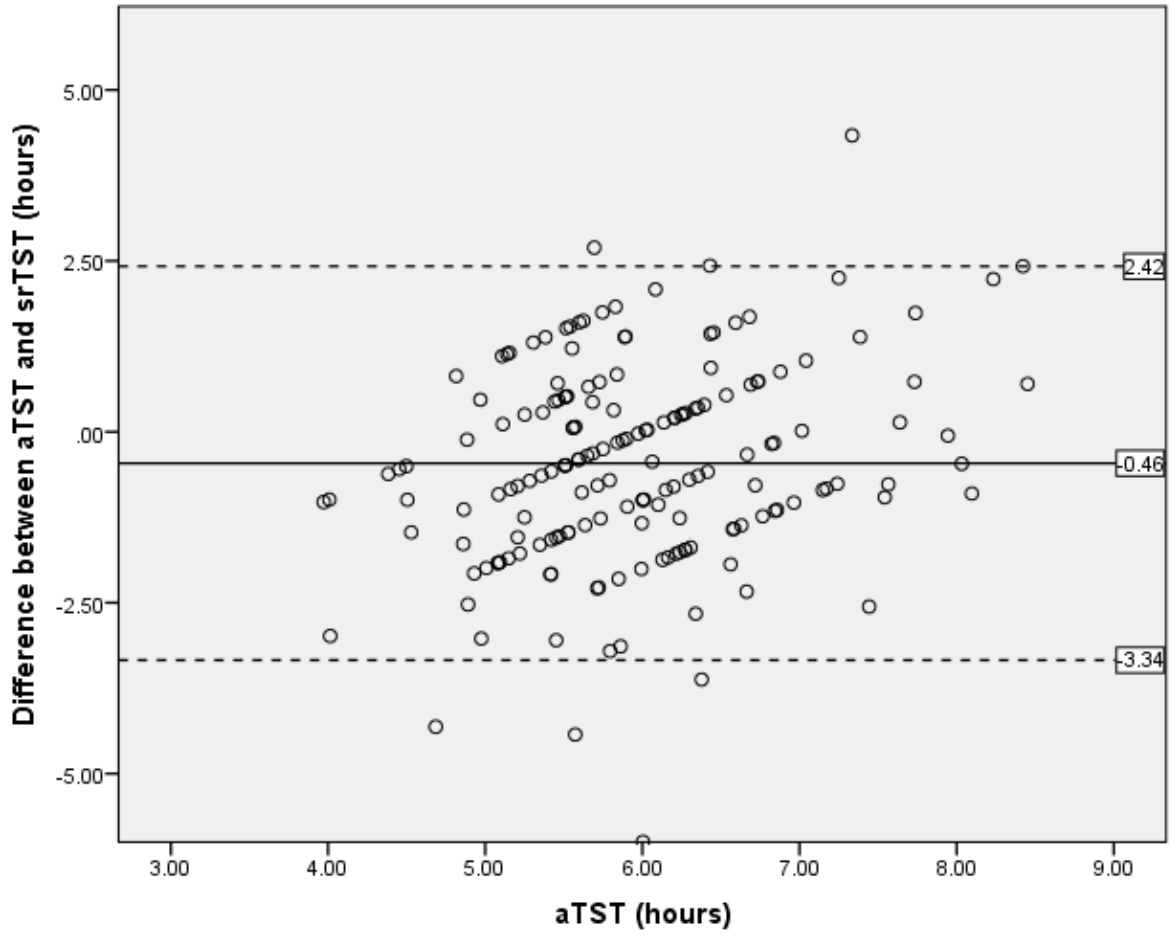
* indicates no significant difference, aTST: actigraphy total sleep time, CI: Confidence Interval, PSQI: Pittsburgh Sleep Quality Index, srTST: self-reported total sleep time, WASO: Wake After Sleep Onset, [a] CoTASS 2 sample does not include actigraphy sample, [b] part-time work considered employed and students counted as unemployed, CoTASS (n= 3318); [c] actigraphy (n=174), CoTASS (n = 3305), [d] actigraphy (n=174), CoTASS 2 (n=3054)

2.5.3 – Sleep Duration

Mean self-report sleep duration was low: averaging 6.4 (SD 1.5) and 6.0 (SD 1.5) hours in the actigraphy and total sample respectively (the actigraphy sample reporting slightly longer sleep duration). Mean objective sleep duration was also low. The sub-study sample spent 6 hours asleep at night and on average spent an accumulated 50 minutes awake each night after sleep onset (WASO). This is reflected in the mean sleep efficiency (SE) of 85%.

The Bland Altman plot (figure 2.2) showed a significant systematic difference between objective and subjective measurements of TST, with participants reporting to have slept on average 27.6 minutes (-0.46h, SD = 1.47) more than recorded using actigraphy. The limit of agreement was extremely large, extending over 5.76 hours and ranging from over-reporting TST by 3.34 hours to under-reporting by 2.42 hours.

Figure 2.2: Bland-Altman plot comparing objective and subjective measurements of total sleep time (TST).



aTST: Actigraphy measured TST, srTST: Self-reported TST. The solid line represents the mean difference of TST at 0.46 hours (SD = 1.47). A 95% CI: (-0.68h, -0.24h) is evidence that systematic bias exists between our datasets. The dotted lines represent limit of agreement (LOA): (-3.34, 2.42).

Assessment of the influence of demographics and sleep quality on bias and agreement (table 2.4) showed that consistent over-reporting was present regardless of sex or age. Those older than the sample's average age (39 yrs.) had less variability, with the maximum under-reporting falling under 2 hours. There was an interaction of sleep quality with bias. Over-reporting only occurred in those who had poorer actigraphic determined WASO (≥31 min) and poorer actigraphic sleep efficiency (<85%).

Table 2.4: Analysis of bias (subjective sleep duration over-reporting vs actigraphy) and agreement after stratification by variables

Variable		n	srTST – aTST hours (SD)	95% CI	95% LOA (h)
Sex	**Male**	73	-0.69 (1.45)	-1.03, -0.35	2.45, -3.18
	Female	102	-0.30 (1.47)	-0.59, -0.01	2.59, -3.18
Age	**< 38.7 (y)**	99	-0.49 (1.66)	-0.82, -0.16	2.77, -3.74
	≥ 38.7 (y)	76	-0.43 (1.18)	-0.70, -0.16	1.89, -2.75
Comorbid disease	**Yes**	155	-0.36 (1.44)	-0.59, -0.14	2.45, -3.18
	No	20	-1.22 (1.55)	-1.94, -0.50	1.81, -4.25
Employed	**Yes**	101	-0.53 (1.29)	-0.78, -0.27	2.01, -3.06
	No	74	-0.37 (1.69)	-0.77, 0.02 [a]	2.93, -3.68
WASO	**< 30 (min)**	37	-0.06 (1.27)	-0.36, 0.48 [a]	2.55, -2.43
	≥ 31 (min)	138	-0.60 (1.49)	-0.85, -0.35	2.32, -3.52
Sleep efficiency	**≥ 85%**	94	-0.08 (1.25)	-0.33, 0.18 [a]	2.37, -2.53
	< 85%	81	-0.90 (1.58)	-1.25, -0.55	2.22, -4.00

aTST: actigraphy total sleep time, srTST: self-reported total sleep time, CI: Confidence Interval, LOA: Limit of Agreement, WASO: Wake After Sleep Onset, SE: Sleep Efficiency. [a] no bias observed

2.5.4 - Criterion Validity

Only one valid criterion cut-point resulted from our cost regularization function (Appendix A) that met the *a priori* conditions of a minimum 10% of objective short and long sleep and a minimum 40% in the “normal” category. Objective cut points of 5.4 hours and 6.4 hours respectively defined short and long sleep duration. This

corresponded to the subjective cut-point of 5.5 hours for short sleep and 7.2 hours for long sleep. Our frequency distribution table (Table 2.5) shows that srTST did not accurately identify short sleepers with only 37% accurately identified as short and 50% misidentified as medium and 13% being long sleeper on actigraphy. There were similar errors for long sleep.

Table 2.5: Validation of correct classification of subjects with short, normal and long sleep durations using subjective measurements against objective methods of measurement.

		Classification by srTST n (% within aTST category)			
		Short	**Normal**	**Long**	**Total**
Classification by aTST [a]	**Short**	17 (37.0%)	23 (50.0%)	6 (13.0%)	46 (100%)
	Normal	20 (26.7%)	39 (52.0%)	16 (21.3%)	75 (100%)
	Long	7 (13.0%)	23 (42.6%)	24 (44.4%)	54 (100%)
	Total	44 (25.1%)	85 (48.6%)	46 (26.3%)	175 (100%)

[a] Classification was determined using objective cut-points of 5.42 and 6.42 hours and subjective cut points of 5.48 and 7.22 hours to determine short, normal and long sleep duration

2.6 – Discussion

To our knowledge, this is the first multi-method description of an adult population sleep duration and measurement validation study conducted in a LMIC in South Asia. A notable finding is a short mean sleep duration in Sri Lankan adults of between 6.0 and 6.4 hours using both objective and subjective methods. This is considerably lower than the 7-9 hours of sleep recommended by HIC consensus groups [67] and the mean self-reported sleep duration of 7.5 hours observed from a combined sample of 71883 individuals from seven LMICs in our recent meta-analysis [61]. While we cannot account for this difference, neither of the estimates above reported on a South Asian adult population. Although not evident, this short sleep duration does match the experience of the authors who live and work in Sri Lanka. Disturbed sleep was prevalent, with almost 80% of our sample having a WASO greater than half an hour and 46% having poor actigraphic sleep efficiency. This is of particular concern given previous studies have shown induced sleep disturbances results in similar physiological consequences to those seen in sleep restriction [77]. If true, short sleep duration paired with poor sleep quality may be an unexplored factor that may partially explain the existing high rates of cardiometabolic comorbidity seen in Sri Lanka [78].

One objective of our study was to validate subjective sleep duration against objective sleep measurements for use in large-scale epidemiology. The Bland Altman plot was selected over other measures of correlation as it is better suited to quantify agreement between two quantitative methods of measurement [76]. We observed systematic bias, with subjective measurements consistently over-reporting sleep on average by almost half an hour. This finding is consistent with previous validation studies between self-report and actigraphy measured sleep duration [21, 46]. Stratification of our

sample showed systematic bias only occurred in those that had the poorest objective measurements of sleep quality. In other words, sleep quality affected the reliability of self-reported sleep duration. When used in descriptive population epidemiology this bias may be less important as it can be accounted for if known through such sub-studies (or even discounted in time trend analysis) unless there is a strong interaction with key demographic variables, which was not seen here.

Our findings are more concerning when applied to analytic epidemiology, e.g. using self-reported sleep duration as an exposure, which requires an accurate ascertainment of sleep duration and low levels of misclassification when using categorical exposures. There was a wide range of individual differences in agreement between the two methods that ranged from being underreported by 2.5 hours and over reported by 3.5 hours. This range of nearly six hours surpassed the 2-hour maximum acceptable difference that we defined *a priori* based on the 2-hour spread between the recommended 7 – 9 hours of sleep a night [67]. This maximum acceptable difference was chosen as anything greater would result in misclassification of those that have had short sleep (<7 hours) potentially being misclassified as having long sleep (>9 hours). This was confirmed through our criterion cut-point validation. Unfortunately, no cut points were identified that could reasonably accurately classify agreed short and long duration sleepers. If a random misclassification did occur, this would bias associations to the null potentially obscuring real associations. However, as we observed over reporting of sleep duration in those with poor sleep quality (good sleepers being quite accurate), this would imply that population-based analyses of the effect of self-reported short sleep duration on, e.g. morbidity, will underestimate the effect size in those who also have poor sleep quality.

There are several reasons why this systematic bias and low agreement between self-reported and objective measurements were observed. We could not account for whether objective sleep included ‘workdays’ or ‘days off’ which would have affected agreement between the two measurements. This has been observed in previous studies, where shorter durations of sleep occurred on weekdays, with sleep duration extended on the weekend to compensate, known as social jetlag [79, 80]. Another reason is the use of the PSQI as our subjective method of measurement and the fact that it requires the participants to provide estimates of sleep over the last 30 days. Lauderdale [46] shows that requesting estimates of time using “… last 30 days” produces temporally restricted estimates of sleep, while Biddle [66] suggests that agreement between subjective and objective measurements could be improved if participants were questioned using specific time periods. A ‘…over the last two weeks’ question has been proposed to improve agreement [66, 81]. The high prevalence of poor sleep quality within our sample also have contributed to a low overall agreement between measurements. Previous studies have shown that those with poor self-reported sleep quality tended to under-report sleep duration [82] in contrast to this study. It suggests movement recorded using actigraphy during the night and classified as “awake” is not perceived as such by the person when using self-report. This may be evidence of a larger validity issue around actigraphy, with low agreement observed between PSG and actigraphy in poor sleepers [83]. This may also be a limitation of the translation of the PSQI, as anecdotal evidence given by research assistants suggested that the wording of questions relating to time of sleep onset being ambiguously interpreted as the time in bed in Sinhala.

The study was limited by several factors. First, there was inevitable attrition. The present study achieved good follow up rates with few systematic differences between (i) COTASS 2 responders and non-responders, and (ii) those whom had actigraphy recorded that were included and excluded from analysis. (Appendix table B1-2). The sampling difference in short sleep duration would make our observations an overestimation, thus highlighting short sleep duration in this setting. Second, a minimum of 4 nights of valid actigraphy was required by each subject to be included in the study; however standard practices recommend a minimum of 7 nights recording [42]. This was decided in part due to logistics; the number of devices available was limited due to financial constraints, and nearly one-third of participants not recording seven nights of actigraphy, having removed the watches for extensive periods. Last, the criterion cut-point method is sample specific. As such, it may be able to identify agreed boundaries for short and long sleep in other samples. Further development of a standardised criterion test would allow for the rapid comparison of two methods of measurement to determine the accuracy in identifying the extremes of distribution that are normally of interest.

2.7 – Conclusion

Sri Lankan adults have a high prevalence of short sleep duration and poor sleep quality in comparison to their HIC counterparts, potentially questioning the applicability of consensus statements [18, 60] derived from HIC samples. Basic things like security, lack of stable housing, poverty and hunger are all factors that affect sleep quality and duration [51, 55, 56] and the effect of geography and climate on sleep is still not fully understood [27, 28]. Anecdotally, these themes were reflected in sleep diaries, with individuals commonly reporting disturbed sleep due to shared sleeping quarters, heat,

and safety concerns. If true, these findings suggest sleep disturbance as a new avenue for assessing the causes of the very high obesity and diabetes rates in Sri Lanka. Nevertheless, low levels of agreement between methods and difficulty in ascertaining reliable classification advocates for caution when interpreting epidemiological findings. This study demonstrates the need for culturally relevant sleep recommendations, consistency in metrics, and exploration into the feasibility, reliability and validity of more economical devices used to measure sleep.

Acknowledgements

Dr. Glozier reports grants by the National Health and Medical Research Council (NHMRC) of Australia; Program Grant 566529

2.8 – Appendix

APPENDIX A: Equations

Equation A1: Threshold function

$$\Theta = [\theta_{lower}, \theta_{upper}]$$

Equation A2: Class and condition equations

$$class_j\ (x_i\ ; \Theta) = \begin{cases} C_1,\ if\ x_i^1 <\ \theta_{lower} \\ C_2,\ if\ x_i^1 >\ \theta_{lower} \text{ and } x_i^1 <\ \theta_{upper} \\ C_3,\ if\ x_i^1 >\ \theta_{upper} \end{cases}$$

$$class_k\ (x_i\ ; \Theta) = \begin{cases} C_1,\ if\ x_i^2 <\ \theta_{lower} \\ C_2,\ if\ x_i^2 >\ \theta_{lower} \text{ and } x_i^2 <\ \theta_{upper} \\ C_3,\ if\ x_i^2 >\ \theta_{upper} \end{cases}$$

Equation A3: Confusion matrix

$$m = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}$$

Where c_{ij} is the number of people belonging to classi in classj . Thus, the diagonal elements of m indicate the correspondence between the two dependent variables given thresholds f, while the off-diagonal elements indicate individuals who are classified into different sleep classes according to the different dependent measures. The sum of the off-diagonal elements of ‘m’ were included in the cost function, along with a regularization term.

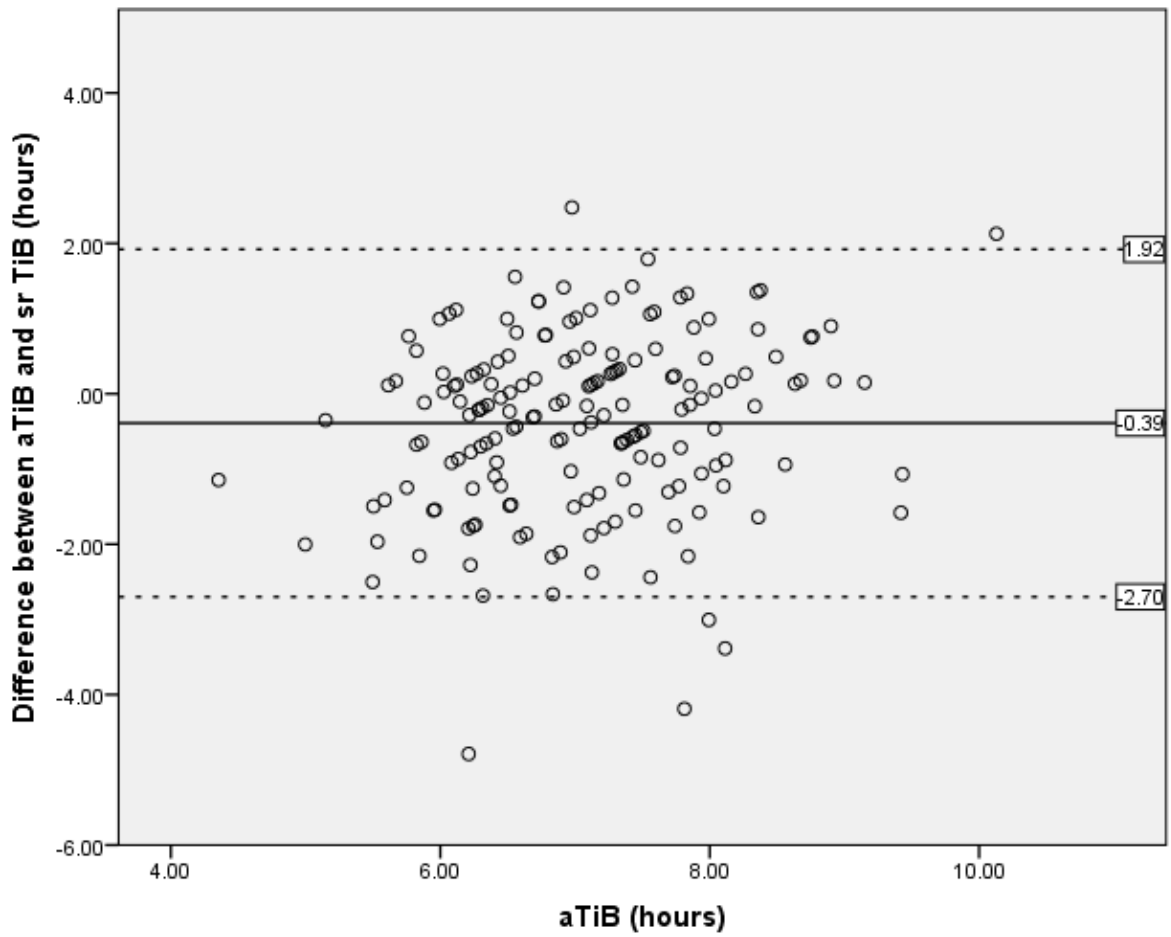
Equation A4: Cost-regularisation function

$$cost = \sum_{i \neq j,\ i,j=1\ldots3} m_{ij} + L\,|\Theta|$$

Where $|\Theta|$ refers to L_2 norm and L is a free parameter controlling regularization. We selected L and accepted the best parameter result, which categorized a minimum of 40 percent of the sample in the `middle' category, and a minimum of 10 percent in the short and long categories.

APPENDIX B: Omitted Analysis

Figure B1: Bland-Altman plot comparing objective and subjective measurements of Time in Bed (TiB).



aTiB: Actigraphy measured TiB, srTiB: Self-reported TiB. The solid line represents the mean difference of TST at 0.39 hours (SD = 1.18). A 95% CI: (-0.57h, -0.21h) is evidence that systematic bias exists between our datasets. The dotted lines represent limit of agreement (LOA): (-2.70, 1.92).

Table B1: Comparison of demographics of those that did not consent to participate in the sleep study vs those that had their actigraphy recorded.

Binary variables (n, %)		**Actigraphy Collected (n=268)**	**Did not consent (n=168)**	**Difference Odd Ratio (95% CI)**
Sex	Male	110 (41.0%)	62 (36.9%)	1.19 (0.80 – 1.77)
	Female	158 (59.0%)	106 (63.1%)	
Comorbid disease [a]	Yes	228 (85.1%)	135 (80.4%)	0.80 (0.47 – 1.35)
Employed [b]		161 (60.1%)	89 (53.0%)	0.78 (0.53 – 1.15)
Alcohol abuse [c]		20 (7.46%)	16 (9.52%)	1.32 (0.66 – 2.63)
Current smoker [e]		28 (10.4%)	19 (11.3%)	1.11 (0.60 – 2.06)
Continuous. variables		**($\bar{x}$, SD)**	**($\bar{x}$, SD)**	**Difference t (df), p**
Age (y)		41.2 (12.4)	42.4 (13.09)	-0.92 (434), p = 0.360
BMI (kg/m^2) [f]		23.2 (4.6)	23.9 (4.2)	-1.48 (398), p = 0.141
srTST (h) [g]		6.5 (1.5)	6.1 (1.4)	3.13 (420), p = 0.002 *

BMI: body mass index, srTST: Self-reported total sleep time. * indicates significant difference. [a]: Actigraphy n=263, consent n=159, [b]: Actigraphy n=268, consent n=165, [c]: Actigraphy n=266, consent n=165, d: Actigraphy n=266, consent n=160, [e]: Actigraphy n=268, consent n=166, [f]: Actigraphy n=265, consent n=135, [g]: Actigraphy n=263, consent n=159.

Table B2: Analysis of demographics and self-reported sleep between those that were included and excluded from the final analysis and had their actigraphy recorded.

		Final analysis of actigraphy		
Binary variables (n,%)		**Included (n=175)**	**Excluded (n=93)**	**Difference Odd Ratio (95% CI)**
Sex	Male	73 (41.7%)	37 (39.8%)	0.92 (0.55 – 1.54)
	Female	102 (58.3%)	56 (60.2%)	
Comorbid disease	Yes	158 (88.6%)	75 (80.6%)	1.61 (0.74 – 3.49)
Employed		101 (57.7%)	60 (64.5%)	0.75 (0.45 – 1.26)
Alcohol abuse		11 (6.3%)	9 (9.7%)	0.62 (0.25 – 2.56)
Current smoker		17 (9.7%)	11 (11.8%)	1.61 (0.74 – 3.49)
Continuous. variables		**($\bar{x}$, SD)**	**($\bar{x}$, SD)**	**Difference t (df), p**
Age (y)		38.7 (11.6)	40.4 (13.8)	-0.82 (266), p = 0.413
BMI (kg/m2)		23.5 (4.5)	22.6 (4.8)	-1.55 (263), p = 0.121
srTST (h)		6.4 (1.5)	6.7 (1.5)	1.46 (261), p = 0.46

BMI: body mass index, srTST: Self-reported total sleep time

Chapter 3 - Sleep Quality Composite Analysis

3.1 – Introduction

3.1.1 – Sleep Quality

Sleep quality has been shown to be a predictor of health and vitality [20]. As discussed in Chapter 1, no consensus on a single definition of sleep quality has been reached by researchers and clinicians, nor a definition widely accepted. As such, for the purposes of this study, we have used the National Sleep Foundations definition of sleep quality, defined to be a measure of 'worth', or how good/bad sleep was (i.e. how restful was your sleep).

3.1.2 – Measuring Sleep Quality

Given the importance of sleep quality as a predictor of health, a validated method of sleep quality is needed to identify those with poor sleep quality. Unsurprisingly given the lack of consensus definition, no gold standard exists to measure sleep quality. Typically, studies have determined how well an individual slept by using subjective methods of measurement such as the PSQI or sleep diary. Self-report questionnaires allow individuals to rate how well they slept over a specified period of time e.g. how well did you sleep over the last month). While easy to interpret and administer, subjective measurements of sleep quality suffer from the variability that exists between each night of sleep and may be influenced by culture (as discussed in Chapter 1). An alternative method of measuring sleep quality has also been suggested by several authors, where objective indices of sleep are measured and combined to reflect how well an individual slept [21, 22]. These include wake after sleep onset (WASO), sleep onset latency (SOL), sleep efficiency and sleep fragmentation which have been shown

to correspond to subjective measures of sleep quality in certain subgroups of sleepers [84].

3.1.3 – Aims and Objectives

Traditionally, research into the sleep-health link has been unidimensional, where single individual indices of sleep have been used to describe 'optimal sleep'. This approach does not account for potential interactions between characteristics or the inherent multidimensional nature of sleep[23]. By combining individual objective indices into single composite scores, we aim to investigate how well these measures correlate to subjective sleep quality ratings in a LMIC setting. Specifically, we aim to:

- Construct composite variables derived from objective indices that measure sleep quality
- Investigate the agreement between single indices and composite objective variables of sleep quality with subjective "self-reported" sleep quality in a LMIC setting.

3.2 – Methods

The methodology for this study describing population, recruitment, data quality and collection of actigraphy and subjective sleep quality have previously been described in Chapter 2, pages 15-17.

3.2.1 – Measures of 'Objective' and 'Subjective' Sleep Quality

Objective indices of sleep quality (i.e. wake after sleep onset (WASO), sleep onset latency (SOL) and sleep efficiency (SE)) were obtained from 175 participants described in Chapter 2. Using the National Sleep Foundations [18] recommendations of objective indices for good sleep quality, WASO, SOL and SE were converted into categorical variables indicating 'good, 'fair and 'bad' sleep quality (table 3.1). Self-reported sleep quality was measured using question 6 of the PSQI: 'During the past month, how would you rate your sleep quality.' Answers were converted from a 4-tiered categorical variable into a 3-tiered categorical variable by converting 'very good' to 'good sleep quality, 'fairly good' to 'fair sleep quality and combining 'fairly bad' and 'very bad' to 'bad sleep quality'.

3.2.2 – Composite Scores of Sleep Quality

Four potential measurements of sleep quality were created by combining measures of objective indices that make up sleep quality. This method was adapted from a similar study that created composite variables to measure sleep quality [21]. Composite variables were created by combining SOL and SE, SOL and WASO, SE and WASO and finally SOL, SE and WASO (seen in Table 3.2). An example of this would be

combining a WASO of ‘fair sleep quality’ and a SOL of ‘bad sleep quality’ would produce a composite score of ‘bad sleep quality’.

3.2.3 – Analyses

Agreement between objective indices and composite variables against subjective measures of sleep quality was assessed using two methods, (i) overall agreement as a percent, (ii) Assessment of inter-rater reliability between the two measurements of sleep quality using Cohen’s weighted kappa [85]. A prevalence index was also created to describe balance of a ‘positive’ and ‘negative’ category within our cohort and to aid in interpreting kappa scores. ‘Good sleep quality’ and ‘fair sleep quality’ were combined into a single ‘good’ category and agreement with ‘bad sleep quality’ assessed with the following equation [86, 87]:

$$\text{Prevalence Index} = \frac{|a - d|}{n}$$

Where a is the number of cases where objective and subjective measures agreed on ‘good sleep quality’ rating, d is the number of cases where objective and subjective methods agreed on ‘bad sleep quality’ rating, and n being the total number of cases.

Table 3.1: Cut-points for converting continuous variables collected using actigraphy into a 4-level categorical variable that corresponding with the 4-levelled answer to self-reported sleep quality question in the PSQI.

3-level categorical variable	SOL (min)	SE (%)	WASO (min)
Good sleep quality	<31	>84	<21
Fair sleep quality	31-45	75-84	21-41
Bad sleep quality	>46	<75	>41

Table 3.2: Criteria for creating a composite sleep quality variable based on participants meeting a quality threshold for sleep onset latency and either efficiency or wake after sleep onset as measured by actigraphy.

Variable 1	**Variable 2**		
	Good	**Fair**	**Bad**
Good	Good	Good	Fair
Fair	Good	Fair	Bad
Bad	Fair	Bad	Bad

3.3 – Results

Classification of individual and composite objective variables and subjective measurements of sleep quality are described in table 3.3. Distribution of individual objective indices, composite variables and subjective measurements of sleep quality are described in table 3.3. Almost 60% of participants self-reported having good sleep quality, with fewer than 6% rating their sleep quality as bad. This distribution was similar to objective classification of sleep quality using SE, with 55% classified as having ‘good’ sleep quality and 7% as ‘bad’. When classified by SOL, almost 98% were classed as having ‘good’ sleep quality, whereas when classified using WASO, 60% were classed as having ‘bad sleep quality’. Distribution of the four composite objective scores was skewed by the opposing directions measured using SOL and WASO. Classification by SOL and SE classed 90% as having ‘good sleep quality’, whereas when classified by WASO and SE, only 40% were classified as having ‘good’ sleep quality. All three objective measures combined into a single objective variable classed 60% of participants as having ‘fair sleep quality’.

Table 3.3: Frequency distribution of converted objective indices, composite sleep variables and subjective sleep quality derived from tables 3.1 and 3.2

		Classification using 3-tiered categorical variable n (% within each category)		
		Good SQ	**Fair SQ**	**Bad SQ**
Objective indices [a]	**SOL**	171 (97.7%)	3 (1.7%)	1 (0.6%)
	SE	94 (53.7%)	68 (38.9%)	13 (7.4%)
	WASO	7 (4.0%)	67 (38.3%)	101 (57.7%)
Composite sleep variable [b]	**SOL x SE**	161 (92.0%)	11 (6.3%)	3 (1.7%)
	SOL x WASO	71 (40.6%)	102 (58.3%)	2 (1.1%)
	SE x WASO	65 (37.1%)	36 (20.6%)	74 (42.3%)
	SOL x SE x WASO	65 (37.1%)	98 (56.0%)	12 (6.9%)
Subjective measurement	**PSQI Q6**	104 (59.4%)	61 (34.9%)	10 (5.7%)

SOL (sleep onset latency), SE (sleep efficiency), SQ (sleep quality), WASO (wake after sleep onset) [a]: Objective indices converted into 3-tiered categorical variable using cut points detailed in table 3.1.[b]: composite sleep variable created using table 3.2

Table 3.4: Frequency distribution observing agreement between classification of objective sleep quality using objective indices and composite sleep variables, and self-reported sleep quality.

		Subjective Sleep Quality n (%)			
		Good SQ (n = 104)	Fair SQ (n = 61)	Bad SQ (n = 10)	Total (n = 175)
Composite score (SOL x SE)	Good SQ	98 (61%)	53 (33%)	10 (6%)	161 (92%)
	Fair SQ	5 (45%)	6 (55%)	0 (0%)	11 (6%)
	Bad SQ	1 (33%)	2 (67%)	0 (0%)	3 (2%)
Composite score (SOL X WASO)	Good SQ	44 (62%)	23 (32%)	4 (6%)	71 (41%)
	Fair SQ	60 (59%)	36 (35%)	6 (6%)	102 (58%)
	Bad SQ	0 (0%)	2 (100%)	0 (0%)	2 (1%)
Composite score (SE x WASO)	Good SQ	40 (62%)	21 (32%)	4 (6%)	65 (37%)
	Fair SQ	22 (61%)	11 (31%)	3 (8%)	36 (21%)
	Bad SQ	42 (57%)	29 (39%)	3 (4%)	74 (42%)
Composite score (SOL x SE x WASO)	Good SQ	40 (39%)	21 (34%)	4 (40%)	65 (37%)
	Fair SQ	60 (58%)	32 (53%)	6 (60%)	98 (56%)
	Bad SQ	4 (4%)	8 (14%)	0 (0%)	12 (7%)

SOL: Sleep onset latency, SE: Sleep efficiency, WASO: Wake after sleep onset, SQ: Sleep quality

Table 3.4 describes the number of individuals that self-reported having “good”, “fair” or “bad” sleep quality that scored the same classification of sleep quality using individual (derived from table 3.1) and composite (derived from table 3.2) objective measurements of sleep quality. Further statistical analysis of the aforementioned frequency distribution is described in table 3.5. Agreement with the subjective sleep quality rating was highest when using the single objective index SOL and composite score using SOL and SE, with an overall agreement of 60% observed. This is in contrast to WASO, where the overall agreement was below 20%.

Inter-rater reliability was assessed using Cohen's weighted kappa and showed poor agreement between classifications of sleep quality using subjective and both single and composite scores of objective indices. This was reflected in the prevalence index, with every model scoring 0.87 or above except WASO and SE x WASO, scoring 0.37 and 0.52 respectively.

Table 3.5: Summary of agreement between composite measures of objective sleep quality and subjective measures of sleep quality.

		Subjective sleep quality Observed agreement n (%)						
		Good SQ n = 104	Fair SQ n = 61	Bad SQ n = 10	Overall Agreement (%)	Kappa (k) (95% CI)	P value	Prevalence Index
Objective indices	SOL	102 (98%)	1 (2%)	0 (0%)	59%	0.01 (-0.04, 0.07)	0.66	0.94
	SE	57 (55%)	23 (38%)	0 (0%)	46%	0.02 (-0.12, 0.16)	0.77	0.87
	WASO	3 (3%)	20 (33%)	6 (60%)	17%	-0.01 (-0.04, 0.06)	0.74	0.37
Composite sleep variable	SOL x SE	98 (94%)	6 (10%)	0 (0%)	59%	0.04 (-0.05, 0.13)	0.44	0.93
	SOL x WASO	44 (42%)	36 (59%)	0 (0%)	46%	0.05 (-0.09, 0.19)	0.47	0.93
	SE x WASO	40 (39%)	11 (18%)	3 (30%)	31%	0.01 (-0.09, 0.12)	0.77	0.52
	SOL x SE x WASO	40 (39%)	32 (53%)	0 (0.0%)	41%	0.05 (-0.08, 0.18)	0.45	0.87

SOL: Sleep onset latency, SE: Sleep efficiency, WASO: Wake after sleep onset, SQ: Sleep quality

3.4 – Discussion

This study suggests that, overall, there is poor agreement between subjective and objective measurements of sleep quality. Whilst overall agreement between self-reported and objective sleep quality classified using SOL and composite model SOL x SE was reasonably high (60%), inter-rater reliability scores were low. This can be explained by the prevalence effect, where the marginal distribution of observed ratings fall under one category of rating at a much higher rate than others [85]. The prevalence index (PI) is a reflection of this, where a PI of 0 indicates no imbalance whereas a PI of 1 indicates that all agreements fall into a single category. As seen in our cohort (table 3.4), almost all models of objective sleep quality bar WASO and SE – WASO composite variable had a PI of over 0.87. An inverse relationship also exists between PI and kappa values, where PI increases when k values decrease and offers an explanation as to why such low agreement is observed. This is because there is a much higher probability of agreement of 'good sleep quality' between subjective and objective measures of sleep quality resulting purely out of chance given that the distribution of both ratings was skewed towards having 'good sleep quality' (i.e. only 6% of participants had a negative self-rating of sleep). This is reiterated by both Hoehler 2000 [88] and Vach 2005 [89], whom both report that in near-homogenous populations, it becomes impossible to identify evidence to support agreement between two measures, and as such, reflected in a low kappa value.

There are several reasons why poor agreement was observed between subjective and objective measures of sleep quality. First, objective and subjective measurements of sleep quality should not be compared as they are measurements of two unique characteristics of sleep. Subjective measurements of sleep quality measure an individual's perception of how they felt about sleep, with perceived satisfaction used

as the scale. This perception is relative to an individual's own experiences, baseline or cultural norms and thus influenced by other covariables (i.e. an individual measuring a SOL of 30 minutes may report sleep quality as excellent if they have a history of taking much longer to sleep or believe that it is normal to do so). This has previously been observed by Kaplan et al. 2017, whom compared 28 objective sleep, clinical and demographic correlates of sleep quality and concluded that commonly obtained measures of polysomnographically-defined sleep contributed little to subjective ratings of prior-night sleep quality [90]. Conversely, objective indices are measured using time as the scale and further categorised by consensus standards from HIC settings [18]. Give that sleep is influenced by culture, geography and climate (see chapter 1), categorising objective indices of sleep measured in a LMIC setting by HIC consensus standards may have attributed to low agreement observed.

Second, that objective indices of sleep are able to be combined to form an overall measurement of sleep quality, as suggested by several authors [21, 22, 91, 92]. This method relies on the assumption that individual objective indices of sleep correlate to overall sleep quality, however, agreement was equally as poor between single objective indices and composite models. This poor agreement may be attributed to each of the objective indices being treated as independent, and wholly separate from other sleep characteristics. Ignoring potential interactions between objective indices of sleep could lead to incorrect classification of what is defined as 'good' objective sleep quality,' i.e. an individual with short SOL would rate as having 'good' sleep quality yet may only fall asleep quickly due to an accumulated sleep debt as a result of short sleep duration.

Third, the use of subjective measurements of sleep quality is temporally flawed. Subjective measurement of sleep quality can be made in two ways, (i) night-by-night

basis or (ii) over a defined period of time (i.e. over the last 30 days). Answering a satisfaction question based on any length of time explicitly invokes a non-systematic review of this period, leaving measures vulnerable to transient influences that draw attention to arbitrary or incomplete information (i.e. opinion of the previous night's sleep)[93]. Similarly, while a night-by-night diary has the benefit of aggregating over several nights and adjectives, it is disadvantaged as no two days (even if intentionally matched) are identical [93].

There are several limitations to this study. There was inevitable attrition. The present study achieved good follow up rates with few systematic differences between responders and non-responders, and those whom had actigraphy recorded that were included and excluded from analysis. While actigraphy was able to calculate individual indices of sleep using algorithms, it is based on wrist movement. This is in contrast to measurements obtained using a PSG, where sleep indices are obtained from sleep architecture (i.e. latency between stages of sleep). Previous studies have found that measurements of SOL were found to be incorrect and unreliable when recorded using actigraphy, where individuals who have insomnia were recorded as being asleep via actigraphy yet were awake and unmoving in bed [94, 95]. Unfortunately, as discussed in Chapter 1 and 2, the cost of performing a PSG for the purpose of epidemiological research in a LMIC setting is uneconomical. Another limitation is the polarised results obtained from subjective ratings of sleep quality. This may be a result of using HIC consensus cut-points to determine what constitutes optimal objective indices of sleep. This may explain why a prevalence effect was observed in our study, where only 5% rated their sleep quality as being 'bad' as a result of cultural and environmental stressors and again highlights the improper application of HIC consensus cut-points on what is classified as "good sleep quality".

3.5 – Conclusion

Although sleep quality is both a predictor of abnormal health and an important measure of sleep, it lacks definitional consensus. In the absence of a definition, objective indices and subjective measurements of sleep quality have been assumed to measure the same attribute of sleep. This is not the case as, at least in our sample, objective indices of SE, SOL and WASO did not agree with each other, and were unable to be combined into more useful composite indices as a 'global' measure of an individual's satisfaction with their sleep. Poor agreement between objective and subjective measures of sleep quality may also have been attributed to the incorrect application of HIC consensus standards used to classify objective indices in a LMIC setting. Given its importance as a predictor of health, vitality and wellbeing, a definitional consensus of sleep quality must be reached on sleep quality.

Chapter 4 ─ General Conclusion and Future Work

Several concluding remarks and suggested opportunities for future work can be made following the research contained in this thesis. First is the improper application of HIC consensus standards of sleep on those in LMIC settings. Sleep recommendations are based on current research; however, most of this comes from HIC settings. These recommendations inform future studies, and public health policy yet are not culturally or region specific. Given that sleep is influenced by factors such as artificial light, latitude, housing and climate (see chapter 1), further research must be done to curate cultural and region-specific recommendations of sleep. These recommendations could then be used as part of a public health campaign, aimed at reducing cardiometabolic disease which is prevalent in Sri Lanka.

Second, no definition of sleep quality has been widely accepted by either medical experts or researchers. Given that a link has been established between sleep quality as a predictor of adverse health outcomes (see chapter 1), a definitive consensus must be reached to ensure future research and public policy remain effective and relevant. The absence of a clear definition of sleep quality has also resulted in objective and subjective measurements of sleep quality incorrectly assumed to measure the same attribute, at least in a LMIC setting.

Third, the field of sleep medicine must move away from defining sleep by individual characteristics (i.e. sleep duration) and instead tread it as a multidimensional construct. Most research into the sleep-health link focuses on individual characteristics of sleep as predictors of health. While this top-down approach provides an overview of sleep, it does not take into account underlying interactions between these variables. A bottom-up approach to sleep may be better suited and instead, observe how factors

that affect sleep interact with each other and influence the overall perception of sleep quality. This is important when looking at sleep-health and improving health outcomes, and calls for a more holistic, lifestyle-oriented approach.

Appendix A

The following article is the unedited version of the journal article published as ‘Agreement Between Subjective and Objective Measures of Sleep Duration in a Low-Middle Income Country Setting’ in the special edition of Sleep Health Journal titled ‘Global perspectives of Sleep’.

Sleep Health 4 (2018) 543–550

Contents lists available at ScienceDirect

Sleep Health

Journal of the National Sleep Foundation

journal homepage: sleephealthjournal.org

Agreement between subjective and objective measures of sleep duration in a low-middle income country setting

Aaron Schokman [a], Yu Sun Bin, PhD [a,b], Guido Simonelli, MD [c], Jonathon Pye [d,e], Richard Morris, PhD [d], Athula Sumathipala, PhD [f,g], Sisira H Siribaddana, MD, FRCP [h], Matthew Hotopf, PhD [g], Fruhling Rijsdijk, PhD [g], Kaushalya Jayaweera, MD [f], Nick Glozier, MBBS, PhD [a,d,e,*]

[a] Central Clinical School, Sydney Medical School, University of Sydney, Australia
[b] Sleep Group, Charles Perkins Centre, University of Sydney, Australia
[c] Behavioural Biology Branch, Walter Reed Army Institute of Research, Silver Spring, Maryland, USA
[d] Brain and Mind Centre, University of Sydney, Australia
[e] Neurosleep, NHMRC Centre of Research Excellence, Australia
[f] Institute for Research & Development, Colombo, Sri Lanka
[g] Institute of Psychiatry, Psychology and Neurology, King's College London, London, United Kingdom
[h] Rajarata University of Sri Lanka, Sri Lanka

ARTICLE INFO

Article history:
Received 23 March 2018
Received in revised form 24 August 2018
Accepted 30 August 2018

Keywords:
Actigraphy
Self-Report
Validation
Criterion Cut-points
Sleep Duration
Sleep Quality
Low-Middle Income Country (LMIC)
Public Health
Analytical Epidemiology

ABSTRACT

Objectives: Describe sleep duration in adult Sri Lankans and determine the bias and agreement of self-report and actigraphic assessments.
Design: Validation sub-study nested within the Colombo Twin and Singleton Study (2012-2015).
Setting: Colombo, Sri Lanka.
Participants: 175 adults with actigraphy, randomly selected from 3497 participants with self-reported sleep assessed in a population-based cohort.
Measurements: Self-reported sleep duration, ascertained by the Pittsburgh Sleep Quality Index (PSQI), was compared to a minimum of four days of actigraphy. Bias and agreement were assessed using the Bland-Altman method and a novel application of criterion cut-point analysis. Objective measurements of wake after sleep onset (WASO) and sleep efficiency were evaluated.
Results: Sri Lankans have short sleep duration; averaging 6.4h (SD 1.5) self-reported and 6.0h (SD 0.9) actigraphically. Poor sleep quality was prevalent with an average WASO of 49 min., and sleep efficiency <85%. Bias was observed, with self-report consistently over-reporting sleep on average by 27.6 min (95% CI: -0.68, -0.24) compared to objective measures, but wide individual variation in disagreement, ranging from over-reporting by 3.34h to under-reporting by 2.42h. A criterion cut-point method also failed to define agreed definitions of short and long sleep duration.
Conclusions: Sleep in Sri Lankan adults, whether measured subjectively or objectively, is of short duration and suboptimal objective quality by High Income Country consensus standards. Given the high cardiometabolic morbidity in Sri Lanka and poor measurement agreement observed, this warrants further investigation and supports the need for culturally appropriate, reliable, and valid assessment for analytic epidemiology in non-Western settings.

Introduction

Interest in sleep epidemiology and the link between sleep and optimal health has grown over the last two decades.[1–4] The majority of this research is conducted in Western, high-income country (HIC) settings[1–13] which warrants the question of how relevant findings and the derived consensus statements are[5,6] to low-middle income countries (LMICs). Sleep duration and quality are known to be affected by cultural, social, environmental and geographical influences[7–9] which

Abbreviations: aTST, Objectively recorded Total Sleep Time; CI, Confidence Interval; HIC, High-Income Country; LMIC, Low-Middle Income Country; LOA, Limit of Agreement; PSG, Polysomnography; SE, Sleep Efficiency; srTST, Self-reported Total Sleep Time; TST, Total Sleep Time; WASO, Wake After Sleep Onset.

* Corresponding author at: Brain and Mind Centre, Faculty of Health and Medicine, University of Sydney, NSW 2050, Australia.
E-mail address: nick.glozier@sydney.edu.au (N. Glozier).

https://doi.org/10.1016/j.sleh.2018.08.008
2352-7218/

vary greatly between LMIC and their HIC counterparts. This is compounded by the rapid demographic and epidemiological transitions occurring in developing countries[10–12] resulting from lifestyle and cultural changes, uptake in technology, and shifts to urban living, which may have substantial population-level effects on sleep.[13,14]

Evidence on sleep from a public health perspective in low and middle-income settings is lacking in part due to logistical and financial limitations. Polysomnography (PSG) is the current "gold standard" diagnostic tool used to study sleep physiology.[15,16] However, its use in population-based sleep epidemiology is untenable, as the equipment is expensive and specialised training is required to conduct and analyse the recordings.[17] A viable alternative to PSG is the use of actigraphy, a wearable device worn on the wrist that quantifies objective sleep measures. Actigraphy has previously been validated against PSG measurements of sleep duration, quality and efficiency,[18] but is better suited for smaller studies as the devices are currently relatively expensive, require multiple days of continuous recording and an accompanying sleep diary to be simultaneously filled out for best practice.[19] A third option is the use of subjective self-reported sleep questionnaires that ask individuals about their sleep habits duration and perceived sleep quality. Even though this method of sleep measurement is relatively inexpensive and thus ideally suited to studies involving large samples, there is a lack of consensus on the validity of its use to assess sleep characteristics compared to "objective" measures.[20,21]

Population-based studies of sleep duration have at least two aims: (i) descriptive epidemiology of sleep parameters in a population, e.g. establishing trends and (ii) analytical epidemiology to ascertain potential risks of negative health outcomes borne by those defined as belonging to a category, e.g. the extremes of sleep duration. The cut-points defining these categories are often determined by consensus based upon HIC setting data[5,22] however may not be applicable in populations with different sleep distributions and demographics. Furthermore, the criterion validity and agreement of self-reported measures of behaviour and symptoms against "gold standards" vary across cultures[6]; thus the results of validation studies from HIC populations may not be applicable to other cultural settings. Validation studies often compare the accuracy of two comparable measurements using pre-set cut points yet fail to explore if other, better-suited cut-points exist, e.g. identifying agreed short and long sleepers with the smallest margin of error.[18,20,21] Criterion cut-point methods can also identify meaningful thresholds within the reference method as well as the corresponding cut-points within a second method of measurement concurrently, rather than using one as "gold standard".

This study aims to describe sleep duration in a well-ascertained sample and validate self-reported measurements of sleep duration against objective measurements in Sri Lanka, a South Asian LMIC. More specifically, the study aims to:

1 Describe the sleep duration of Sri Lankan adults using self-report and actigraphic methods.
2 Validate self-reported "subjective" sleep duration against "objective" sleep diary informed actigraphy in this setting.
3 Demonstrate whether the application of a criterion cut-point algorithm could identify agreed values discriminating short and long subjective and actigraphic measured sleep duration for use in analytic epidemiology in resource-poor settings.

Methods

This study was designed as an actigraphic sub-study nested within the Colombo Twin and Singleton Study (CoTASS) 2. The study received ethical approval from the Psychiatry, Nursing & Midwifery Research Ethics Subcommittee, King's College London, UK (ref: PNM/10/11-124), the Faculty of Medical Sciences University of Sri Jayewardenepura Ethical Review Committee (USJP ERC) (ref: 596/11) and the Research Integrity & Ethics Administration, The University of Sydney, Australia (ref: 2012/2181).

Population-based sample

CoTASS is a cohort study of twins randomly selected from the Colombo twin registry and a sample of singletons randomly selected from non-twin households, stratified by Grama Niladhari Divisions (GND, the smallest administrative unit, with approximately 4,000 people in each) from which the twins were ascertained, in Colombo, the capital of Sri Lanka. Baseline data collection took place between 2005 and 2007 as described in detail in Siribaddana et al. (2008).[23] CoTASS 2 was a follow-up study that took place between 2012 and 2015 involving 3969 participants (2934 twins and 1035 singletons) and focused on genetic and environmental influence on mental health and cardiovascular disease. Trained researchers conducted face to face interviews and collected biometric samples and validated self-report questionnaires semantically translated into Sinhala, including the Pittsburgh Sleep Quality Index (PSQI), as described in further detail in Jayaweera et al. (2017).[24]

Nested Actigraphy sub-study

Invitations to participate in the sleep sub-study were given to 626 randomly selected participants traced from the COTASS 2 sample (Fig. 1). Participants were required to wear an actigraphy wristwatch that recorded activity over a period of seven days while simultaneously completing a sleep diary. Of those invited, 95 individuals (15.2%) refused to participate, and a further 73 (11.7%) were uncontactable during recruitment (Fig. 1). The self-reported sleep duration of the 168 not consenting to, or unable to be contacted for, the actigraphy sub-study was less (6.1h) than those in our analysed sample, (6.5h), $t(420) = 3.13$, $P = 0.002$, but did not differ on other health or demographic variables.

Four hundred and fifty-eight (73.2%) cohort participants accepted the invitation to participate. However, only 268 (58.5%) had their actigraphy recorded as a result of time restrictions on COTASS2 fieldwork assessments and delays on actiwatch importation. Of those with actigraphy, 16 (6.0%) were excluded from analysis due to the participant recorded as constantly moving even when asleep. Another 24 (9.0%) were excluded as fewer than 4 nights of data were recorded, and a further 6 (2.2%) due to extended periods of time where the device was removed. Two Actiwatch devices were found to be faulty, with several recordings made with the same devices recording extensive periods of maximum movement and zero light, and all 16 (6.0%) recordings associated with these two devices were excluded. A further 10 (3.7%) failed to complete the PSQI and 21 (7.8%) did not complete the accompanying sleep diary, leaving 175 (65.3%) included in the analysis. Of the 268 participants whom had actigraphy recorded, no demographic health and sleep duration differences were observed between those who were included and excluded from the final analysis.

Data Quality

Of the 3969 participants in COTASS 2, 3672 (94.9%) answered the Pittsburgh Sleep Quality Index (PSQI). Seven (0.2%) were removed as outliers having reported spending more than 15 hours in bed (>3 s.d) and a further 168 (4.6%) removed due to an interpretation error where total sleep time (TST) was reported longer than time spent in bed leaving 3497 (95.2%) participants with subjective measurements of sleep duration included in the final sample.

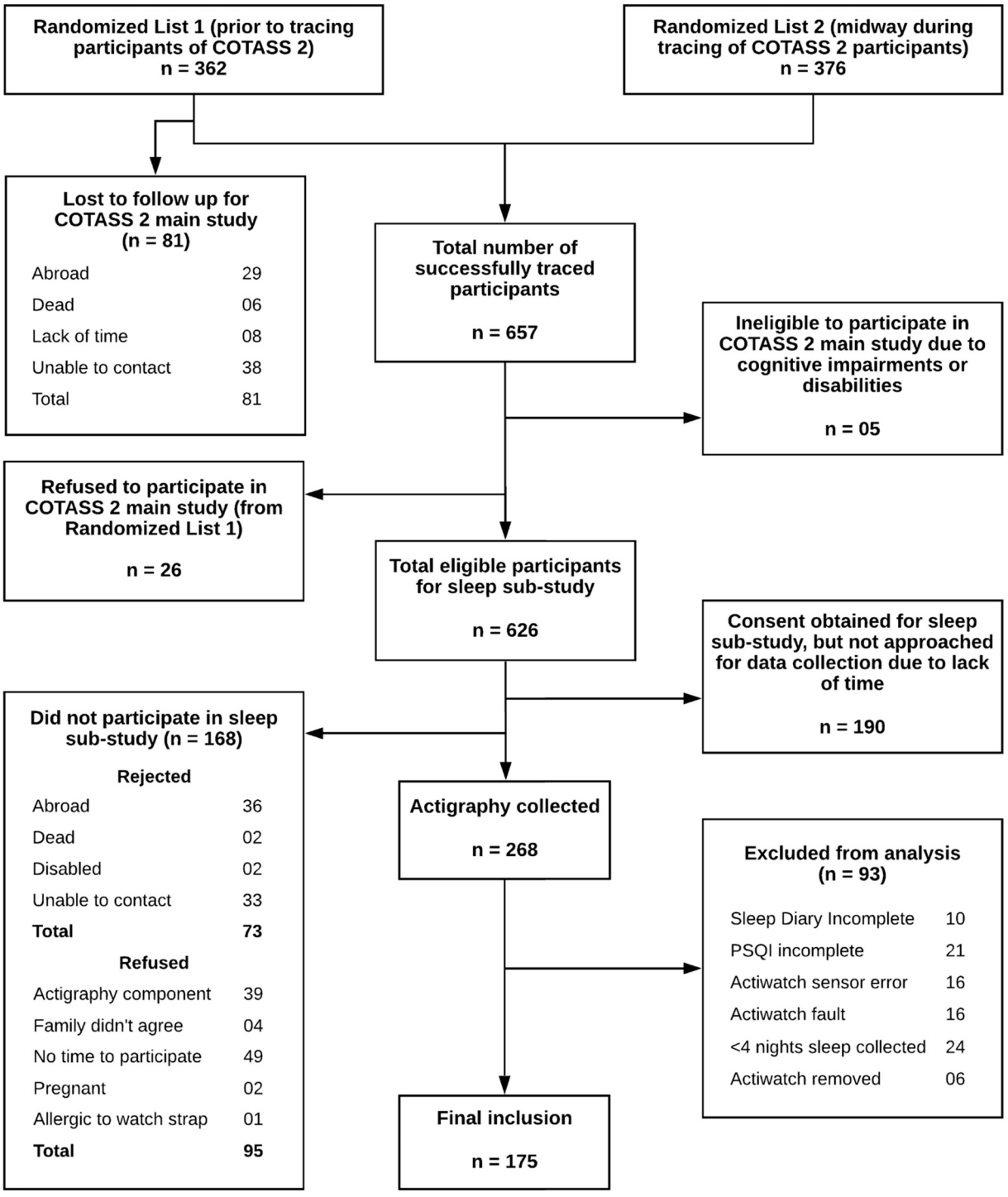


Fig. 1. Flowchart describing participant recruitment for sleep sub-study.

Measures

"Objective" sleep duration

Actigraphic measurements were collected using wrist-worn actigraphy devices (Actiwatch Spectrum Pro, Phillips Respironics, USA, firmware: 01.01.2009) which have produced reliable sleep statistics.[25] Measurements were collected in 30-second epochs over a 7-day period, with a minimum of four days of valid recording to be included in the study. Participants were required to complete a daily sleep diary and indicated bedtime and rise time using the event button on the Actiwatch device. Sleep-wake detection algorithms were used by Actiware 6.0 software (Phillips Respironics), set to a medium sensitivity threshold and 10 minutes of immobility for sleep onset.[19,26] Manual scoring of actigraphy was conducted by one of the authors (AS) based on visual inspection, sleep diary entries and Actiwatch timestamps according to standardized guidelines.[19] Average scores of total sleep time (TST) were calculated, defined as the total amount of time scored as sleep during the main rest period, which was nights for most participants. To ensure internal scoring, validation and minimization of bias, a second researcher (JP) scored

a random selection of 40 actigraphy recordings (20% of total sample) and agreement of these assessed as per standardized guidelines.[19] The resulting inter-scorer agreement fell within acceptable limits with no systematic bias observed (95% CI: -0.17, 0.01 and limit of agreement: -0.58, 0.42).

"Subjective" sleep duration

The PSQI is a standardized self-reported questionnaire that retrospectively assesses sleep of the prior 30 days.[27] Self-reported total sleep time (srTST) was calculated using question 4 of the PSQI 'during the past month, how many hours of actual sleep did you get at night?'.

Health and sociodemographic measures

Both health and sociodemographic information were collected during CoTASS 2. Height and weight were measured, and body mass index (BMI) was calculated for each. Depressive symptoms were assessed using the Beck Depression Inventory,[28] and the presence of depression was defined as a score of over 21. Alcohol abuse was assessed using the Alcohol Use Disorders Identification Test.[29] A score greater than eight indicated alcohol abuse. Participants were asked about whether they currently smoked or not ('smoker' or 'non-smoker'). Participants were also interviewed about their medical history, including hospital visits, surgery and targeted questions regarding specific illnesses and chronic diseases. A positive answer during the interview indicated the presence of chronic disease. Finally, employment status was similarly queried in the interview, and the information later dichotomized into two broad categories: "employed" (including part-time work) and "unemployed" (including students/retired).

Statistical analysis

We compared the actigraphy subsample to those of the entire CoTASS 2 population to assess representativeness. Binary variables were compared using chi-square analysis with variability assessed by odds ratio scores while continuous variables including self-reported sleep measurements were assessed using independent-sample t-tests.

Subjective and objective sleep measurements

Validation of subjective against objective measurements of sleep duration was conducted using an adapted version of the Bland-Altman plot method.[30] Actigraphy was plotted on the x-axis (as opposed to the average of the two methods) as actigraphy is used as the gold standard reference method. [31] Mean difference, confidence intervals, and limit of agreement were used to examine agreement between methods. These were recalculated following stratification by age dichotomized at the mean age of sample and sleep efficiency dichotomized at 85% as recommended by the National Sleep Foundation.[5] As only two participants had a wake after sleep onset (WASO) of < 20 min, the next cut-points of <30 min and ≥ 31 min were used to stratify measurements.[5]

Criterion validation of short and long sleep duration in this sample

We employed a criterion method defining cut points for short and long sleep duration in this sample using data from both measurement approaches. We imposed a criterion on our algorithm that at least 10% of the total sample be assigned to the short and long sleep categories and that a minimum of 40% of the sample be assigned to "normal" duration, with no result produced if these conditions are not met (Eq. A2).

Each subject was assigned two scores 'x_i', created using converted z score of self-reported TST and objective measurement of TST. Three categories were created (C_1, C_2, C_3) that corresponded to short, normal and long sleep duration and defined by the parameter Θ that consisted of a lower and upper threshold.

$$\Theta = \left[\theta_{lower}, \theta_{upper}\right]$$

θ_{lower} and θ_{upper} refer to the defining cut-points between each category and once applied to 'x_i', allowed for the creation of two class assignments, $class_j$ and $class_k$ that correspond to self-report and objective measurements respectively. A confusion matrix 'm' was created to distinguish between classes, with the sum of the diagonal directional elements a representation of error or "cost" (Eq. A3). This cost function (Eq. A4) was minimized using the default Nelder-Mead Simplex algorithm in scikit-learn (Python 3.6.1). The function was minimized over 400 starting seeds spread evenly across parameter space between z = 1.28 and -1.28. The top 5 results of this cost function (i.e. when agreement errors were the least) were applied separately to both self-report and actigraphy datasets, creating corresponding duration cut-points that identified objective and subjective short, average and long sleepers with agreement between these categories assessed using a frequency analysis table.

Results

The actigraphy subsample and overall CoTASS 2 sample characteristics are described in Table 1. The actigraphy sample was younger, less likely to abuse alcohol and possibly less depressed than the CoTASS 2 sample. No significant difference between samples was observed in sex, current smoking status, or BMI. Rates of chronic disease were high in both samples, with more than 85% of both samples reporting to suffer from chronic physical illness (ex. dental disease).

Sleep duration

Mean self-report sleep duration was low: averaging 6.4 (SD 1.5) and 6.0 (SD 1.5) hours in the actigraphy and total sample respectively (the actigraphy sample reporting slightly longer sleep duration). Mean objective sleep duration was also low. The sub-study sample spent 6 hours asleep at night and on average spent an accumulated 50 minutes awake each night after sleep onset (WASO). This is reflected in the mean sleep efficiency (SE) of 85%.

The Bland-Altman plot (Fig. 2) showed a significant systematic difference between objective and subjective measurements of TST, with participants reporting to have slept on average 27.6 minutes (-0.46h, SD = 1.47) more than recorded using actigraphy. The limit of agreement was extremely large, extending over 5.76 hours and ranging from over-reporting TST by 3.34 hours to under-reporting by 2.42 hours.

Assessment of the influence of demographics and sleep quality on bias and agreement (Table 2) showed that consistent over-reporting was present regardless of sex or age. Those older than the sample's average age (39 yrs.) had less variability, with the maximum under-reporting falling under 2 hours. There was an interaction of sleep quality with bias. Over-reporting only occurred in those who had poorer actigraphic determined WASO (≥31 min) and poorer actigraphic sleep efficiency (<85%).

Criterion validity

Only one valid criterion cut-point resulted from our cost regularization function (Appendix A) that met the *a priori* conditions of a minimum 10% of objective short and long sleep and a minimum 40% in the "normal" category. Objective cut-points of 5.4 hours and 6.4 hours respectively defined short and long sleep duration. This

Table 1
Comparison of demographics of sub-study and CoTASS 2 sample

Binary variables (n, %)		Actigraphy (n=175)	CoTASS (n=3322) [a]	Difference Odd Ratio (95% CI)
Sex	Male	73 (41.7%)	1379 (41.5%)	0.99 (0.73 – 1.35)
	Female	102 (58.3%)	1943 (58.5%)	
Comorbid disease	Yes	155 (88.6%)	2865 (86.3%)	1.23 (0.99 – 1.98)
Employed [b]		101 (57.7%)	1845 (55.6%)	1.09 (0.80 – 1.48)
Alcohol abuse		11 (6.3%)	411 (12.4%)	0.48 (0.26 – 0.89)
Depressed [c]		2 (1.2%)	172 (5.3%)	0.30 (0.07 – 1.22)
Current smoker		17 (9.7%)	422 (12.7%)	0.74 (0.44 – 1.23)
Continuous. variables		**($\bar{x}$, SD)**	**($\bar{x}$, SD)**	**Difference t (df), p**
Age (y)		38.7 (11.6)	43.3 (14.4)	4.2 (3495), $p < 0.01$
BMI (kg/m^2) [d]		23.5 (4.5)	23.8 (4.6)	0.9 (3226), $p = 0.38$
Sleep measurements		**($\bar{x}$, SD)**		
srTST (h)		6.4 (1.5)	6.0 (1.5)	3.3 (3495), $p < 0.01$
aTST (h)		6.0 (0.9)	-	-
WASO (min)		48.6 (22.8)	-	-
Sleep Efficiency (%)		84.6 (5.9)	-	-

aTST: actigraphy total sleep time, CI: Confidence Interval, PSQI: Pittsburgh Sleep Quality Index, srTST: self-reported total sleep time, WASO: Wake After Sleep Onset
[a] CoTASS 2 sample does not include actigraphy sample
[b] Part-time work considered employed and students counted as unemployed, CoTASS (n= 3318)
[c] Actigraphy (n=174), CoTASS (n = 3305)
[d] Actigraphy (n=174), CoTASS 2 (n=3054)

corresponded to the subjective cut-point of 5.5 hours for short sleep and 7.2 hours for long sleep. Our frequency distribution table (Table 3) shows that srTST did not accurately identify short sleepers with only 37% accurately identified as short and 50% misidentified as medium and 13% being long sleeper on actigraphy. There were similar errors for long sleep.

Discussion

To our knowledge, this is the first multi-method description of an adult population sleep duration and measurement validation study conducted in a LMIC in South Asia. A notable finding is a short mean sleep duration in Sri Lankan adults of between 6.0 and 6.4 hours using both objective and subjective methods. This is considerably lower than the 7-9 hours of sleep recommended by HIC consensus groups[22] and the mean self-reported sleep duration of 7.5 hours observed from a combined sample of 71883 individuals from seven LMICs in our recent meta-analysis.[7] While we cannot account for this difference, neither of the estimates above reported on a South Asian adult population. Although not evident, this short sleep duration does match the experience of the authors who live and work in Sri Lanka. Disturbed sleep was prevalent, with almost 80% of our sample having a WASO greater than half an hour and 46% having poor actigraphic sleep efficiency. This is of particular concern given previous studies have shown induced sleep disturbances results in similar physiological consequences to those seen in sleep restriction.[32] If true, short sleep duration paired with poor sleep quality may be an unexplored factor that may partially explain the existing high rates of cardiometabolic comorbidity seen in Sri Lanka.[33]

One objective of our study was to validate subjective sleep duration against objective sleep measurements for use in large-scale epidemiology. The Bland-Altman plot was selected over other measures of correlation as it is better suited to quantify agreement between two quantitative methods of measurement.[31] We observed systematic

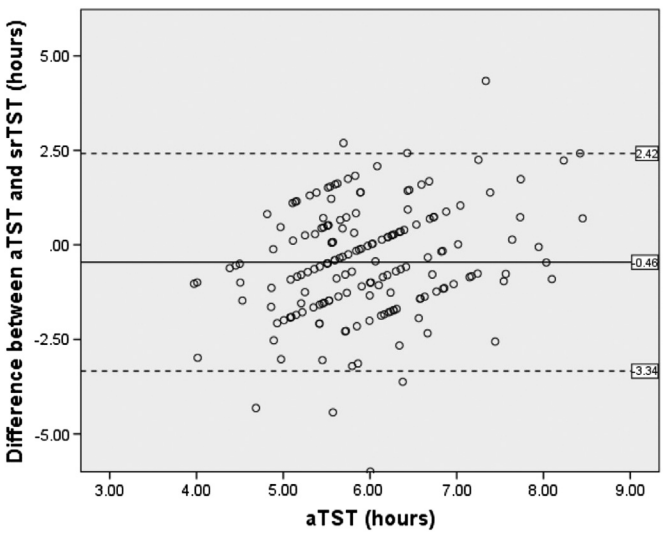


Fig. 2. Bland-Altman plot comparing objective and subjective measurements of total sleep time (TST). aTST: Actigraphy measured TST, srTST: Self-reported TST. The solid line represents the mean difference of TST at 0.46 hours (SD = 1.47). A 95% CI: (-0.68h, -0.24h) is evidence that systematic bias exists between our datasets. The dotted lines represent limit of agreement (LOA): (-3.34, 2.42).

Table 2
Analysis of bias (subjective sleep duration over-reporting vs actigraphy) and agreement after stratification by variables

Variable		n	srTST – aTST hours (SD)	95% CI	95% LOA (h)
Sex	Male	73	-0.69 (1.45)	-1.03, -0.35	2.45, -3.18
	Female	102	-0.30 (1.47)	-0.59, -0.01	2.59, -3.18
Age	< 38.7 (y)	99	-0.49 (1.66)	-0.82, -0.16	2.77, -3.74
	≥ 38.7 (y)	76	-0.43 (1.18)	-0.70, -0.16	1.89, -2.75
Comorbid disease	Yes	155	-0.36 (1.44)	-0.59, -0.14	2.45, -3.18
	No	20	-1.22 (1.55)	-1.94, -0.50	1.81, -4.25
Employed	Yes	101	-0.53 (1.29)	-0.78, -0.27	2.01, -3.06
	No	74	-0.37 (1.69)	-0.77, 0.02 [a]	2.93, -3.68
WASO	< 30 (min)	37	-0.06 (1.27)	-0.36, 0.48 [a]	2.55, -2.43
	≥ 31 (min)	138	-0.60 (1.49)	-0.85, -0.35	2.32, -3.52
Sleep efficiency	≥ 85%	94	-0.08 (1.25)	-0.33, 0.18 [a]	2.37, -2.53
	< 85%	81	-0.90 (1.58)	-1.25, -0.55	2.22, -4.00

aTST: actigraphy total sleep time, srTST: self-reported total sleep time, CI: Confidence Interval, LOA: Limit of Agreement, WASO: Wake After Sleep Onset, SE: Sleep Efficiency.
[a] No bias observed.

bias, with subjective measurements consistently over-reporting sleep on average by almost half an hour. This finding is consistent with previous validation studies between self-report and actigraphy measured sleep duration.[21,34] Stratification of our sample showed systematic bias only occurred in those that had the poorest objective measurements of sleep quality. In other words, sleep quality affected the reliability of self-reported sleep duration. When used in descriptive population epidemiology this bias may be less important as it can be accounted for if known through such sub-studies (or even discounted in time trend analysis) unless there is a strong interaction with key demographic variables, which was not seen here.

Our findings are more concerning when applied to analytic epidemiology, e.g. using self-reported sleep duration as an exposure, which requires an accurate ascertainment of sleep duration and low levels of misclassification when using categorical exposures. There was a wide range of individual differences in agreement between the two methods that ranged from being underreported by 2.5 hours and over reported by 3.5 hours. This range of nearly six hours surpassed the 2-hour maximum acceptable difference that we defined *a priori* based on the 2-hour spread between the recommended 7 – 9 hours of sleep a night.[22] This maximum acceptable difference was chosen as anything greater would result in misclassification of those that have had short sleep (<7 hours) potentially being misclassified as having long sleep (>9 hours). This was confirmed through our criterion cut-point validation. Unfortunately, no cut points were identified that could reasonably accurately classify agreed short and long duration sleepers. If a random misclassification did occur, this would bias associations to the null potentially obscuring real associations. However, as we observed over reporting of sleep duration in those with poor sleep quality (good sleepers being quite accurate), this would imply that population-based analyses of the effect of self-reported short sleep duration on, e.g. morbidity, will underestimate the effect size in those who also have poor sleep quality.

There are several reasons why this systematic bias and low agreement between self-reported and objective measurements were observed. We could not account for whether objective sleep included 'workdays' or 'days off' which would have affected agreement between the two measurements. This has been observed in previous studies, where shorter durations of sleep occurred on weekdays, with sleep duration extended on the weekend to compensate, known as social jetlag.[35,36] Another reason is the use of the PSQI as our subjective method of measurement and the fact that it requires the participants to provide estimates of sleep over the last 30 days. Lauderdale[34] shows that requesting estimates of time using "… last 30 days" produces temporally restricted estimates of sleep, while Biddle[20] suggests that agreement between subjective and objective measurements could be improved if participants were questioned using specific time periods. A '…over the last two weeks' question has been proposed to improve agreement.[20,37] The high prevalence of poor sleep quality within our sample also have contributed to a low overall agreement between measurements. Previous studies have shown that those with poor self-reported sleep quality tended to under-report sleep duration[38] in contrast to this study. It suggests movement recorded using actigraphy during the night and classified as "awake" is not perceived as such by the person when using self-report. This may be evidence of a larger validity issue around actigraphy, with low agreement observed between PSG and actigraphy in poor sleepers.[39] This may also be a limitation of the translation of the PSQI, as anecdotal evidence given by research assistants suggested that the wording of questions relating to time of sleep onset being ambiguously interpreted as the time in bed in Sinhala.

The study was limited by several factors. First, there was inevitable attrition. The present study achieved good follow up rates with few systematic differences between responders and non-responders, and those whom had actigraphy recorded that were included and excluded from analysis. The sampling difference in short sleep duration would make our observations an overestimation, thus highlighting short sleep duration in this setting. Second, a minimum of 4 nights of valid actigraphy was required by each subject to be included in the study; however standard practices recommend a minimum of 7 nights recording.[19] This was decided in part due to logistics; the number of devices available was limited due to financial constraints, and nearly one-third of participants not recording seven nights of actigraphy, having removed the watches for extensive periods. Last, the criterion cut-point method is sample specific. As such, it may be able to identify agreed boundaries for short and long sleep in other samples. Further development of a standardized criterion test would allow for the rapid comparison of two methods of measurement to determine the accuracy in identifying the extremes of distribution that are normally of interest.

Conclusion

Sri Lankan adults have a high prevalence of short sleep duration and poor sleep quality in comparison to their HIC counterparts,

Table 3
Validation of correct classification of subjects with short, normal and long sleep durations using subjective measurements against objective methods of measurement

		Classification by srTST n (% within aTST category)			
		Short	Normal	Long	Total
Classification by aTST[a]	Short	17 (37.0%)	23 (50.0%)	6 (13.0%)	46 (100%)
	Normal	20 (26.7%)	39 (52.0%)	16 (21.3%)	75 (100%)
	Long	7 (13.0%)	23 (42.6%)	24 (44.4%)	54 (100%)
	Total	44 (25.1%)	85 (48.6%)	46 (26.3%)	175 (100%)

[a] Classification was determined using objective cut-points of 5.42 and 6.42 hours and subjective cut points of 5.48 and 7.22 hours to determine short, normal and long sleep duration.

potentially questioning the applicability of consensus statements[5,6] derived from HIC samples. Basic things like security, lack of stable housing, poverty and hunger are all factors that affect sleep quality and duration[10,40,41] and the effect of geography and climate on sleep is still not fully understood.[9,42] Anecdotally, these themes were reflected in sleep diaries, with individuals commonly reporting disturbed sleep due to shared sleeping quarters, heat, and safety concerns. If true, these findings suggest sleep disturbance as a new avenue for assessing the causes of the very high obesity and diabetes rates in Sri Lanka. Nevertheless, low levels of agreement between methods and difficulty in ascertaining reliable classification advocates for caution when interpreting epidemiological findings. This study demonstrates the need for culturally relevant sleep recommendations, consistency in metrics, and exploration into the feasibility, reliability and validity of more economical devices used to measure sleep.

Disclosure

Material has been reviewed by the Walter Reed Army Institute of Research. There is no objection to its presentation and publication. The opinions or assertions contained herein are the private views of the author, and are not to be construed as official, or as reflecting the views of the Department of the Army or the Department of Defence. The investigators have adhered to the policies for protection of human subjects as prescribed in AR 70–25.

Conflicts of Interest

None of the authors have any relevant conflicts of interest to report.

Appendix A

Equation 1: Threshold function

$$\Theta = \left[\theta_{lower}, \theta_{upper}\right]$$

Equation 2: Class and condition equations

$$class_j\,(x_i;\Theta) = \begin{cases} C_1, & if\ \ x_i^1 < \theta_{lower} \\ C_2, & if\ \ x_i^1 > \theta_{lower} \text{ and } x_i^1 < \theta_{upper} \\ C_3, & if\ \ x_i^1 > \theta_{upper} \end{cases}$$

$$class_k\,(x_i;\Theta) = \begin{cases} C_1, & if\ \ x_i^2 < \theta_{lower} \\ C_2, & if\ \ x_i^2 > \theta_{lower} \text{ and } x_i^2 < \theta_{upper} \\ C_3, & if\ \ x_i^2 > \theta_{upper}. \end{cases}$$

Equation 3: Confusion matrix

$$m = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}$$

Where c_{ij} is the number of people belonging to classi in classj. Thus, the diagonal elements of m indicate the correspondence between the two dependent variables given thresholds f, while the off-diagonal elements indicate individuals who are classified into different sleep classes according to the different dependent measures. The sum of the off-diagonal elements of 'm' were included in the cost function, along with a regularization term.

Equation 4: Cost-regularization function

$$cost = \sum_{i \neq j,\ \ i,j=1\ldots3} m_{ij} + L\ |\,\Theta\,|$$

Where $|\Theta|$ refers to L_2 norm and L is a free parameter controlling regularization. We selected L and accepted the best parameter result, which categorized a minimum of 40 percent of the sample in the `middle' category, and a minimum of 10 percent in the short and long categories. Professor Glozier reports grants by the National Health and Medical Research Council (NHMRC) of Australia; Program Grant 566529.

References

1. Cappuccio FP, Cooper D, D'Elia L, Strazzullo P, Miller MA. Sleep duration predicts cardiovascular outcomes: a systematic review and meta-analysis of prospective studies. *Eur Heart J*. 2011;32(12):1484–1492.
2. Goel N, Rao H, Durmer JS, Dinges DF. Neurocognitive Consequences of Sleep Deprivation. *Semin Neurol*. 2009;29(4):320–339.
3. Cappuccio FP, D'Elia L, Strazzullo P, Miller MA, et al. Sleep Duration and All-Cause Mortality: A Systematic Review and Meta-Analysis of Prospective Studies. *Sleep*. 2010;33(5):585–592.
4. Knutson KL, Spiegel K, Penev P, Van Cauter E. The Metabolic Consequences of Sleep Deprivation. *Sleep Med Rev*. 2007;11(3):163–178.
5. Ohayon M, Wickwire EM, Hirshkowitz M, et al. National Sleep Foundation's sleep quality recommendations: first report. *Sleep Health*. 2017;3(1):6–19.
6. Goldberg DP, Gater R, Sartorius N, et al. The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psychol Med*. 1997;27(1):191–197.
7. Simonelli G, Marshall NS, Grillakis A, Miller CB, Hoyos CM, Glozier N. Sleep Health Epidemiology in Low and Middle-Income Countries: a Systematic Review, and Meta-Analysis of the Prevalence of Poor Sleep Quality and Sleep Duration. *Sleep Health*. 2018 [In Press].
8. Steger b, Kroll-smith S. Negotiating Daytime Sleep: An Intercultural Comparison of Social and Cultural Aspects of Sleep Patterns. A joint symposium of the WFSRSMS and the ESRS; 2006.
9. Brockmann PE, Gozal D, Villarroel L, Damiani F, Nuñez F, Cajochen C, et al. Geographic latitude and sleep duration: A population-based survey from the Tropic of Capricorn to the Antarctic Circle. *Chronobiol Int*. 2017;34(3):373–381.
10. Stranges S, Tigbe W, Gómez-Olivé FX, Thorogood M, Kandala N-B. Sleep Problems: An Emerging Global Epidemic? Findings From the INDEPTH WHO-SAGE Study Among More Than 40,000 Older Adults From 8 Countries Across Africa and Asia. *Sleep*. 2012;35(8):1173–1181.
11. Salomon JA, Murray CJ. The epidemiologic transition revisited: compositional models for causes of death by age and sex. *Popul Dev Rev*. 2002;28(2):205–228.
12. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med*. 2006;3(11):e442.
13. Hillman DR, Murphy AS, Antic R, Pezzullo L. The economic cost of sleep disorders. *Sleep*. 2006;29(3):299–305.
14. Daley M, Morin CM, LeBlanc M, Grégoire J-P, Savard J. The economic burden of insomnia: direct and indirect costs for individuals with insomnia syndrome, insomnia symptoms, and good sleepers. *Sleep*. 2009;32(1):55–64.
15. Kushida CA, Littner MR, Morgenthaler T, et al. Practice Parameters for the Indications for Polysomnography and Related Procedures: An Update for 2005. *Sleep*. 2005;28(4):499–523.
16. Hobson JA. A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects: A. Rechtschaffen and A. Kales (Editors).(Public Health Service, US Government Printing Office, Washington, DC, 1968, 58 p., $4.00). *Electroencephalogr Clin Neurophysiol*. 1969;26(6):644.
17. Kushida CA, Chang A, Gadkary C, Guilleminault C, Carrillo O, Dement WC. Comparison of actigraphic, polysomnographic, and subjective assessment of sleep parameters in sleep-disordered patients. *Sleep Med*. 2001;2(5):389–396.
18. Marino M, Li Y, Rueschman MN, et al. Measuring Sleep: Accuracy, Sensitivity, and Specificity of Wrist Actigraphy Compared to Polysomnography. *Sleep*. 2013;36(11):1747–1755.
19. Neurosleep. Manual for Actigraphy Standardisation. Data Acquisition, Analysis and Reporting. Australia: National Health and Medical Research Council Centre of Research Excellence; 2017.
20. Biddle DJ, Robillard R, Hermens DF, Hickie IB, Glozier N. Accuracy of self-reported sleep parameters compared with actigraphy in young people with mental ill-health. *Sleep health*. 2015;1(3):214–220.
21. Girschik J, Fritschi L, Heyworth J, Waters F. Validation of Self-Reported Sleep Against Actigraphy. *J Epidemiol*. 2012;22(5):462–468.
22. Hirshkowitz M, Whiton K, Albert SM, et al. National Sleep Foundation's sleep time duration recommendations: methodology and results summary. *Sleep Health*. 2015;1(1):40–43.
23. Siribaddana SH, Ball HA, Hewage SN, et al. Colombo Twin and Singleton Study (CoTASS): A description of a population based twin study of mental disorders in Sri Lanka. *BMC Psychiatry*. 2008;8:49.

24. Jayaweera K, Aschan L, Pannala G, et al. The Colombo Twin and Singleton Follow-up Study: a population based twin study of psychiatric disorders and metabolic syndrome in Sri Lanka. *BMC Public Health*. 2018;18(1):145.
25. Respironics P. Equivalence of activity recordings and derived sleep statistics; Actiwatch-64, Actiwatch 2 and Actiwatch Spectrum; 2009.
26. Jobert M, Wilson FJ, Roth T, et al. Guidelines for the recording and evaluation of pharmaco-sleep studies in man: the International Pharmaco-EEG Society (IPEG). *Neuropsychobiology*. 2013;67(3):127–167.
27. Buysse DJ, Reynolds Iii CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Res*. 1989;28(2):193–213.
28. Beck AT, Steer RA, Brown GK. Beck depression inventory-II. *San Antonio*. 1996;78(2):490–498.
29. Saunders JB, Aasland OG, Babor TF, De la Fuente JR, Grant M. Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption-II. *Addiction*. 1993;88(6):791–804.
30. Bland JM, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;327(8476):307–310.
31. Giavarina D. Understanding Bland-Altman analysis. *Biochem Med (Zagreb)*. 2015;25(2):141–151.
32. Kahn M, Fridenson S, Lerer R, Bar-Haim Y, Sadeh A. Effects of one night of induced night-wakings versus sleep restriction on sustained attention and mood: a pilot study. *Sleep Med*. 2014;15(7):825–832.
33. Katulanda P, Constantine G, Mahesh J, et al. Prevalence and projections of diabetes and pre-diabetes in adults in Sri Lanka—Sri Lanka Diabetes, Cardiovascular Study (SLDCS). *Diabet Med*. 2008;25(9):1062–1069.
34. Lauderdale DS, Knutson KL, Yan LL, Liu K, Rathouz PJ, et al. Self-reported and measured sleep duration: how similar are they? *Epidemiology*. 2008;19(6):838–845.
35. Monk TH, Buysse DJ, Kennedy KS, Potts JM, DeGrazia JM, Miewald JM. Measuring sleep habits without using a diary: the sleep timing questionnaire. *Sleep*. 2003;26(2):208–212.
36. Roepke SE, Duffy JF. Differential impact of chronotype on weekday and weekend sleep timing and duration. *Nat Sci Sleep*. 2010;2:213.
37. Wolfson AR, Carskadon MA, Acebo C, et al. Evidence for the validity of a sleep habits survey for adolescents. *Sleep*. 2003;26(2):213–216.
38. Regestein QR, Friebely J, Shifren JL, et al. Self-reported sleep in postmenopausal women. *Menopause*. 2004;11(2):198–207.
39. Sadeh A. The role and validity of actigraphy in sleep medicine: an update. *Sleep medicine reviews*. 2011;15(4):259–267.
40. Simonelli G, Patel SR, Rodríguez-Espínola S, et al. The impact of home safety on sleep in a Latin American country. *Sleep Health*. 2015;1(2):98–103.
41. Simonelli G, Leanza Y, Boilard A, et al. Sleep and Quality of Life in Urban Poverty: The Effect of a Slum Housing Upgrading Program. *Sleep*. 2013;36(11):1669–1676.
42. Ohayon MM, Milesi C. Artificial Outdoor Nighttime Lights Associate with Altered Sleep Behavior in the American General Population. *Sleep*. 2016;39(6):1311–1320.

References

1. Barbera, J., *Sleep and dreaming in Greek and Roman philosophy.* Sleep Medicine, 2008. **9**(8): p. 906-910.
2. Ekirch, A.R., *At Day's Close: Night in Times Past*. 2006: W. W. Norton.
3. Economics, D.A., *Asleep on the job: Costs of inadequate sleep in Australia*. 2017, Report for Sleep Health Foundation.
4. Lallukka, T., et al., *Association of sleep duration and sleep quality with the physical, social, and emotional functioning among Australian adults.* Sleep Health: Journal of the National Sleep Foundation, 2018. **4**(2): p. 194-200.
5. Cappuccio, F.P., et al., *Sleep duration predicts cardiovascular outcomes: a systematic review and meta-analysis of prospective studies.* European Heart Journal, 2011. **32**(12): p. 1484-1492.
6. Buxton, O.M., et al., *Sleep Restriction for 1 Week Reduces Insulin Sensitivity in Healthy Men.* Diabetes, 2010. **59**(9): p. 2126.
7. Irwin, M.R., R. Olmstead, and J.E. Carroll, *Sleep Disturbance, Sleep Duration, and Inflammation: A Systematic Review and Meta-Analysis of Cohort Studies and Experimental Sleep Deprivation.* Biological Psychiatry. **80**(1): p. 40-52.
8. Freeman, D., et al., *The effects of improving sleep on mental health (OASIS): a randomised controlled trial with mediation analysis.* The Lancet Psychiatry, 2017. **4**(10): p. 749-758.
9. Mehta, A., et al. *MINDFULLNESS-BASED STRESS REDUCTION IMPROVES TOTAL SLEEP TIME IN VETERANS WITH PTSD*. in *SLEEP*. 2012. AMER ACAD SLEEP MEDICINE ONE WESTBROOK CORPORATE CTR, STE 920, WESTCHESTER, IL 60154 USA.
10. Neikrug, A.B., M.R. Crawford, and J.C. Ong, *Behavioral Sleep Medicine Services for Hypersomnia Disorders: A Survey Study.* Behav Sleep Med, 2017. **15**(2): p. 158-171.
11. Ko, H.J. and C.H. Youn, *Effects of laughter therapy on depression, cognition and sleep among the community-dwelling elderly.* Geriatrics & Gerontology International, 2011. **11**(3): p. 267-274.
12. Cross, N., et al., *Napping in older people 'at risk'of dementia: relationships with depression, cognition, medical burden and sleep quality.* Journal of sleep research, 2015. **24**(5): p. 494-502.
13. Consensus Conference, P., et al., *Recommended Amount of Sleep for a Healthy Adult: A Joint Consensus Statement of the American Academy of Sleep Medicine and Sleep Research Society.* Sleep, 2015. **38**(6): p. 843-844.
14. Hirshkowitz, M., et al., *National Sleep Foundation's sleep time duration recommendations: methodology and results summary.* Sleep Health: Journal of the National Sleep Foundation. **1**(1): p. 40-43.
15. Bliwise, D.L. and T.B. Young, *The Parable of Parabola: What the U-Shaped Curve Can and Cannot Tell Us about Sleep.* Sleep, 2007. **30**(12): p. 1614-1615.
16. Cappuccio, F.P., et al., *Sleep Duration and All-Cause Mortality: A Systematic Review and Meta-Analysis of Prospective Studies.* Sleep, 2010. **33**(5): p. 585-592.
17. Buysse, D.J., *Sleep Health: Can We Define It? Does It Matter?* Sleep, 2014. **37**(1): p. 9-17.
18. Ohayon, M., et al., *National Sleep Foundation's sleep quality recommendations: first report.* Sleep Health, 2017. **3**(1): p. 6-19.
19. Gradisar, M., et al., *The Sleep and Technology Use of Americans: Findings from the National Sleep Foundation's 2011 Sleep in America Poll.* Journal of Clinical Sleep Medicine : JCSM : Official Publication of the American Academy of Sleep Medicine, 2013. **9**(12): p. 1291-1299.
20. Tang, N.K.Y., et al., *Changes in Sleep Duration, Quality, and Medication Use Are Prospectively Associated With Health and Well-being: Analysis of the UK Household Longitudinal Study.* Sleep, 2017. **40**(3): p. zsw079-zsw079.

21. Girschik, J., et al., *Validation of Self-Reported Sleep Against Actigraphy.* Journal of Epidemiology, 2012. **22**(5): p. 462-468.
22. Krystal, A.D. and J.D. Edinger, *Measuring sleep quality.* Sleep medicine, 2008. **9**: p. S10-S17.
23. Matricciani, L., et al., *Rethinking the sleep-health link.* Sleep Health: Journal of the National Sleep Foundation, 2018.
24. Czeisler, C.A., et al., *Human sleep: its duration and organization depend on its circadian phase.* Science, 1980. **210**(4475): p. 1264-1267.
25. Peuhkuri, K., N. Sihvola, and R. Korpela, *Diet promotes sleep duration and quality.* Nutrition research, 2012. **32**(5): p. 309-319.
26. Driver, H.S. and S.R. Taylor, *Exercise and sleep.* Sleep medicine reviews, 2000. **4**(4): p. 387-402.
27. Ohayon, M.M. and C. Milesi, *Artificial Outdoor Nighttime Lights Associate with Altered Sleep Behavior in the American General Population.* Sleep, 2016. **39**(6): p. 1311-1320.
28. Brockmann, P.E., et al., *Geographic latitude and sleep duration: A population-based survey from the Tropic of Capricorn to the Antarctic Circle.* Chronobiology International, 2017. **34**(3): p. 373-381.
29. Knutson, K.L., *Sociodemographic and cultural determinants of sleep deficiency: Implications for cardiometabolic disease risk.* Social Science & Medicine, 2013. **79**: p. 7-15.
30. Hafner, M., et al., *Why Sleep Matters—The Economic Costs of Insufficient Sleep: A Cross-Country Comparative Analysis.* Rand Health Quarterly, 2017. **6**(4): p. 11.
31. Dewald, J.F., et al., *The influence of sleep quality, sleep duration and sleepiness on school performance in children and adolescents: A meta-analytic review.* Sleep Medicine Reviews, 2010. **14**(3): p. 179-189.
32. Kushida, C.A., et al., *Practice Parameters for the Indications for Polysomnography and Related Procedures: An Update for 2005.* Sleep, 2005. **28**(4): p. 499-523.
33. Douglas, N.J., S. Thomas, and M.A. Jan, *Clinical value of polysomnography.* The Lancet, 1992. **339**(8789): p. 347-350.
34. Hobson, J.A., *A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects: A. Rechtschaffen and A. Kales (Editors).(Public Health Service, US Government Printing Office, Washington, DC, 1968, 58 p., $4.00).* Electroencephalography and clinical neurophysiology, 1969. **26**(6): p. 644.
35. Standards of Practice Committee of the American Sleep Disorders, A., *Practice Parameters for the Use of Polysomnography in the Evaluation of Insomnia.* Sleep, 1995. **18**(1): p. 55-57.
36. Martin, J.L. and A.D. Hakim, *Wrist Actigraphy.* Chest, 2011. **139**(6): p. 1514-1527.
37. Marino, M., et al., *Measuring Sleep: Accuracy, Sensitivity, and Specificity of Wrist Actigraphy Compared to Polysomnography.* Sleep, 2013. **36**(11): p. 1747-1755.
38. Kushida, C.A., et al., *Comparison of actigraphic, polysomnographic, and subjective assessment of sleep parameters in sleep-disordered patients.* Sleep Medicine, 2001. **2**(5): p. 389-396.
39. Wang, S.-Y., H.-J. Chang, and C.-C. Lin, *Sleep disturbances among patients with non-small cell lung cancer in Taiwan: congruence between sleep log and actigraphy.* Cancer nursing, 2010. **33**(1): p. E11-E17.
40. Sánchez-Ortuño, M.M., et al., *Home is where sleep is: an ecological approach to test the validity of actigraphy for the assessment of insomnia.* Journal of clinical sleep medicine: JCSM: official publication of the American Academy of Sleep Medicine, 2010. **6**(1): p. 21.
41. Natale, V., G. Plazzi, and M. Martoni, *Actigraphy in the assessment of insomnia: a quantitative approach.* Sleep, 2009. **32**(6): p. 767-771.
42. Neurosleep, *Manual for Actigraphy Standardisation*, in *Data Acquisition, Analysis and Reporting*. 2017, National Health and Medical Research Council Centre of Research Excellence: Australia.

43. Jean-Louis, G., et al., *Sleep detection with an accelerometer actigraph: comparisons with polysomnography.* Physiology & behavior, 2001. **72**(1-2): p. 21-28.
44. Cole, R.J., et al., *Automatic sleep/wake identification from wrist activity.* Sleep, 1992. **15**(5): p. 461-469.
45. Buysse, D.J., et al., *The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research.* Psychiatry Research, 1989. **28**(2): p. 193-213.
46. Lauderdale, D.S., et al., *Self-reported and measured sleep duration: how similar are they?* Epidemiology (Cambridge, Mass.), 2008. **19**(6): p. 838-845.
47. Silva, G.E., et al., *Relationship between reported and measured sleep times: the sleep heart health study (SHHS).* Journal of clinical sleep medicine: JCSM: official publication of the American Academy of Sleep Medicine, 2007. **3**(6): p. 622.
48. Mezick, E.J., R.R. Wing, and J.M. McCaffery, *Associations of self-reported and actigraphy-assessed sleep characteristics with body mass index and waist circumference in adults: moderation by gender.* Sleep medicine, 2014. **15**(1): p. 64-70.
49. Biddle, D.J., et al., *Accuracy of self-reported sleep parameters compared with actigraphy in young people with mental ill-health.* Sleep Health: Journal of the National Sleep Foundation. **1**(3): p. 214-220.
50. Walker, M., *Why We Sleep: Unlocking the Power of Sleep and Dreams*. 2017: Scribner.
51. Stranges, S., et al., *Sleep Problems: An Emerging Global Epidemic? Findings From the INDEPTH WHO-SAGE Study Among More Than 40,000 Older Adults From 8 Countries Across Africa and Asia.* Sleep, 2012. **35**(8): p. 1173-1181.
52. Steger, b. and S. Kroll-smith, *Negotiating Daytime Sleep: An Intercultural Comparison of Social and Cultural Aspects of Sleep Patterns. A joint symposium of the WFSRSMS and the ESRS.* 2006.
53. King, A.C., et al., *Moderate-intensity exercise and self-rated quality of sleep in older adults: a randomized controlled trial.* Jama, 1997. **277**(1): p. 32-37.
54. Grandner, M.A., R.A.L. Gallagher, and N.S. Gooneratne, *The use of technology at night: impact on sleep and health.* Journal of clinical sleep medicine: JCSM: official publication of the American Academy of Sleep Medicine. **9**(12): p. 1301.
55. Simonelli, G., et al., *Sleep and Quality of Life in Urban Poverty: The Effect of a Slum Housing Upgrading Program.* Sleep, 2013. **36**(11): p. 1669-1676.
56. Simonelli, G., et al., *The impact of home safety on sleep in a Latin American country.* Sleep Health, 2015. **1**(2): p. 98-103.
57. Epstein, R., et al., *Changing from communal to familial sleep arrangement in the Kibbutz: effects on sleep quality.* Sleep, 1997. **20**(5): p. 334-339.
58. Goel, N., et al., *Neurocognitive Consequences of Sleep Deprivation.* Seminars in neurology, 2009. **29**(4): p. 320-339.
59. Knutson, K.L., et al., *The Metabolic Consequences of Sleep Deprivation.* Sleep medicine reviews, 2007. **11**(3): p. 163-178.
60. Goldberg, D.P., et al., *The validity of two versions of the GHQ in the WHO study of mental illness in general health care.* Psychological medicine, 1997. **27**(1): p. 191-197.
61. Simonelli, G., et al., *Sleep health epidemiology in low and middle-income countries: a systematic review and meta-analysis of the prevalence of poor sleep quality and sleep duration.* Sleep Health, 2018. **4**(3): p. 239-250.
62. Salomon, J.A. and C.J. Murray, *The epidemiologic transition revisited: compositional models for causes of death by age and sex.* Population and Development Review, 2002. **28**(2): p. 205-228.
63. Mathers, C.D. and D. Loncar, *Projections of global mortality and burden of disease from 2002 to 2030.* PLoS medicine, 2006. **3**(11): p. e442.
64. Hillman, D.R., et al., *The economic cost of sleep disorders.* Sleep, 2006. **29**(3): p. 299-305.

65. Daley, M., et al., *The economic burden of insomnia: direct and indirect costs for individuals with insomnia syndrome, insomnia symptoms, and good sleepers.* Sleep, 2009. **32**(1): p. 55-64.
66. Biddle, D.J., et al., *Accuracy of self-reported sleep parameters compared with actigraphy in young people with mental ill-health.* Sleep health, 2015. **1**(3): p. 214-220.
67. Hirshkowitz, M., et al., *National Sleep Foundation's sleep time duration recommendations: methodology and results summary.* Sleep Health, 2015. **1**(1): p. 40-43.
68. Siribaddana, S.H., et al., *Colombo Twin and Singleton Study (CoTASS): A description of a population based twin study of mental disorders in Sri Lanka.* BMC Psychiatry, 2008. **8**: p. 49-49.
69. Anandakumar, D., et al., *Validation of the Sinhala version of the Pittsburgh sleep quality index.* Ceylon Medical Journal, 2016. **61**(1).
70. Jayaweera, K., et al., *The Colombo Twin and Singleton Follow-up Study: a population based twin study of psychiatric disorders and metabolic syndrome in Sri Lanka.* BMC Public Health, 2018. **18**(1): p. 145.
71. Respironics, P., *Equivalence of activity recordings and derived sleep statistics; Actiwatch-64, Actiwatch 2 and Actiwatch Spectrum. 2009*.
72. Jobert, M., et al., *Guidelines for the recording and evaluation of pharmaco-sleep studies in man: the International Pharmaco-EEG Society (IPEG).* Neuropsychobiology, 2013. **67**(3): p. 127-167.
73. Beck, A.T., R.A. Steer, and G.K. Brown, *Beck depression inventory-II.* San Antonio, 1996. **78**(2): p. 490-8.
74. Saunders, J.B., et al., *Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption-II.* Addiction, 1993. **88**(6): p. 791-804.
75. Bland, J.M. and D. Altman, *Statistical methods for assessing agreement between two methods of clinical measurement.* The lancet, 1986. **327**(8476): p. 307-310.
76. Giavarina, D., *Understanding bland altman analysis.* Biochemia medica: Biochemia medica, 2015. **25**(2): p. 141-151.
77. Kahn, M., et al., *Effects of one night of induced night-wakings versus sleep restriction on sustained attention and mood: a pilot study.* Sleep Medicine, 2014. **15**(7): p. 825-832.
78. Katulanda, P., et al., *Prevalence and projections of diabetes and pre-diabetes in adults in Sri Lanka—Sri Lanka Diabetes, Cardiovascular Study (SLDCS).* Diabetic Medicine, 2008. **25**(9): p. 1062-1069.
79. Monk, T.H., et al., *Measuring sleep habits without using a diary: the sleep timing questionnaire.* Sleep, 2003. **26**(2): p. 208-212.
80. Roepke, S.E. and J.F. Duffy, *Differential impact of chronotype on weekday and weekend sleep timing and duration.* Nature and science of sleep, 2010. **2**: p. 213.
81. Wolfson, A.R., et al., *Evidence for the validity of a sleep habits survey for adolescents.* Sleep, 2003. **26**(2): p. 213-216.
82. Regestein, Q.R., et al., *Self-reported sleep in postmenopausal women.* Menopause, 2004. **11**(2): p. 198-207.
83. Sadeh, A., *The role and validity of actigraphy in sleep medicine: an update.* Sleep medicine reviews, 2011. **15**(4): p. 259-267.
84. Buysse, D.J., et al., *Recommendations for a standard research assessment of insomnia.* Sleep, 2006. **29**(9): p. 1155-1173.
85. Hallgren, K.A., *Computing inter-rater reliability for observational data: an overview and tutorial.* Tutorials in quantitative methods for psychology, 2012. **8**(1): p. 23.
86. Byrt, T., J. Bishop, and J.B. Carlin, *Bias, prevalence and kappa.* Journal of clinical epidemiology, 1993. **46**(5): p. 423-429.

87. Sim, J. and C.C. Wright, *The kappa statistic in reliability studies: use, interpretation, and sample size requirements.* Physical therapy, 2005. **85**(3): p. 257-268.
88. Hoehler, F.K., *Bias and prevalence effects on kappa viewed in terms of sensitivity and specificity.* Journal of clinical epidemiology, 2000. **53**(5): p. 499-503.
89. Vach, W., *The dependence of Cohen's kappa on the prevalence does not matter.* Journal of clinical epidemiology, 2005. **58**(7): p. 655-661.
90. Kaplan, K.A., et al., *When a gold standard isn't so golden: Lack of prediction of subjective sleep quality from sleep polysomnography.* Biological psychology, 2017. **123**: p. 37-46.
91. Åkerstedt, T., et al., *The Subjective Meaning of Good Sleep, An Intraindividual Approach Using the Karolinska Sleep Diary.* Perceptual and Motor Skills, 1994. **79**(1): p. 287-296.
92. Harvey, A.G., et al., *The Subjective Meaning of Sleep Quality: A Comparison of Individuals with and without Insomnia.* Sleep, 2008. **31**(3): p. 383-393.
93. Krueger, A.B. and D.A. Schkade, *The reliability of subjective well-being measures.* Journal of public economics, 2008. **92**(8-9): p. 1833-1845.
94. Hauri, P.J. and J. Wisbey, *Wrist actigraphy in insomnia.* Sleep, 1992. **15**(4): p. 293-301.
95. Chambers, M.J., *Actigraphy and insomnia: A closer look Part 1.* Sleep, 1994. **17**(5): p. 405-408.