DTU Library

Global and large-scale study of complex patterns in high-resolution sleep activity data

Jónasdóttir, Sigriður Svala

Publication date:
2021

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
Jónasdóttir, S. S. (2021). *Global and large-scale study of complex patterns in high-resolution sleep activity data*. Technical University of Denmark.

Global and large-scale study of complex patterns in high-resolution sleep activity data

A DISSERTATION PRESENTED
BY
SIGGA SVALA JÓNASDÓTTIR
TO
THE DEPARTMENT OF APPLIED MATHEMATICS AND COMPUTER SCIENCE

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN THE SUBJECT OF
COMPUTATIONAL SOCIAL SCIENCE

TECHNICAL UNIVERSITY OF DENMARK
COPENHAGEN, DENMARK
FEBRUARY 2021

Global and large-scale study of complex patterns in high-resolution sleep activity data

Abstract

Sleep is a complex physiological process influenced by intrinsic and extrinsic factors. Everyday, each person on earth sleeps and if they do not, it will impact them. For that reason sleep has been studied across the population at large scale, but the research has been limited by self-reported and subjective data-sets known to recall biases. Some of the key metrics and methods were developed to cater to these types of data, but today sleep recording technology has been revolutionised and wearable devices enable objective recordings in-situ over long period of time. With rising numbers of wearable device owners and studies using this technology, I see a great potential to accelerate our understanding of human sleep in modern society. This study sets out to develop new methods and metrics appropriate for multi-night recordings of sleep in-situ. Furthermore, I investigate whether current knowledge regarding sleep patterns persist when explored with a global, large-scale and high-resolution sleep activity data-set, but also seek to expand on some the fundamental knowledge.
I find detailed sleep trajectories to have complex and multidimensional patterns across the population. I introduce new features and visualisation methods, and a novel data-driven metric which may be indicative of whether individual physiological sleep needs are met or not. Furthermore, I study age-related changes in sleep timing, duration and life-stage dependent gender differences. I find novel and unprecedented results regarding associated changes in sleep due to travel, and show that regional policy and cultural context exerts strong influence on sleep behavior.

Contents

Listing of figures

This thesis is based on the following papers:

1. Jonasdottir, Sigga Svala, Kelton Minor, and Sune Lehmann. "Gender differences in nighttime sleep patterns and variability across the adult lifespan: a global-scale wearables study." Sleep 44.2 (2021): zsaa169.

2. Jonasdottir, Sigga Svala, James P. Bagrow, and Sune Lehmann. "Travel serves to balance skewed sleep patterns." (Submitted to Nature Human Behavior (January 2021) and currently under review).

3. Jonasdottir, Sigga Svala, Jari Saramaki, and Sune Lehmann. "Quantifying complex patterns in high-resolution sleep activity data." (2021)

Other publications:

- Minor, Kelton, et al. "Ambient heat and human sleep." arXiv preprint arXiv:2011.07161 (2020).

In the dissertation, I reference or cite my work as 'paper1-3', respectively to the list above. The papers can be found after the list of references or at the end of the thesis. For curious readers, the poems at the beginning of each chapter are Icelandic lullabies to harmonise all the content about sleep.

Acknowledgments

To my supervisor Sune. First and foremost, thanks for seeing potential in me and giving me this amazing opportunity. The passed couple of years have been a lesson of a lifetime, and I will carry everything you have taught me into future endeavours — not only for professional purposes but also just for life in general. Moreover, having gone around DTU and other places the last three years, I have encountered smiles, twinkled eyes and heard all the nice things people have to say about you. I know from first hand that you are deeply admired and appreciated, not only for your brilliant work but also because you are such a nice person.

To my collaborator Jim. Thanks for a superb collaboration, your questions were always so fresh and unexpected, and really pulled our work to the next level. You also taught me how to create a proper outline for a paper, and it is one of the most useful things I have learned over the passed three years!

To my collaborator Jari. Thanks for hosting me at Aalto University. Even though I only managed to stay a few weeks, I was warmly welcomed by you and your research group. I have also really enjoyed our collaboration on trying to understand the complex patterns of sleep.

To my collaborator Kelton. My favourite landpirate. You took me under your wing during my masters thesis and were the first person I got to know who was really into science. Your enthusiasm was contagious! Luckily, I also became PhD student and today I am so proud of the work we have accomplished together! Thanks for all the amazing inspiring discussion we had about science, food, beers, politics, pastries, and other silly things. I will hold the crazy train adventures to Lund close to my heart #teamÖresund.

To Laura. Planet Laura. My favourite fast-drinking beer buddy. I am so thankful you showed up just when I was couple months into my PhD. We have shared so many amazing moments; Enys wedding in Budapest, being stuck in Lund, beers at Gorki's and more beers at Gorki's. Not only do I cherish our friendship, but your scientific work is really inspiring and sets a standard for the rest of us. I know I can always come to you for advice, and you somehow always understand my vague and unformulated questions and thoughts. And, of course, million thanks to Marco De Nadai as well!

To my collaborators at DTU & KU. Anna, Kristoffer, Germans, Peter, Asger, Andreas, Sebastiano, Bjarke, Ulf, Radu, Enys. Thanks for all the burger lunches, group meetings, silly discussion and good times. Your talent inspires and motivates me everyday! It really has been an amazing experience to be surrounded by such intelligent and wonderful people in your work environment.

To mom, dad, Helga, Ari & Geir Thanks for everything. You have always supported my decisions and choices, and never doubted me – not once. Your relentless belief and pride in me has carried further than you can imagine. Thanks for the countless visits to Denmark and running circles around me when I came to visit Iceland. You are my entire world, and like you say in Icelandic; botnlaus ást og umhyggju <3

To Halla, Sigrún and Ási. Klettarnir í hafinu. My favourite people and the fixed points in my existence. Thank you so much for the love and support. You made feel so special when I was growing up, and you still do today.

To Edda. You are the one that experienced me at the lowest low and highest high throughout these three years. I think I would not have made through without the fixed point of you in my surroundings everyday. Thanks for listing to all the 'ranting' and being just as excited as me about all small victories. I cherish our friendship so unimaginably.

To Anna Hildigunnur. Nobody, and I mean nobody had to listen to me complain as much the passed three months as you. And each time I complained, you were so understanding, loving and supportive. I cherish you for everything you are and so grateful for our friendship.

To all my friends and family. Thank you for all the support! Every one of you owns a piece of me, thus also a piece of this work.

Sofðu unga ástin mín

Úti regnið grætur

Mamma geymir gullin þín

Gamla leggi og völuskrín

Við skulum ekki vaka um dimmar nætur

Jóhann Sigurjónsson

1

The Physiological, Biological, Behavioral & Contextual Aspects of Sleep

Sleep is a natural daily recurring state that has developed as humans have evolved throughout time. Our living conditions have changed drastically the last 200-300 years and currently our environment is governed by artificial light, screen use, smartphone notifications, caffeine consumption, abundance of information at all time and myriads of other things. Due to how fast our daily life and environments have changed, the physiology process of sleep has unlikely caught up with our new ways of living. While these changes have taken place the passed couple of decades, studies have shown that chronic sleep deprivation and sleep disorders are increasingly more common [1]. Here I provide high-level overview of the physiological architecture of sleep and introduce biologically based differences that can effect sleep. Furthermore, I provide evidence to exemplify how behavioral and contextual factors in human everyday life can effect sleep.

1.1 The Physiological process of sleep

Sleep is a complex physiological process involving changes in brain activity, neurotransmitters, parasympathetic nervous system, muscle tone and multiple other functionalities. Nighttime sleep is typically broken down into different stages that are characterised by certain neuro-physiological activities. These sleep stages are measured with the gold standard of sleep recording, *the Polysomnygraphy* (PSG) [2,3,4].

Sleep stages The changes that occur in the brain during sleep are grouped together into two stages; non-rapid eye movement (NREM) and rapid eye movement (REM) which alternate cyclically throughout the night [5]. Non-rapid eye movement is either split into three or four sub-stages, dependent on classification standards [6,7]. *NREM1* is characterised by alpha waves (8-12 Hz), emergence of theta waves (4-7 Hz) and slow-rolling eye movements. *NREM2* is marked by the presence of transient electrical phenomena; k-complexes (large-amplitude rapidly fluctuating burst of brain activity) and sleep spindles (12-15 Hz oscillating signals lasting 0.5-2 seconds). The purpose of these have not been fully established, but they are believed to support memory consolidation and filter sensory input. *NREM3* is often referred to as slow-wave sleep (SWS) because of high amplitudes and low frequency (1 Hz) in brain activity. As for the other sleep stages, the purpose of SWS is not fully understood but it is believed to discharge sleep pressure accumulated throughout waking

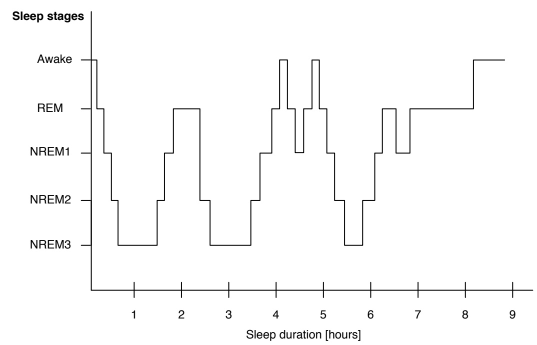


Figure 1.1: Example of a hypnogram. Hypnogram is a form of visualisation to depict how an individual cycles through sleep stages (marked on the y-axis) during nighttime sleep (x-axis represent sleep duration).

hours of the day [8,9,5]. The end of SWS is followed by *REM* sleep which is characterised by brain activity similar to when an individual is awake, all the while there is loss of muscle tone (except for eye muscles). This stage of sleep is sometimes referred to as the "paradoxical sleep", since people seem awake judging from brain activity but are obviously sound asleep if observed in reality [5]. Slow-wave sleep (NREM3) has been considered the most restorative sleep stage and often associated with sleep quality, however Matthew Walker makes an important point in his book "Why we sleep?" that all sleep stages must serve an important role since we cycle through them repeatedly throughout the night [10,11,12]. Figure 1.1 exemplifies a hypnogram, which visualises the alternation between sleep stages during a night. Figure 1.1 illustrates one example of a hypnogram, but they vary by individuals, nights and sleep duration, but the overall structure is fundamentally the same.

The Two Process Model A phenomena called the *two process model* ensures that sleep takes place every night in healthy humans [13,14]. The effects at play stem from *the process of homeostasis* and *the circadian system* [15,16]. Humans accumulate adenosine from last awakening, which is believed to drive sleep propensity [8,9]. The circadian system optimizes bodily functions for wakefulness and sleep for certain intervals of the approximate 24-hours day. In fact, more than 100 variables (e.g.

body temperature, melantonin, serotonin, dopamine and more), both physiological and psychological, have been shown to change their values rhythmically over the course of the day[17]. All of these daily changes contribute to wakefulness, which is close to reaching a minimum at night while sleep propensity is measured the highest – this state is referred to as *the sleep gate* and where sleep onset is likely imposed and nighttime sleep begins[17,5]. The circadian system will be reviewed more intricately in Chapter 2.

1.2 Biologically based differences

Few human biological traits are known to impact sleep – the most prominent ones are *age* and *gender*. Additionally, adverse changes in body-mass index (BMI) have been associated with modification in sleep which will be addressed below.

Age Age-associated changes include decrease in total time asleep, slow wave and REM sleep, as well as increased sleep latency, sleep fragmentation, NREM1 and NREM2 sleep[18,19,20]. It is undetermined whether changes in sleep occur because older adults need less sleep or they are unable maintain or produce as much sleep. A study revealed that older adults were more resistant to deprivation of slow-wave sleep than younger adults[21]. However, research has shown that reduced sleep negatively impacts cognitive performance, irrespective of age, and a recent review confirmed that there is more evidence that supports the hypothesis that older adults have an impaired ability to generate sleep rather than a reduced sleep need[21,22,23]. Age-related changes are not only characterized by changes in sleep-brain activity and sleep duration, but also by alteration in the circadian system which include phase advancement and diminished amplitude in daily rhythmicity of variables that have been related to sleep (e.g. core temperature, melatonin and cortisol)[24,25]. These phase advancements are also detected in sleep timing, where bed and wake times, as well as mid-sleep, occur at earlier hours with increasing age[26,27,15,28,29,30].

Sleep variability is an important qualitative indicator and often used in sleep research, since a growing number of research indicates that irregular sleep is associated with adverse effects on human health [31,32,33,34,35]. In sleep epidemiology inter-individual (between-individual) variability is common to use, due the nature of epidemiological data (single or few data-points per individual), but a more informative measure is intra-individual (within-individual) variability. Between-individual variability for sleep onset, offset, duration and mid-sleep declines with age and a sleep diary-based study found within-individual variability in sleep duration to decrease with age [36,31,37]. However, limited evidence exists about age-related changes in sleep variability within individuals, particularly for separate estimates within weekdays and weekends [37,38,31].

Gender Age-related changes in sleep are moderated by gender, and women peak in phase delay earlier than men [39,15]. Women are found to have a significantly higher melatonin amplitude and lower core body temperature, and a shorter circadian period compared to men [40,41]. These evidence rationalise the fact that women on average are earlier chronotypes than men (reported via the Munich Chronotype Questionnaire) and more likelier to be 'morning types' (reported via Morningness–Eveningness questionnaire) [15,42,43]. The overall difference between the genders in phase preferences decreases as people move further into adulthood, and disappear around 50-60 years of age which is the period during the lifespan where menopause usually takes place for women. These differences between the genders in daily rhythmic behavior goes along with the fact that women sleep on average more than men up until age 50-60 [44,45,46,47].

Body-mass Index Another biological trait associated with sleep behavior is body mass index (BMI), but BMI is typically classified into four group, underweight/normal weight/overweight/obese, designed by the World Health Organization [48]. The prevalence of obesity (BMI$\geq$30) has increased the passed decades while average sleep duration has decreased [49,50,51,52,53,54,55,56,57,58]. There is evidence that the co-occurrence of these two might be related, where both epidemiological evidence and clinical interventions studies tie the two together[59,60,61,62,63]. A clinical study with approximately thousand participants revealed that individuals with short sleep had reduced leptin and elevated ghrelin measures, which is a state likely to increase appetite and possibly explaining the increased BMI observed with short sleep duration[64]. A systematic review with about 20 studies (both clinical and epidemiological) concluded that short sleep duration might risk obesity due to adverse effect on parameters of glucose regulation, including insulin resistance and certain disregulation of appetite leading to excessive food intake[38,65]. Furthermore, a study with about 1200 twins found short sleep duration associated with elevated BMI (adjusting for genetics and shared environment)[66]. However, the causality between short sleep duration and obesity has not been fully established, since majority of clinical studies take place over too short of a period and lack standardized methodological approach[67,68].

1.3 Environmental and Contextual effect on sleep

Light exposure Humans have adapted to natural environment of sunlight during the day and darkness during night, which has created the daily rhythm of rest and activity. In fact, light is considered the strongest zeitgebar on the circadian system[69,70,71,72]. Light is absorbed through the eyes and signals are forwarded to the suprachiasmitic nucelus (SCN) to synchronise and coordinates all cellular circadian clocks in the body[73,74]. Currently, we are not only exposed to natural light, but our surroundings in everyday life are illuminated with artificial light and blue-light from computer screens and smart phones, all of which has been shown to suppress melantonin onset[75,76,77,78]. The extent and magnitude to which this effects human sleep is not fully understood. More longitudinal studies are required to further confirm melantonin suppression due to artificial light exposure, and a deeper understanding of influences from different illumination levels and duration of exposure is needed[79].

Temperature Another environmental variable that fluctuates over the course of a day, month and year is outdoor temperature. Core body temperature is an important variable in the physiological process of sleep, but it drops before sleep onset is imposed at night[80]. Indoor, outdoor and intrinsic core temperature are interconnected and therefore likely to influence the daily rhythmicity of sleep. Cold temperature have been shown to increase frequency and duration of nighttime awakenings, and the length of REM sleep periods, while hot temperature is associated with later sleep onset, shorter duration, day-time napping and self-reported sleep deficiencies[81,82,83].

Seasonality and Geographic Location The two key variables that change the most with seasonal fluctuations are daylight and temperature. As discussed above, light exposure and outdoor temperature influence the daily rhythm of rest and activity, and thus sleep patterns tend to vary across the year[84,85,86]. Seasonal variation are more marked in extreme latitudes compared to equatorial regions, and tends to exacerbate seasonal variation in sleep behavior[87,88].

Cultural Context There are large disparities in sleep patterns across cultures, with the most prominent contrast between Eastern (Asia) and Western (Europe and North America) geographic regions. Studies have shown that sleep duration is shorter among individuals residing in the East compared to those living in the West[30,86,46,89]. People in the East tend to go later to bed than those living in the West, while both groups are waking up at the same time[30,86]. Weekend-weekday misalignment (measured with social jetlag via Munich Chronotype questionnaire) has been reported lower in large samples of individuals residing in China and Japan, compared to a European one[90,91,44]. Weekend-weekday differences in sleep patterns are also observed minuscule for Eastern countries in a large scale ($\sim$ 24 0000 users) study with objective measures from wearable devices[46]. An unexplored factor in global sleep differences is day-time napping which is more accepted and common in some parts of the world compared to others. For example in Japan, the phenomena of *Inemuri* (napping; literally 'to be asleep while present') is interpreted as the result of exhaustion from devotion to work and and is considered a subtle method of showing commitment to work[92]. 'Siesta' or day-time napping are also a culturally accepted phenomenon in the Mediterranean region, especially during the summer months to escape the grueling heat. However, whether these day-time rest periods play into global differences in the daily rhythm of rest and activity is yet to be explored in the literature.

Socioeconomic status A recent literature review concluded that low socioeconomic status (SES) is associated with higher rates of sleep disturbances (difficulty falling asleep and maintaining sleep) and lower sleep quality[93]. A study concluded that individuals with lower education were more likely to experience insomnia (also when controlling for ethnicity, gender, and age) and another showed that individuals raised in households with lower SES spend more time in NREM2 sleep and less time in slow-wave sleep (SWS) than those with higher childhood SES[94,95]. Some racial groups are more probable to hold a low socioeconomic status, and in fact, African Americans (in the US) and other racial minorities are likely to have short and long average sleep duration which are associated with increased mortality[96]. Socioeconomic status is considered a fundamental driver in differences considering population health. Marginalized groups are routinely exposed to stressful situations such as discrimination and job strain, which require coping mechanism and extra energy that will impact on sleep in the long run[97].

1.4 How can behavior influence sleep?

Physical Exercise Epidemiological studies have consistently associated physical exercise with better sleep, observed across multiple age groups, gender, race and demographics[98]. Higher than average levels of physical activity are associated with less likelihood of insufficient sleep and fewer sleep disturbance[99,100]. A meta-analysis with 41 studies summarized the effect of acute or regular physical activity on a self-reported or biological measure of sleep, and provided compelling evidence supporting exercise as an intervention to improve perceived and objective metrics of sleep in healthy individuals[101]. In summary, physical exercise does not only provide acute improvement of sleep quality, but is also inexpensive and accessible to most people, and often a cheaper option than other standard sleep medicine treatments[79].

Dietary consumption Above I discussed that sleep deprivation might risk weight gain and obesity due to its adverse effects on metabolic regulators. Furthermore, the relationship between food-intake and sleep is likely bi-directional, meaning the composition of ones diet can influence sleep quantity and quality. For example, epidemiological studies have associated lower sleep quality with high carbohydrate food intake, while good sleep quality is associated with more vegetable and fish consumption[102,103,104]. Not only does the composition food intake influence sleep but also the timing of the consumption where late caloric intake has been associated with increased risk of obesity[105,106,107].

Probably the most notorious fluid that people associate with tiredness and sleep is coffee. Adenosine is present throughout the central nervous system (CNS) and believed to mediate the effect wakefulness[108,109]. Caffeine causes most of its effects by antagonizing adenosine receptors and consequently relieving the feeling of tiredness. Caffeine consumption is believed to have negative effect on sleep and large sample studies find association between daily caffeine intake and sleep problems, as well as daytime sleepiness[110,111,112]. Generally, good sleep hygiene rules suggest to forgo caffeine consumption after midday, but studies have shown that caffeine consumption 6h before sleep can negatively effect sleep[113].

Smartphones & Sleep Technology continuous to play large role in the human daily life and much of it actually takes place online (especially during the COVID-19 pandemic which is ongoing during the writing of this dissertation). Smartphones are by our side nearly all hours of every day, vibrating and lighting up with notification which likely distract us from what we are physically engaged in. The effect that smartphones have on human sleep is a relatively untouched territory in terms of research. However, there do exist some interesting studies which have associated screen-time and smartphone addiction with poor sleep and sleep quality [114,115]. Late-night use of smartphone has shown to interrupt sleep, and makes people more tired and less ready to work the morning after [116]. A recent review concluded that the use of digital and social media can have positive effect on children and adolescents (e.g. early learning, exposure to new ideas and knowledge, increased opportunities for social contact and support) but also adverse impact in terms of health, *sleep*, attention and confidence [117].

When studying the effect of smartphones on sleep there are many aspects to consider. There exist multiple types of apps (e.g. social media, dating, utility, media, games and more) which might impact us differently. For example the use of utility app such as Google Maps is improbable to leave a mark on us 'emotionally' while social media or dating apps might. Another important aspect is the time of the day the usage takes place – but usage closer to normal bed-time or during the night might have more impact on sleep quantity and quality. There is also a potential bi-directional relationship between sleep quality and smartphone use, where a poor night's sleep could result in more day-time fatigue and lack of attention leading to more screen use. The effect of light exposure from the screens of electronic devices before bed have actually studied closely, and found to prolong the time it takes to fall asleep, induce delays for circadian clock and REM sleep, and reduce alertness the following morning [77,118,78].

Other How humans choose to move through life can impact their health, and decisions ranging from small daily choices to life altering changes can have lasting impact on sleep. For example, having a child is likely to impact individual's sleep quality. Pregnancy and the postpartum period are associated with physiological changes (for women) and behavioral demands (both parents), known to disturb sleep quality (more for women than men) [119,120,121]. Other long lasting changes such as moving to a new apartment could impact the daily rhythm of rest, since environmental factors such as noise (e.g. excessively loud neighbours, airplanes and car traffic) can have a negative effect on sleep health [122,123,124]. A small innocent decisions such as taking a couple days vacation, can also impact sleep adversely. Travel and new resting environments are known to influence sleep quantity and quality, and the phenomena of the *First night effect* (FNE) was documented in 1964. FNE is characterised with sleep-initiation problems and prolonged sleep-onset latency, found to take place on the first night of sleep in new environment [125,126]. This is actually a consequence of a single brain-hemisphere displaying elevated alertness in new and unfamiliar surroundings [127]. Furthermore, the journey to the destination can also induce sleep complication due to *travel fatigue* [128,129], that can be exacerbated with jetlag, which occurs due to desynchronisation of the body's internal clock and the new time zone an individual enters after long-distance travel [130,131,132].

Conclusion

Sleep is a complex physiological process, involving changes in different parts of the body which is though highly influenced by behavioral and contextual aspects in everyday life. The physiological changes involve alteration in brain activity that are grouped into four different stages, which humans cycle through repeatedly over the course of a night. These sleep stages are believed to be restorative and serve important functionality for human health. Two independent processes called sleep homeostasis and the circadian system, control the daily rhythm of rest and activity where different bio-markers vary over the course of $\sim$ 24 hour day to optimise the body for sleep and to impose the feeling of tiredness at night and ensure that sleep takes place. Gender and age exert strong influences on sleep, which is observed in multiple large scale epidemiological studies. Sleep timing or the phase preference is advanced to earlier hours with increasing age, and gender differences are life-stage dependent. The environment and context of human existence also influences sleep behavior, were for example natural light, seasonality, geographic location and cultural context can have strong effects on sleep behavior. Furthermore, physical activity, use of blue-light emitting screens, dietary choices, timing of consumption, travel and multiple other behavioral choices can also influence the human daily rhythm of rest and activity.

Bí, bí og blaka,

álftirnar kvaka,

ég læt sem ég sofi,

en samt mun ég vaka.

Bíum bíum bamba,

börnin litlu ramba

fram um fjalla kamba

að leita sér lamba.

Þjóðvísa (Icelandic Folk Song)

2

Methods & Metrics from Sleep Epidemiology

Every human on earth sleeps, and if they do not it will impact them. Short and irregular sleep duration contributes to molecular, immune, and neurological changes that play a role in disease development, increasing, for example, the risk of obesity and cardiovascular diseases, and substantially affecting mood, motor and cognitive performance[133,134,135,136,59,137,138,139]. For this reason it has been considered important to study sleep across the population where the field of Sleep Epidemiology dates back to around 1980 with the first documented modern epidemiological studies being conducted[140,141]. Sleep epidemiology is defined as *the study of distribution and determinants of sleep, sleep-related symptoms and sleep disorders, and the application of this study to improve sleep health and sleep-health related conditions, including studies of how sleep influences health and disease*[142,143]. This chapter reviews important measures and methods from the field of sleep epidemiology, mostly focusing on two aspects; data collection and sleep metrics.

2.1 Large Scale Data Collection

Large-scale studies have been based on self-reported sleep estimates where key metrics were limited to quantities people could reasonably be expected to recall. There are three main methods that have been used to collect data about people's sleep behaviour at scale; questionnaires, surveys and diaries. All of these methods are classified as *self-report*, meaning that subjects estimate quantities and qualities about their behavior themselves. Wearable devices or wrist actigraphies are becoming more common to use in epidemiological studies, however they are the main subject in Chapter 4[144,145,146]. Although it seems simple to ask an individual about their sleep, it is actually problematic since humans have never been considered very good at assessing their own behaviour. Multiple studies have estimated the extent to which self-reports of sleep duration (via sleep diaries) reflect on objectively measured estimates with wearable devices, and found them to correlate poorly[147,148,149,150,151]. Sleep surveys contain straightforward questions such as "How many hours of sleep do you usually get a night (or when you usually sleep)?"[152]. The quality of the assessment obtained from surveys has not found corresponded well with objective measures of sleep assessed using actigraphy, as well as corresponding poorly with estimates from self-reported sleep diaries.[153,154,155,156]. Sleep questionnaires pose a range of questions where the answers are used to obtain a score or estimates for some particular aspects of the subject's sleep behavior. Examples of some popular ones are i) The Pittsburgh Sleep Quality Index (PSQI), a self-rated questionnaire which assesses sleep quality and disturbances over a 1-month interval and ii) Morningness-Eveningness questionnaire (MEQ) which aims to measure whether a person's circadian rhythm produces peak alertness in the morning, in the evening, or in between [157,158]. These are widely used in a range of different studies and have been validated thoroughly [159,160,161,162,163]. The PSQI and MEQ ask subjective questions, therefore impossible to use them for objective measured sleep. However, another widely used questionnaire is the Munich Chronotype Questionnaire (MCQT) which asses the timing of sleep within the 24 hour day, and was actually developed to capture the same characteristics as MEQ [16,164]. The framework of MCQT enables the use of objectively measured sleep, and renders assessment of the biological phase preference (chronotype) and misalignment between the social and biological clock (social jetlag). These metrics, and others, will discussed more closely in the next section.

2.2 Phase preference and Misalignment

Professor Kleitman proposed the existence of the daily rhythm of rest and activity, or the circadian rhythm, in his book, Sleep & Wakefulness, first written in 1939 [165]. Daily cycles have in fact been observed for organisms ranging from unicellular marine creatures to mammals [15]. The circadian system - referred to as chronotype - is found within each cell of the human body, where signalling proteins produce approximately a 24 hour day where certain functionalities are optimized for the *biological day-time* while shifting neurobiological activities to favor sleep when it is *biological night*. A person's chronotype is believed to depend on specific alleles of genes but is also a manifestation of gene-environment interaction [166,167,168,169,170,171]. The shift between day and nighttime produces environmental signals (e.g. light, temperature and access to resources) which act as zeitgebars to synchronise the intrinsic biological clock. Nevertheless the circadian rhythm actually persists in the absence of these environmental cues and is self-sustained - but how can that be? [15,172,71]. An individual's phase of preference is believed to be *entrained* by environmental cues, not fully controlled by them and chronotype rather varies by their strength. There are multiple examples of this; i) when individuals exchange urban lives (weak signals due to indoor life and limited natural light exposure) for natural light conditions, their sleep timing and dim-light melantonin onset (DLMO) advance significantly, ii) sleep timing is earlier in populations with no access to electricity compared to those with access to artificial light, iii)chronotype is earlier in rural areas than in urban ones and iv) average chronotype correlates with the position within a time zone and the further to the East the earlier the chronotype [173,174,175,176,177,178]. The assessment of chronotype with the MCQT has been validated against bio-markers, such as dim-light melantonin onset and cortisol, as well as objective behavioral measures of sleep, and is believed to be the best estimate of phase preference via self-reports [179,28,180,181,182,171,15,183].

The Munich Chronotype Questionnaire poses 17 questions about bed-time, wake-up time, sleep latency and more, where the answers are used to estimate mid-sleep on work-free days (MSF), which is used to assess chronotype. Work-free days are believed to better reflect the circadian phase preference since there is probably less pressure of social or work obligations [42]. Early morning work schedule and alarms do truncate nighttime sleep on weekdays for many individuals, resulting in shorter sleep duration than preferred [184,44]. This recurrent temporal pattern of sleep deprivation, sometimes referred to as sleep debt, is often compensated for with longer sleep duration on the weekends rather than earlier bed-times on weekdays [42]. To make sure that the chronotype is not influenced by sleep debt, it is corrected for in the calculations:

$$MSF_{sc} = \begin{cases} SO_{free\,days} + \frac{SD_{free\,days}}{2} & \text{if } SD_{free\,days} \leq SD_{work\,days} \\ MSF - \frac{SD_{free\,days} - SD_{week}}{2} = SO_{free\,days} + \frac{SD_{week}}{2} & \text{if } SD_{free\,days} > SD_{work\,days} \end{cases} \tag{2.1}$$

Where MSF is midsleep on free days while MSF_{sc} is corrected midsleep on free days. SD refers sleep duration where the subscript 'week' denotes weekly average sleep duration, 'free days' weekend averages and 'work days' weekday averages. SO refers to sleep onset (point in time where people fall asleep).

Chronotype is a continuous variable (with a fixed range), measured in time (hh:mm), and distributes normally across the population [15,42]. Individuals who are late chronotypes (sometimes referred to as owls) typically sleep less on weekdays and longer on weekends, while the opposite applies to the earliest risers [184,15].

Two temporal dimension have been discussed which govern the daily rhythm of rest and activity; the biological (innate preferences) and environmental clock. However, there is a third dimension that might be the most influential one on human sleep - *the social clock*. A large part of the population, or about 75 % of the US and European labor force, maintain a conventional 5 day work week from 9 to 5 which constrains their weekly sleep patterns [185,186]. Due the fact that chronotype is a normally distributed trait with a wide range of behavior, many of these individual are contingent to day schedule that is out of sync with their phase preference. Wittman et. al. (2006) developed a concept to describe this misalignment between the biological and social clock called *social jetlag*, and is estimated by calculating the difference between midsleep on free days and work days [184].

$$\textit{Social jetlag} = MSF - MSW \tag{2.2}$$

where MSF denotes midsleep on free days (weekends) and MSW midsleep on work days (weekdays). Groups with high social jetlag, or living against the clock, have been associated with negative behavioral outcomes such smoking, obesity, less healthy dietary patterns, worse academic performance, symptoms of depression and more [184,187,183,188,44,189,190,191,192,193,194]. Sleep researchers have for years advocated for more flexible school and work schedules to support those who have a biological clock far out of sync with the social clock, with little results.

2.3 Insufficient sleep and poor sleep quality

An important measure used to quantify sleep behavior at scale is sleep duration. As mentioned in the section above (*Large Scale Data Collection*), there do exist different methods to collect data about sleep duration at scale - most commonly subjective and retrospective estimates via sleep diaries and surveys. These estimates are acknowledged to recall biases, and habitual self-reported sleep duration should rather be thought of as time in bed than actual physiological sleep [195]. The term *insufficient sleep* is thought of as nighttime sleep too brief to meet physiological needs. Scientist have debated about what is sufficient sleep duration for the passed years, but today there is a general consensus that recommended amount of sleep for adults is 7-9 hours per night[79]. Sleeping less than 7 hours per night on a regular basis is associated with adverse health outcomes and sleeping more than 9 hours per night on a regular basis is associated with health risks [196,197,198,199]. The only way to study the prevalence of insufficient sleep across the population is by comparing self-reported estimates of habitual sleep duration to guidelines concerning recommended amount of sleep for adults. There do exists ways to explore *sleep quality*, for example measures of *sleep fragmentation*; number of awakenings after sleep onset and/or wake after sleep onset (WASO). Sleep fragmentation can have detrimental effect, and they are known to reduce subjective assessment of mood, decreases mental flexibility and sustained attention [200,201]. *Sleep latency* is also considered a measure for sleep quality, and used to examine how long it takes an individual to fall asleep, after he or she gets into bed. *Sleep efficiency* is another importance parameter and refers to the percentage of total time in bed actually spent asleep. In a way sleep efficiency summarizes the severity of sleep fragmentation and sleep latency simultaneously [202]. Most of these is metrics are difficult to collect via self-reports, but perhaps the most accessible one is the number of awakenings per night.

2.4 Sleep variability

Evidence suggest that excessive time in bed and irregular sleep–wake timing, may contribute to the development of insomnia[203,204]. For example, if an individual experiences a poor nighttime sleep, he or she may try to compensate by staying in bed longer the morning after or go earlier to bed the consecutive night. However, this compensatory behavior may have negative consequences where homeostatic pressure might be relieved and could reduce the ability to sleep[205,206]. Sleep variability is a measure to quantify sleep behavior in epidemiology and has been associated with adverse health outcomes such as cardio-metabolic risks, poorer sleep quality and lower plasma levels of inflammatory markers in older adults, lower academic performance and lasting impact on brain development in adolescents[207,208,209,210]. A meta-analysis of 53 peer-reviewed empirical publications has also associated sleep variability with age (more variability among younger people), race/ethnicity (more variability for non-whites), physical health conditions, body-weight (BMI & body-weight gain associated with higher variability), psychopathology (symptoms of depression and bipolar), insomnia and stress[31].

Sleep variability is typically quantified as the standard deviation of multi-night recordings within individual (referred to as intra-individual variability)[31,38]. The variability can be estimated for different measured, e.g. sleep duration, chronotype and more. Due to the nature of data-sets in sleep epidemiology, they rarely contains numerous data-points per individual, therefore has inter-individual variability (between subject variability) also become common to use[86,36]. Nevertheless, new tools and techniques are constantly being developed, and a recent example is sleep regularity index (SRI), constructed by Philips et al. (2017) - a measure which captures changes in sleep timing on a day-to-day (circadian) timescale[209].

2.5 Bed & Wake Time

Bed time (sleep onset) and wake time (sleep offset) represent the point of time when an individual falls asleep and wakes up. These are *not* commonly used to study sleep timing across the population and most large scale sleep studies rather use mid-sleep or chronotype [86,36]. In fact, midsleep or chronotype have been shown to correlate well with multiple bio-markers that are indicators of sleep – thus midsleep is believed to be the best proxy to measure the biological phase of preference [179,28,180,181,182,171,15,183]. However, in modern society biological preferences might be suppressed rendering a mismatch between innate inclinations for sleep timing and need, and the actual outcome. Bed and wake time therefore entail important information to help understand behavioral choices concerning sleep, which in some cases is stronger determinant for the human daily rhythm of rest.

Conclusion

In this chapter I review some of the important metrics and methods used in sleep epidemiology to study daily rhythm of rest across the population. There exist limited ways to collect sleep data at large scale, and most studies rely on self-reported estimates collected via surveys, questionnaires and diaries – all known to recall biases. Consequently, the quantitative metrics used for analysis in sleep epidemiology have been constrained by these data types. The most common measures are retrospective and habitual sleep duration, and questionnaire based estimates for phase preference (chronotype) and misalignment (social jetlag). Sleep variability is an important element for good sleep hygiene since irregular sleep behavior has been linked to adverse physiological effects. It is typically quantified as standard deviation of within individual recordings, but due to the nature of the data in sleep epidemiology (typically constrained to single or few measures per user), between-individual variability is also often employed. Interestingly, bed and wake time have mostly been excluded when studying sleep across the population, but I believe they entail important information regarding behavioral aspects of sleep.

Dvel ég í draumahöll

og dagana lofa.

Litlar mýs um löndin öll

liggja nú og sofa.

Sígur ró á djúp og dal,

dýr til hvílu ganga.

Einnig sofna skolli skal

með skottið undir vanga.

Kristján frá Djúpalæk

3

Hierarchical Data Analysis

Hierarchical data typically refers to a data-set containing time series observations from number of individuals or groups, thus observations involve at least two dimensions; one to identify the individual (or group) and a time series dimension[211,212]. This type of data can have a complicated structure with hierarchical group levels, crossed design and unbalanced sampling rate. Hierarchical data is also sometimes referred to as panel data, longitudinal data or time series cross-sectional data[213,212]. The chapter splits into two main parts; first part discusses what has to be conducting analysis with raw hierarchical data and the second part reviews an important regression analysis method used to model panel data.

3.1 Analysis of Raw Data and Error propagation

The data-set for this project is hierarchical with a set of behaviours measured across thousands of individuals where multiple effects are at play. Naturally I explore and analyze the data with individual and population level visualisations and statistics. When doing so, it is important to be considerate of what is aggregated and averaged at all time, and a key point is that behind every average or any summary statistic is an underlying distribution with a range of behavior. Thus, even though I observe statistically significant differences between measures for different groups, the underlying distributions might overlap extensively. Important element to all of this is how to propagate uncertainty and what is the origin of uncertainty? Since users have multiple nights recorded for a set of behaviors, the measures span a distribution and the average is accompanied by uncertainty. In most cases, I choose to propagate the uncertainty from individual level. The alternative is to estimate the uncertainty for the distribution of individual summary statistics, see schematic explanation on Figure 3.1. I quantify uncertainty of an average (μ) with the standard error of the mean (SEM), calculated as $\partial\mu = \frac{\sigma}{\sqrt{n}}$ where σ is the standard deviation of the distribution and n is the number of data-points. The standard error of the mean (SEM) can both be estimated for distribution at the population level (Figure 3.1B), and individual level (Figure 3.1A). If I choose to estimate SEM at the individual level, it must be propagated in the following manner[214]:

$$\mu_{group} = \frac{\mu_1 + \mu_2 + ... + \mu_m}{M} \; where \; \mu_i \pm \partial\mu_i \; then \; \partial\mu_{group} = \frac{\sqrt{\partial\mu_1^2 + ... + \partial\mu_m^2}}{M} \tag{3.1}$$

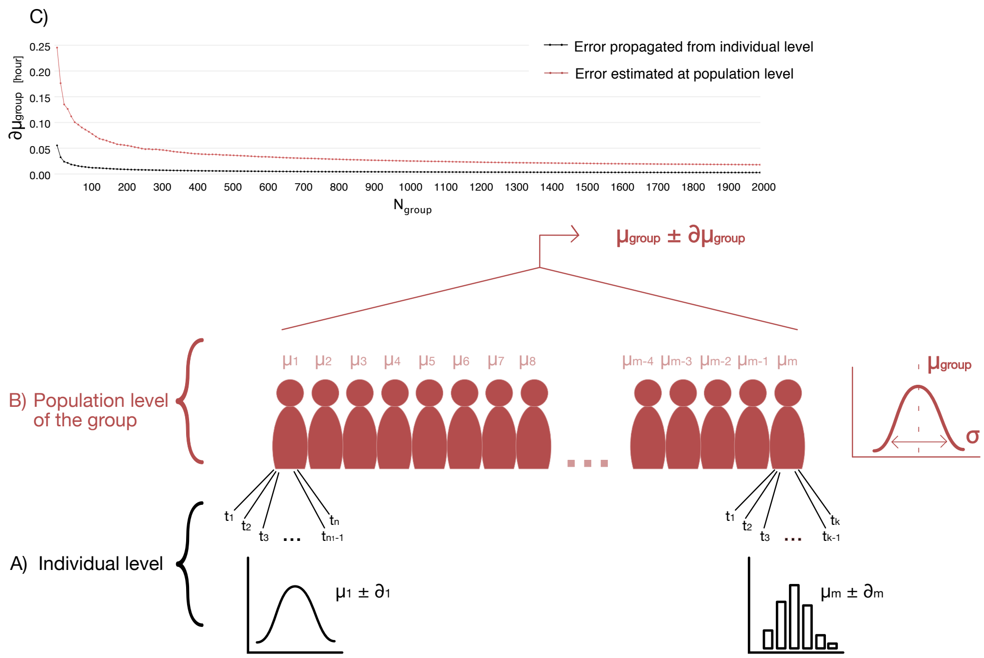


Figure 3.1: Error propagation A schematic illustration to show that for hierarchical data-structure, error or uncertainty group averages can be quantified in two ways; **B** from the distribution at the population level or **A** by propagating uncertainty from the individual level. **C** compares the two possibilities (A & B) quantitatively by randomly selecting a group of users with varying size ($N_{step} = 10$) and estimating the uncertainty ($\partial\mu_{group}$) for the group average (μ_{group}) at every step. The red curve represents error estimates at the population level and black curve the error estimates when SEM is propagated from the individual level.

where μ_i are averages for individual's distributions and $\partial\mu_i$ the respective SEM. On Figure 3.1 I explore the quantitative difference between the two approaches, by varying a selection of users (who all reside in the same country) and estimate the uncertainty for group average of sleep duration. Error estimate at the population level are always slightly higher than those propagated from the individual level. The difference are substantial for small user groups but when $N_{group} \geq 500$ they become minuscule. One should keep in mind that this comparison might vary by selection of users and metrics.

3.2 Generalized Linear models

Statistical tools such as analysis of variance and regression analysis are based on strict assumptions concerning i) normality (each sample is drawn from a normally distributed population), ii) variance homogeneity (samples are drawn from populations with equal variance) and iii) independence (samples are independent within and between groups)[215]. Here I introduce Generalized Linear Models which provide similar framework as typical regression models but with relaxed assumptions. The development of Generalized Linear Models dates bake to the 1930's but studies from mid 70's introduced the modern framework[216,217,218]. They can be viewed as population-average models in which the parameters are interpreted as quantifying effects of covariates on the marginal mean value of the dependent variable for the entire population. Generalized Linear models consist of two main types; *Fixed effects* and *Mixed effects*. For the purpose of this study I only use Mixed effects model and therefore the main subject of this chapter. However, I do touch upon how the two differentiate and justify the choice of Mixed Effects Model.

3.2.1 Data Structure for Generalized Linear Models

The data typically analysed with generalized linear models is panel-data and consists of groups of cross-sectional units observed over time (as mentioned before). The data can be multilevel and typically correlated, where there is a source of dependence within a level[211,216]. A classical example is a data-set with patient records. Patients can have multiple doctors and be treated at different hospitals. Observations from the same doctor and/or hospital are likely dependent, where the outcome of treatments can rely on experience and dedication of a doctor and the resources of a hospital.

If I define the case more precisely and consider twenty doctors (independent of one another) where each has some sample of patient records. The aim is to examine linear dependence between two measured variables, where ordinary least square (OLS) linear regression can not be applied since there is dependence between data-points sampled from the same doctor. One way to deal with the correlations would be to aggregate samples into averages by doctor, and then perform an OLS linear regression. However, then there is not taken advantage of the full data-set and aggregation can result in scenario of a Simpson paradox, or a case where trend appears in several different subgroups of the data but disappears or reverses when these groups are combined, see example on Figure 3.2[219]. Another approach would be to fit twenty different models (one for each doctor). However, the results are then "scattered" and one needs to boil them together to obtain an overall conclusion. Mixed effects model can be thought off as a trade-off between these two approaches, where their framework controls for the dependence in the data-set and allows for the hierarchical structure, while yielding an overall assessment of the linear relations and effects of covariates.

The data-set for this project has a hierarchical structure and is comprised of measures for nighttime sleep recorded over different periods of time for thousands of users. These users have different countries of residence, which are known effect sleep, and therefore an added source of dependence on top of the user level.

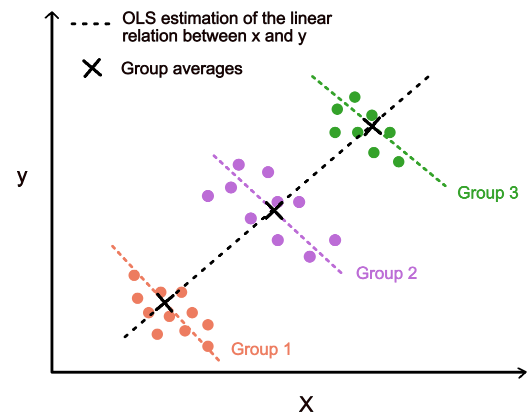


Figure 3.2: The Simpson paradox Example of how linear trend can appear differently in subgroups than across group-averages

3.2.2 Fixed or Random Effects?

Generalized linear models are used to infer a relationship between a response variable and covariates, just like conventional linear models. However, in generalized linear models one of the covariates is categorical and a source of dependence within the data, usually representing experimental or observational units of the data-set (also referred to as groups, levels or effects). If the set of possible levels of the covariate is fixed the data is modelled just using fixed-effects parameters, on the other hand if the levels are a random sample of all possible outcomes, we implement random effects model[220]. The crucial distinction between fixed and random effects models is whether the unobserved individual effect, or the error term, correlates with any independent variables in the model. In the case of random effects, the error term is assumed to be uncorrelated with all independent variables, which allows for time-invariant variables to play a role as explanatory variables (opposite in the case of Fixed effects). However, one needs to specify these characteristics (random effects) that may or may not influence the predictor variable, and in some instances these are unknown or unavailable leading to omitted variable bias[221,211,215].

It is common to model panel data with either fixed or random effects, but in some instances the choice between the two can depend on ones point of view and the subject is still a debate among scientists[222,223,224,213,225]. The argument typically used against random effects model is the fact that first-level independent variables are required to be uncorrelated with the random effects. Some consider that an unrealistic premise since these independent variables vary both within and between clusters (random effects)[224,226]. To the contrary, Bell & Jones (2014) make the rather bold statement that Random Effects model are in fact preferred in almost any occasion if implemented correctly[213]. Random effects model have also rendered same results as fixed effects and performed well even though normality assumptions were violated[227,223]. Nevertheless, discourse concerning the preferred choice between mixed or fixed effects model can get lengthy and detailed, and therefore I rather list examples when the choice between the two is obvious:

- *Fixed effects model*
  - If groups are unique entities and the number of groups is small
  - The modelling aims to understand the characteristics of each level
  - When there is interested in analysing the impact of variables that vary over time.
  - **Examples of Fixed Effects**: Gender, nationality, eye-color, height and more
- *Random effects model*
  - If groups are regarded as a sample from a wider population
  - If the groups are multilevel or, there is at least two sources of random variation
  - If groups are small and multiple, and data-points can be in multiple groups (crossed effects).
  - **Examples of Random Effects**: Individuals, hospitals, schools (randomly sampled from the entire population) and more

Throughout the project I use mixed effects model for different parts of the study and there are two main reasons for that: **i)** The grouping factors (or effects) and the source of dependence originates from subjects (individuals) who are only a sample of the wider population **ii)** Fixed effects models require repeated measurements for grouping factors under range of conditions of the primary independent variable. The primary independent variable is age and median sleep duration and the dependent variable is only measured under one or few conditions. **iii)** The data set is multi-level, or has two sources of dependence (user and country).

3.3 Random effects model

Random effects models are also referred to as mixed effects model because of the way they are defined, but they must always include at least one fixed effect hence, a mix of fixed and random effects. The mixed effects model with random intercept is specified in matrix form as

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}u + \varepsilon, \qquad \text{with } u \sim N_q(0, \mathbf{G}) \text{ and } \varepsilon \sim N_n(0, \mathbf{R}), \tag{3.2}$$

with β representing the fixed effects parameters, u representing the random effects, X representing the $n \times p$ design matrix for the fixed-effects parameters, and Z the $n \times q$ design matrix describing the random effects[215]. The covariance matrix for the error term is $\mathbf{R} = var(\varepsilon)$ has dimension $n \times n$ in many examples, $R = \sigma^2\mathbf{I}$. The covariance matrix for the random effect coefficients, $\mathbf{G} = var(u)$ has dimension $q \times q$, where is q is the number of random effect coefficients. If all random effect coefficients are independent, then $\mathbf{G}$ is a diagonal matrix where $\mathbf{G} = var(u)\mathbf{I}$[215].

One should consider that it is possible to have group averages of the independent variable - meaning there is also a random element to the slope. Groups are then characterized by two random effects; their intercept and slope[225]. This variant of the model is not used in the analysis and therefore not explained further.

Parameter estimation and significance

Parameter estimations for mixed effects model can not be written exactly and therefore a maximum likelihood estimation (MLE) typically applied. The mixed effects model parameters are a vector of the fixed effect, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$ which is the vector of parameters used in the two covariance matrices **G** and **R**. Thus, the likelihood function is a function of the observations and the model parameters, which returns probability of observing a particular observation **y**, given a set of model parameters. However, MLE tends to underestimate the random effects and therefore an alternative criterion is used; the restricted (or residual) maximum likelihood (REML), which is considered the gold standard of parameter optimization in mixed effects models[215,221,218,228]. The point of interest in mixed effects model, is usually estimates of fixed effects but rather random effects, but section A.1 in Appendix A elicits further details about parameter estimation and significance on fixed effects.

Multilevel mixed effects model

The mixed linear model defined in equation 3.2 was defined for a single grouping level, but can be adapted to multilevel grouped data. The matrix notation for a two-level mixed effects model will then be:

$$\mathbf{y}_{ij} = \mathbf{X}_{ij}\beta + \mathbf{Z}_{1,ij}\mathbf{b}_i + \mathbf{Z}_{2,ij}\mathbf{b}_{ij} + \varepsilon_{ij} \tag{3.3}$$

where $\mathbf{b}_i \sim N_{q_1}(0, \mathbf{D}_1)$, $\mathbf{b}_{ij} \sim N_{q_2}(0, \mathbf{D}_2)$ and $\varepsilon_{ij} \sim N_{n_{ij}}(0, \mathbf{R_{ij}})$. Observations are grouped into N first-level groups (indexed $i = 1, ..., N$) each with second level subgroups (indexed by $j = 1, ..., n_i$) where random vectors $\mathbf{b_i}$, $\mathbf{b_{ij}}$ and ε_{ij} are independent of each other. Random effects can either have nested or crossed designs. Crossed design refers to the random effects in mixed effects models, and occur when multiple measurements are associated with multiple grouping variables. In a completely crossed design, all subjects provide responses for all conditions/time-points[225].

In practice The mixed effects models for this study are implemented using R Version 3.5.1 and the `lmerTest` package, the `lmer` function to fit the data set which applies REML and Satterthwaite's degrees of freedom to estimate fixed effects parameters and their significance[229,230,231]. Models are reduced by removing insignificant fixed effects (one at a time) with the `drop1` function which utilizes *F*-test for its estimates.

Conclusion

The data-set used for this study is hierarchical, and users are sampled at different time and rate. There is an added complexity to the structure, since sleep behavior is influenced by cultural context, and a subject's country of residence is also source of dependence. Above, I discuss what is important to consider when analysing the raw data, specifically when aggregating estimates into averages and ways in which error or uncertainty can be quantified. The largest part of the chapter is spent discussing generalized linear models which are type of regression models that have relaxed assumption concerning normality, variance of homogeneity, and independence. In generalized linear models one of the covariates is categorical and a source of dependence within the data. These models are comprised of two main types; *fixed* and *mixed* effects models. The fixed effects variant has a fixed number of possible levels for the covariate with source of dependence, while in mixed effects they are a random sample of all possible outcomes. For the study sample I use mixed effects model, since i) the source of dependence or grouping factors are users who are only a sample of the wider population, ii) the data-set has a two-level structure and iii) fixed effects models require repeated measurements for levels under range of conditions of the primary independent variable, which is not the case for the model set-up. Lastly, I introduce the mixed effects model analytical framework, parameter estimations and practical items for implementation.

Bíum bíum bambaló,

Bambaló og dillidillidó

Vini mínum vagga ég í ró

En úti bíður andlit á glugga

Jónas Árnason

4

Validity of Sleep Activity Data from Wearable Devices

Over the passed couple of decades it has become more common to use wearable devices to measure sleep for the purpose of research[232]. The most common types of these are wrist-actigraph worn on the surface of the skin, and monitor movements that are used to infer sleep. In this chapter I review the advantages and disadvantages for the utilisation of these. Furthermore, I introduce the origin of the data source used for this project and review how the sleep recordings were validated.

4.1 The Data Source

The data was collected with consumer wearable devices from 2015 to 2019 designed to track physical activity and sleep behavior from users all across the world (see Figure 4.1). When users first connect their wearable devices to smartphone, they receive a visual instruction on how and where (wrist) to place the device and they are advised to wear it on their dominant side. The devices use proprietary, *internally validated* algorithms based on movement registered by an internal accelerometer used to infer sleep and wake states in 1-minute interval, or epochs. Epochs are aggregated into nights with sleep onset, offset and duration, and sleep fragmentation is quantified as wake after sleep onset (WASO). Thus, for each registered night WASO is the total time an individual is recorded awake (after defined sleep onset, but also occurring before defined sleep offset)[202].

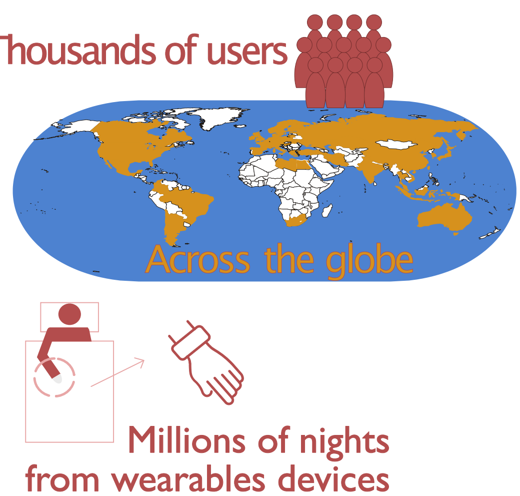


Figure 4.1: The data-set The data stems from consumer wearable devices, which records sleep activity for thousands of users from all over the globe

Data from all devices this study were wirelessly transmitted via Bluetooth to an accompanying mobile phone application, which registered user mobile application (app) usage. Furthermore, the location data originates from GPS traces; these are not collected at a fixed sampling rate but estimates are updated when there is a change is the motion-state of the device (if the accelerometer registers a change). Figure 4.2A shows examples of bed time, wake time and sleep duration recorded over a period of time for three randomly selective users. These estimates are averaged within individuals for all users, and the distribution are illustrated on Figure 4.2B, as well as the distribution for median WASO on Figure 4.2C. Average sleep onset, offset and duration are clearly normally distributed across the population except for WASO which has a large fraction of users with zero instances of WASO (approximately 85 %). That is not unexpected, since the most problematic validity issue with actigraphy is the low specificity in detecting wakefulness within sleep periods, which will be discussed in more details below[232].

4.1.1 General consensus concerning validity of wearable devices

There do exist studies dating back to the mid 90's using wrist-worn actigraphy (wearable devices) to measure sleep for the purposes of research, and the use of these has increased ever since[232]. Multiple studies have been conducted to test the validity of wearable devices and recent review by Kolla *et. al* (2016), reached same conclusion as Sadeh *et al.* (2011), and found wrist-worn actigraphy to underestimate sleep disruptions and overestimate total sleep duration in normal subjects compared to measurements from polysomnography[233]. Another literature review found wrist-actigraphy to consistently overestimate PSG-determined sleep onset latency, *but to a comparatively low degree*[234]. Despite this shortcomings there is a general consensus that wearable devices are usable, and the comprehensive review by Sadeh *et. al* concluded that actigraphy has reasonable validity and reliability in assessing sleep-wake patterns in normal individuals[232]. Furthermore a validation study from 2016 for five different commercial devices concluded that wrist-worn actigraphies can be used to specify total sleep duration[235]. These devices provide cost-effective methods to objectively assess

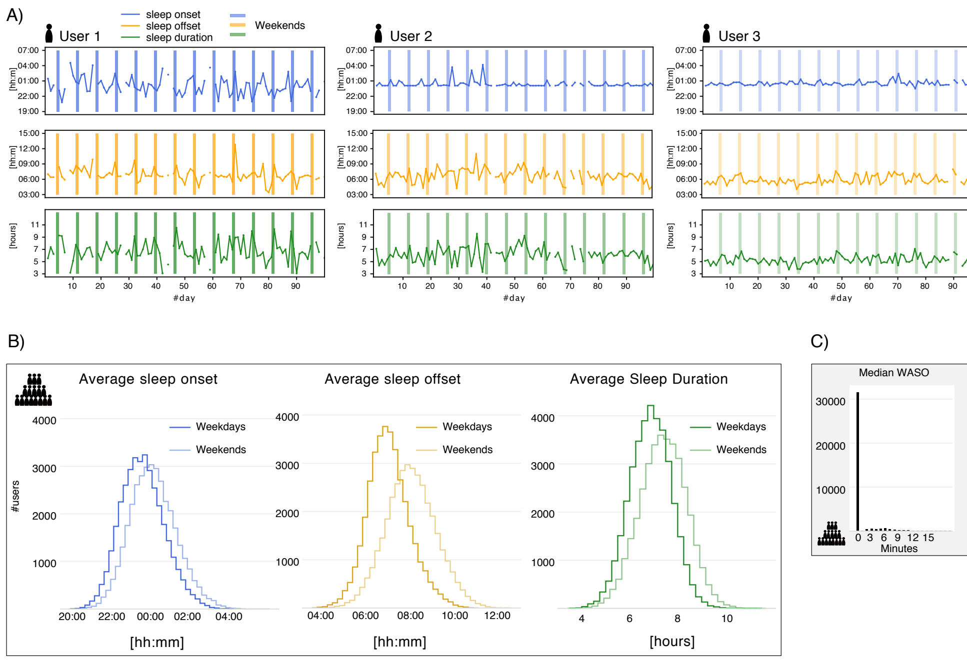


Figure 4.2: Introduction to the data-set

A: Examples of consecutive recordings (100 days) of sleep onset (blue color), offset (yellow color) and duration (sleep duration) for three randomly selected users. The x-axis spans days (0-100) and y-axis shows time for sleep onset and offset, but hours for sleep duration. **B:** The distribution of average sleep onset (blue), offset (yellow) and duration (green) separately for weekday (darker colors) and weekend-nights (lighter colors) across the population **C:** The distribution of median WASO for all users in the data-set sleep in-situ over long period of time. They are not considered a substitute for clinical interviews or overnight polysomnography recordings, but rather provide useful information about sleep in natural environment for extended period of time enabling a more holistic examination of the human daily rhythm of rest and activity[236].

4.2 Data validation

A workshop on wearable devices for measurements of sleep, held by the Sleep Research Society in 2015, recommended that actigraphy or wearable devices should be validated with the gold standard of polysomnography[237]. Unfortunately, that has not been conducted for the wearable device that records sleep for users in this observational study, but I developed an alternative validation approaches. The data-set used for the validation is the same as in my paper *Gender differences in nighttime sleep patterns and variability across the adult lifespan: a global-scale wearables study* and contains 11.14 million nights from 69.650 users, all of who have at least 8 nights of recorded sleep[26].

4.2.1 Convergence with Large & Global Scale Data-sets

To explore the convergence of the study sample with other global and large scale sleep data-sets, I assess whether average estimates of sleep onset, offset and duration at the country-level, and in some instances also by age-group or gender converge with prior published estimates from several other sleep studies. Specifically, I compare my sample to results from Walch *et al.* (A global quantification of "normal" sleep schedules using smartphone data)[86], Roenneberg *et al.* (Epidemiology of the human circadian clock)[15], Ong *et al.* (Large-scale data from wearables reveal regional disparities in sleep patterns that persist across age)[46] and Ford *et al.* (Trends in Self-Reported Sleep Duration among US Adults from 1985 to 2012)[53].

Walch *et al.* (2016) I infer country level estimates of sleep onset, offset and duration from Figures 3A) and B) in the Walch *et al.* paper, thus they may differ marginally from their raw estimates[86]. The data in Walch *et al.* paper is collected with self-reports of 'typical bed and wake-time' rounded to the nearest hour for 5450 users. In order to generate comparable statistics from my study sample, I first estimate individual averages by day-type (weekday and weekend-nights separately), and then compute weighted overall averages using a standard weekday-weekend ratio (2/7 weekend-nights and 5/7 weekday nights average). The reader should note a few items before looking at the results of the comparison; **i)** there are fewer users (5450) in the Walch *et al.* sample, **ii)** it is uncertain how comparable the samples are in terms of underlying demographics (especially age and gender) at the country level and **iii)** I use objective multi-night recordings to obtain country-level averages, while Walch *et al.* used self-reported typical bed and wake-up hours (single estimates) and did not disclose whether these pertained to weekdays, weekends or overall average behavior.
Tables B.1-B.3 in the Appendix B illustrate the comparison between the two samples. The estimates of country-level average sleep duration is higher in Walch *et al.* sample (0.94 hrs at the most), but the relative order of magnitude by countries matches well between the two samples – e.g. both report the Netherlands to have the highest average sleep duration while Japan and Singapore have the lowest. Similarly, country-level averages of bed and wake time were earlier in Walch *et al.* sample, but the countries with earliest and latest bed and wake-up time are the same across data-sets. I use a statistical measure, the Spearman rank correlation, to quantify how well the three measurements correlate between the two data-sets and find $\rho_{onset} = 0.67$, $\rho_{offset} = 0.74$ and $\rho_{duration} = 0.78$ where all three estimates are statistically significant ($p<0.05$).

Roenneberg *et al.* (2007) Now I compare estimates of sleep duration to those reported by Roenneberg *et al.* (2007) but the data was collected with the Munich Chronotype Questionnaire and therefore entails retrospective habitual estimates of sleep duration on weekdays and weekends separately. Since the users in Roenneberg *et al.* sample are predominantly from Germany, Austria, Netherlands and Switzerland, I only include users residing in those geographic regions for the comparison [15]. Table 4.1 reveal that estimates of average sleep duration across the two data-sets closely corresponds, with a weekday average absolute deviation of 3.9% and weekend average absolute deviation of 5.1%. For both weekdays and weekends, my sample has a higher ratio of users in the middle group (7-7.5 hours weekdays, 7.5-8 hours weekends) but fewer are in the group with the longest sleep duration. The percentage of users in the group with shortest sleep duration matches well (1.4% avg. absolute deviation).

Sleep duration	Roenneberg et al. [% users]	Study sampe [% users]
WEEKDAYS		
< 7.0 hours	41.0 %	38.9 %
7.0 - 7.5 hours	21.0 %	26.8 %
> 7.5 hours	38.0 %	34.3 %
WEEKENDS		
< 7.5 hours	34.0 %	34.7 %
7.5 - 8.0 hours	15.5 %	22.7 %
> 8.0 hours	50.5 %	42.6 %

Table 4.1: Estimates of %-point of users within certain range of sleep duration (separately for weekday and weekend-nights) for Roenneberg *et al.* data-set and study sample

ONG *ET AL.* (2019) Ong *et al.* conducted a study on regional differences with nearly half a million objectively measured nights from approximately 24 000 users living in five different countries. The data-set from Ong *et al.* and the study sample might be the most compatible for comparison since they both consist of objective multi-night recordings in-situ. However, they differentiate on couple factors; **i)** there are different types of wearable devices used to measure sleep (Fitbit in Ong *et al.* paper), **ii)** the sample from Ong *et al.* has higher number of users per country but fewer countries, **iii)** there is a higher proportion of female users in Ong *et al.* study and **iv)** a slightly wider age range than in the study sample of this project.

I compare country-estimates of sleep duration by looking at the percentage of users sleeping more than 7 hours, separately on weekends and weekdays (see Table 4.2). These proportions correspond closely, with a weekday country-level averages differentiate by 3.6 percentage points and weekend differentiate by 3.5 percentage points. The smallest national deviation between samples was for Hong Kong (.4%) on weekdays, and the largest difference between the samples was for South Korea on weekdays, where my sample had 7.3% fewer users sleeping 7 hours or more.

Country	**Australia**	**Hong Kong**	**Singapore**	**South-Korea**
WEEKDAYS				
Ong *et al.* **% users w/ duration >7 hrs**	61.0 %	34.0 %	27.0 %	29.0 %
Study sample **% users w/ duration >7 hrs**	65.7 %	33.6 %	25.0 %	21.7 %
WEEKENDS				
Ong *et al.* **% users w/ duration >7 hrs**	74.0 %	58.0 %	51.0 %	52.0 %
Study sample **% users w/ duration >7 hrs**	76.5 %	56.6 %	57.8 %	48.6 %

Table 4.2: %-point of users sleeping 7 hours or more (separately on weekday and weekend-nights) for Ong *et al.* and study sample

Furthermore, I compare country estimates separately by gender (male/female) for sleep onset, off-set and duration on weekdays. The averages for Ong *et al.* data-set are estimated from Figure 2A) in the paper, thus uncertainties might be imposed[46]. Error estimates were reported on the figure in Ong *et al.* paper, but were too small to read off, nonetheless I report averages with standard error of the mean from the study sample (see Table 4.3). The estimates for women never fall within range of standard error of the mean and differences are larger for averages of sleep onset (ranging 27 - 36 minute difference), while sleep duration and wake times match relatively well (ranging 2 - 14 minute difference). Similar differences are identified for men, while national estimates of sleep onset, off-set and duration correspond well across Ong *et al.* sample and the study sample. The differences for sleep onset across data-sets cannot be explained directly, but the data-sets might differentiate in terms of age and gender representation of users, or the devices measure sleep onset differently.

	Australia		Hong Kong		Singapore		South-Korea	
WOMEN	**Ong *et al.***	**Study sample**	**Ong *et al.***	**Study sample**	**Ong *et al.***	**Study sample**	**Ong *et al.***	**Study sample**
Sleep onset [hh:mm ± mm]	22:51	23:22 ± 4	00:14	00:50 ±4	23:57	00:24 ± 7	23:51	00:21 ± 3
Sleep offset [hh:mm ± mm]	06:45	06:57 ± 3	07:34	07:45 ±4	07:06	07:18 ± 7	07:11	07:13 ± 3
Sleep duration [hours]	7.28	7.46 ±0.05	06:09	6.84 ± 0.07	6.56	6.80 ± 0.08	6.71	6.74 ± 0.03
MEN	**Ong *et al.***	**Study sample**	**Ong *et al.***	**Study sample**	**Ong *et al.***	**Study sample**	**Ong *et al.***	**Study sample**
Sleep onset [hh:mm ± mm]	23:06	23:40 ± 3	00:27	00:45 ± 3	00:00	00:29 ± 4	00:00	00:34 ± 1
Sleep offset [hh:mm ± mm]	06:43	06:52 ± 3	07:35	07:27 ± 3	07:00	07:04 ± 3	07:07	07:02 ± 1
Sleep duration [hours]	7.0	07.09 ± 0.04	6.5	6.6 ± 0.04	6.45	6.50 ± 0.05	6.5	6.37 ± 1

Table 4.3: Average sleep onset, offset and duration (with SEM for the study sample) by country and gender separately for Ong *et al.* data-set and study sample

FORD *ET AL.* (2015) Lastly, I compare measures of average sleep duration by gender and age group for a subset users residing in the US to self-report data from the US National Health Interview Survey conducted in 2012. The results are listed in Table 4.4. The estimates for men differentiate the most for the youngest and oldest groups (18-24 and 55-65), while the standard error of the mean overlaps for other age groups (except slight deviation for age group 35-44). The differences across the two data-sets is larger for women, but the sample of women is smaller than for men ($N_{women} = 317$ and $N_{men} = 624$). Furthermore, we can not know how well the sociodemographic composition of the two samples correspond.

Age group	**18-24**	**25-34**	**35-44**	**45-54**	**55-64**
Men NHIS data-set [hrs ± hrs]	7.45 ± 0.05	7.08 ± 0.03	6.99 ± 0.03	6.94 ± 0.04	7.09 ± 0.03
Men study sample [hrs ± hrs]	7.08 ± 0.07	7.08 ± 0.05	6.86 ± 0.06	6.84 ± 0.04	6.82 ± 0.1
Women NHIS data-set [hrs ± hrs]	7.46 ± 0.04	7.13 ± 0.03	7.05 ± 0.03	6.98 ± 0.03	7.05 ± 0.03
Women study sample [hrs ± hrs]	7.31 ± 0.2	7.53 ± 0.08	7.41 ± 0.08	7.37 ± 0.08	6.94 ± 0.1

Table 4.4: Comparison of average sleep duration (with SEM) by gender and age group for users in the study sample residing in the US to estimates from the US National Health Interview survey sample (2012)

4.2.2 Consistency of measurements over time

To ensure that hardware or firmware changes over the period of data collection did not influence the measurements, I estimate average sleep onset, offset and duration for the top three countries for full years of data collection (2016, 2017 and 2018). The estimates in are illustrated in Table 4.5 and show no evident trends between years and deviations are small and random.

Country & Year	Sleep Onset hh:mm ± mm		Sleep Offset hh:mm ± mm		Sleep duration hrs ± hrs	
	Weekdays	Weekends	Weekdays	Weekends	Weekdays	Weekends
Japan 2016	00:18 ± 1	00:40 ± 1	06:44 ± 1	07:38 ± 1	6.31 ± 0.01	6.84 ± 0.01
Japan 2017	00:14 ± 1	00:38 ± 1	06:42 ± 1	07:40 ± 1	6.35 ± 0.01	6.84 ± 0.01
Japan 2018	00:13 ± 1	00:34 ± 1	06:39 ± 1	07:32 ± 1	6.32 ± 0.01	6.82 ± 0.02
Germany 2016	23:31 ± 1	00:21 ± 1	06:50 ± 1	08:20 ± 1	7.21 ± 0.02	7.82 ± 0.02
Germany 2017	23:36 ± 1	00:18 ± 2	06:46 ± 1	08:14 ± 1	7.21 ± 0.02	7.78 ± 0.02
Germany 2018	23:27 ± 2	00:13 ± 2	06:44 ± 2	08:05 ± 2	7.16 ± 0.02	7.70 ± 0.03
UK 2016	23:45 ± 2	00:17 ± 2	07:09 ± 2	08:10 ± 2	7.28 ± 0.02	7.75 ± 0.02
UK 2017	23:43 ± 2	00:18 ± 2	07:07 ± 2	08:11 ± 2	7.28 ± 0.03	7.74 ± 0.03
UK 2018	23:47 ± 3	00:18 ± 3	07:09 ± 3	08:11 ± 3	7.24 ± 0.04	7.74 ± 0.05

Table 4.5: Average sleep onset, offset and duration with SEM (separately for weekday and weekend-nights) for full year of data collection (2016, 2017 and 2018) and the three countries with most users

Furthermore, I visualise the median sleep duration by day over the entire period of data collection on Figure 4.3. There are no obvious jumps outside of characteristic seasonal, monthly and weekly patterns and no year-over-year trends.

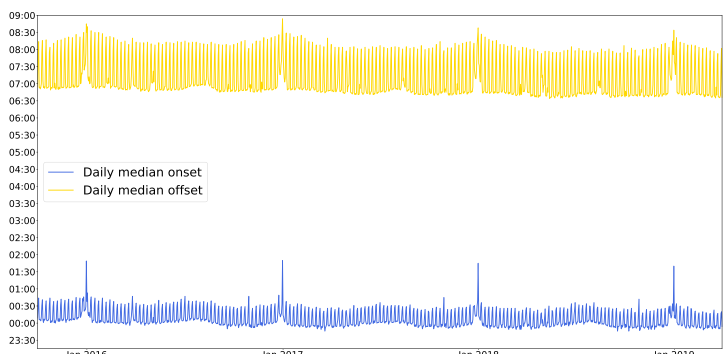


Figure 4.3: Consistency over time
Daily median estimate of bed and wake-time for the entire data-set, throughout the period of data collection

4.3 Demographic representation

The data-set is global with thousands of users residing in ~ 150 countries and large scale with millions of recorded nights. In terms of country of residence for users, some geographic areas more dominant than others, and for example about 1/4 of users live in Japan, and about 50 % of the users reside in five countries. Users are anonymous and self-report their age, gender, height and weight. All data analysis was carried out in accordance with the EU's General Data Protection Regulation 2016/679 (GDPR) and the regulations set out by the Danish Data Protection Agency. The GDPR describes regulations for data protection and privacy in the European Union and the European Economic Area; it also addresses the transfer of personal data outside the EU and EEA areas.
About 1/3 of the users are female and span age range 19-67. To explore representation by age, I compare how well median age at country-level matches with statistics from the United Nation Population Division (UN) for the top five countries (see Table B.4 in Appendix B)[238]. The median values in our sample and the overall population correspond well: users from Japan are slightly younger (by 1 year) while those from Taiwan and the United Kingdom match their respective reference populations. Users from Germany are younger (by 7 years) and also those from Russia (by 5 years). Additionally, I compare age standardized BMI statistics of the study sample to population estimates provided by the World Health Organization (WHO) in Table B.5 in Appendix B[239,240]. Both men and women from all countries fall within or place near the 95% confidence intervals (CIs) of the WHO reference values. Women from the UK fall 0.5 points above the 95% CI and women from Japan average 0.5 points below the 95% CI reference range. The data is observational and users are included if they choose to buy the commercial devices. One must consider that owners of wearable devices may not be representative of the wider population due to potential unobserved factors. Ownership of physical activity tracking devices have been associated with post-secondary education and higher prevalence of physical activity[241,242]. However, there are not many studies that have

explored biases among wearable device owners and most studies were conducted within Western countries (N-America or Europe). Thus, I can not know how these results translate to a larger sample and wider range of geographic regions.

4.4 Data Pre-processing

To reduce the risk of including sleep observations from those suffering from insomnia, shortened night due to users ceasing wristband use in the middle of the resting period, observations from night-shift workers or any other possible data errors, outliers were removed. Up until now, I have only managed to find one example of a study which describes standard filters for sleep onset and offset (Walch *et al.* 2016) which was though applied to data with self-reports of typical bed and wake time rounded to the nearest hour. Users reporting typical wake time before 03:00 or after 11:00, and those with bed times before 19:00 and after 03:00 were filtered out. In sleep epidemiology there exist multiple examples of studies which use standard filters for outlier detection for sleep duration. Typically sleep duration is required to be more than 3 hours, and less than 13 hours[15], or even more conservative (two hours less inclusive) where the criteria is 4 hrs < sleep duration < 12 hrs[243,86].
I did try to design filters to adopt to individual level distributions for sleep onset, offset and duration. For example I experimented with number of standard deviation around the mean, number of interquartile range around the median, isolation forest and others methods I developed myself. Unfortunately, I considered none good enough since in some instances they excluded entries that could potentially be nighttime sleep. Later it will become evident how wide ranged and unpredictable sleep behavior actually is, which renders the task of creating individual level outlier detection complex. Throughout the project I filter sleep duration by standard filters (3 hrs < sleep duration < 13 hrs) and apply filtering to sleep onset and offset with different number of standard deviation around the mean for distributions of all nights.

Conclusion

The data-set originates from consumer wearable devices, and records sleep activity in 1-minute epochs which are aggregated into nights with bed time (sleep onset), wake time (sleep offset), sleep duration and WASO (wake-time after sleep onset). The wristbands have not been publicly validated using the gold standard of polysomnography as recommended in the Sleep Research Society Workshop on wearable devices for the measurement of sleep[237]. However, I find the aggregates of average sleep onset, offset and duration to converge with country-level estimates from other global and large-scale sleep studies. The largest discrepancies are detected in the comparison for sleep onset across data-sets, which could be anticipated since that is were both self-reports and different types of wearable devices are likeliest to deviate. First considering self-reported estimates of sleep, the variation occurs since it can be difficult for people to be aware of time when lying bed at night. Their estimates therefore rely on i) their perception of time and ii) how much time has passed since they last saw a clock. The deviation in sleep onset between different types of wrist-worn actigraphies can be rationalised due differences in hardware and/or firmware to measure sleep specificity (detecting wakefulness while lying in bed), which is acknowledged to be the characteristic in which these devices are most imprecise[232]. Furthermore, I examine whether hardware or firmware changes over the period of data collection influenced the data, but I find the recordings to demonstrate consistency over the period of data collection. Lastly, I explore country-level demographic representation and find good agreement between the World Health Organization's estimates of median age and age standardized BMI.

Illa dreymir drenginn minn;
Drottinn sendu engil þinn
vöggu hans að vaka hjá,
vondum draumum stjaka frá.
Láttu hann dreyma líf og yl
ljós og allt sem gott er til,
ást og von og traust og trú.

Bergþóra Árnadóttir

5

The Wild and Mysterious ways of sleep

Large scale sleep studies have been constrained to self-reported data from sleep surveys, diaries and questionnaires which are known to recall biases. Consequently, this type of research has been restricted to analysis with a limited set of variables. Here I explore whether any salient features of sleep may have been missed by these limitation. I present a new visualization method which illustrates individual characteristic sleep patterns, which inspire a set of features to study sleep patterns across the population. Furthermore, I propose a novel data-driven metric which may allow us to estimate whether or not an individual's physiological sleep need are met or not. Note that all figures in this chapter are reused from `paper 3`.

5.1 The Multifaceted process of sleep

I begin by studying the range of behaviors that can result in similar estimates of traditional sleep epidemiology measures such as chronotype, social jetlag and habitual sleep duration. To do that, I select four users (labelled 1-4) to exemplify how wild and complex sleep actually is when measured objectively in-situ over long period of time. Estimates of median sleep onset, offset and duration (separately for weekday and weekend-nights), as well as social jetlag and chronotype are provided for the users 1-4 in Table 5.1. These summary statistics reveal that there are some similarities but also distinct differences which I further explore on Figure 5.1.

Visual sleep trajectories A convenient way to visualise sleep recordings over time is shown in Figure 5.1C-F. The x-axis spans consecutive nights or day number, and the y-axis spans a range of time (from 20:00 to 12:00). Sleep onset and offset are represented with blue and yellow dots respectively, and weekends are marked with a gray shaded area of the sleep interval. I call these *visual sleep trajectories* and they capture many aspects of sleep patterns, such as typical bed and wake time, sleep regularity, temporal variations, and differences between weekend and weekday behavior.

Comparing the four sleep trajectories Figure 5.1A shows how user 1 and 2 have nearly the same chronotype ($\sim$04:30) selected within the range of the dotted lines and similarly. However, from observing their sleep trajectories on Figure 5.1C and D it is evident that user 1 has high variability both in terms of bed and wake time but user 2 is super-regular with substantial difference in wake-time between weekends and weekdays. Figure 5.1A shows how user 3 and 4 have nearly the overall same median sleep duration ($\sim$7.5 hrs), while in terms of long term patterns user 3 has high

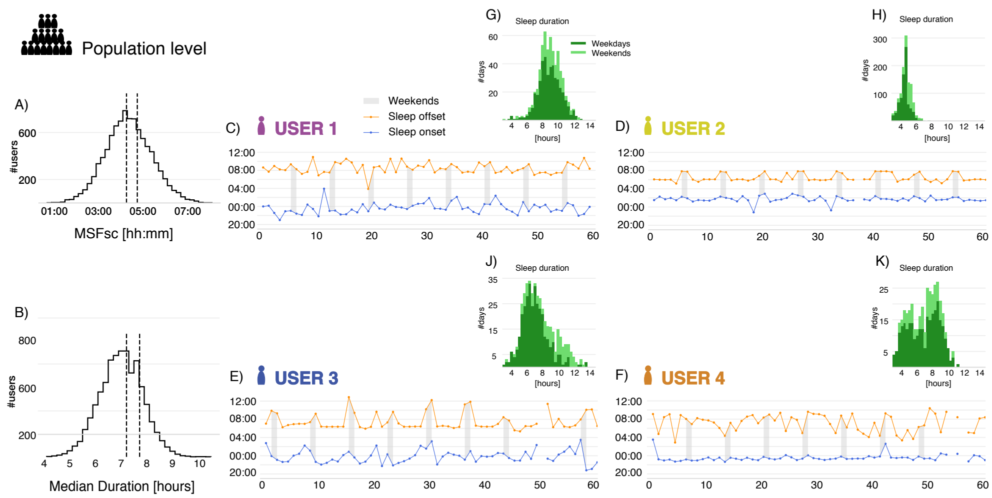


Figure 5.1: The multifaceted process of sleep
A: The distribution of chronotype (MSF_{sc}) across the population, and the dotted vertical line represent the range in which user 1 and 2 were selected from. **B:** The distribution of median sleep duration across the population, and the dotted vertical lines represent the range in which user 3 and 4 were selected from. **C-F:** Temporal patterns, or *sleep trajectories*, for 60 consecutive nights of bed and wake time. The y-axis shows the hours from 20:00 to 12:00. Sleep onset and offset are marked with blue and yellow dots respectively, and weekend-nights are shown by a gray shading of the sleep-interval. Users 1 & 2 have nearly the same chronotype ($\sim$ 04:30) and users 3 & 4 have nearly the same overall medians sleep duration (~ 7.5 hrs). **G-K:** The stacked distribution of sleep duration (dark green color represents weekday-nights and lighter green color weekends) for users 1-4 respectively.

USER	1	2	3	4
Overall				
MSFsc [hh:mm]	04:26	04.35	05:13	03:42
Social jetlag [hrs]	0.27	0.88	1.45	0.13
Median duration [hrs]	9.10	4.82	7.30	7.4
Weekdays				
Median Onset [hh:mm]	23:27	01:17	23:56	23:24
Median Offset [hh:mm]	08:28	06:05	06:56	07:42
Median duration [hrs]	9.00	4.7	7.0	7.3
Weekends				
Median Onset [hh:mm]	23:29	02:00	00:12	23:34
Median Offset [hh:mm]	09:08	07:15	09:52	07:56
Median duration [hrs]	9.2	5.2	9.65	7.5

Table 5.1: Estimates of median bed and wake time (separately for weekday and weekend-nights), as well as chronotype, social jetlag and overall median sleep duration for four selected users (labelled 1-4)

weekend-weekday behavioral differences and user 4 has an interesting behavior of falling asleep at similar time every night, while he/she wakes up at wide range of hours (see Figure 5.1E and F). Additionally, I present three other users (labelled 5-7) in Table 5.2, all which have approximately the same social jetlag (SJ) $\sim$0.75 hrs. These estimate reveal that user 1 has similar wake time on weekends and weekdays, but bed time is advanced to later hours on weekends leading to overall less sleep duration. User 2 shifts both bed and wake-time to later hours on weekends and obtains nearly the same sleep duration on weekends and weekdays. User 3 goes earlier to bed and wakes up later on weekends, thus obtains substantial more amount of sleep on weekends (over 3 hours more).

Users with SJ $\sim$ 0.75 hrs	5	6	7
Onset weekdays [hh:mm]	23:09	23:10	02:00
Offset weekdays [hh:mm]	08:15	06:00	07:37
Onset weekends [hh:mm]	01:16	23:53	00:48
Offset weekends [hh:mm]	08:06	06:59	10:28
Weekend-weekday duration difference [hrs]	-1.3	0.0	3.3

Table 5.2: Estimates of median bed and wake time (separately for weekday and weekend-nights), and weekend-weekday median sleep duration difference for three selected users with the same social jetlag ($\sim$ 0.75 hrs)

A NEED FOR NEW METRICS? The temporal patterns on Figure 5.1 and estimates in Table 5.2 reveal that even though one characteristic of sleep is measured approximately the same across users, there can be very different underlying patterns and behaviors at play. I generalise that conclusion by providing the distributions of various sleep metrics for users with the similar chronotype (∼04:30), social jetlag (∼ 0.75 hrs) and median sleep duration (∼7.5 hrs) on Figures C.1-C.3 in Appendix C. These distributions illustrate all the different characteristic that groups of users with similar epidemiological measures of sleep can comprise.

5.2 New visualisation method: Sleep Portraits

Thus far, we have observed some weaknesses in regards to traditional sleep epidemiology metrics. To help understand the complex patterns we have gotten sense of, I introduce *the sleep portrait* that visualises variation in bed and wake time as 2d-histograms (with half hour bins) separately for weekday and weekend nights. I provide examples of these for users 1-4 in Figure 5.2E-H where each square represents amount of nighttime sleep, and to make the visualisation more accessible for interpretation, I mark recommended sleep duration (7-9 hrs) with grey stepped lines[196,197,198,199]. The sleep portraits reveal and highlight patterns that were just vaguely present in the sleep trajectories for users 1-4 in Figure 5.2A-D.

INDIVIDUAL-LEVEL SLEEP CHARACTERISTICS The sleep portrait in Figure 5.2E shows that user 1 has small weekend-weekday differences in sleep behavior, except that wake time is slightly advanced to later hours on weekends. There is large variability in wake and bed time, clearly identified from the broad areas that the points cover on the sleep portrait. User 1 tends to sleep a lot, often more than 9 hours. User 2 has an extremely regular behavior observed on the focused sleep portrait in Figure 5.2F where the data-points cover a small area. Furthermore, user 2 consistently attains less sleep than recommended, on weekends as well, since all the data-points are below the line of the

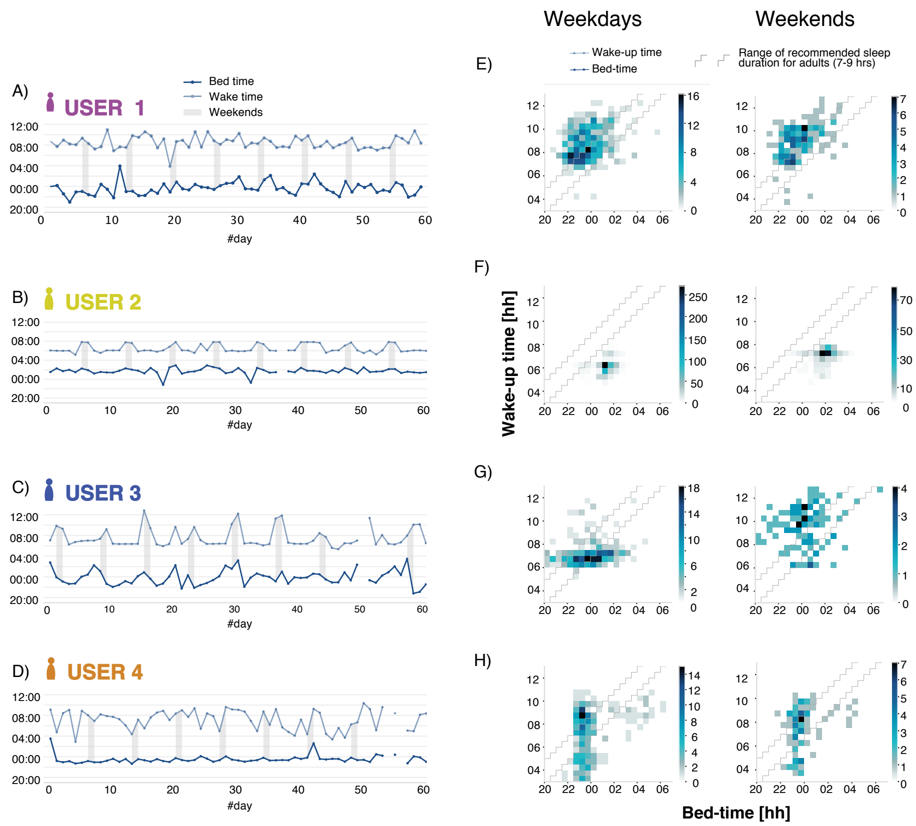


Figure 5.2: Capturing the complexity with Sleep portraits

A-D: Sleep trajectories for users 1-4 with $\sim$60 consecutive nights of bed and wake time. The y-axis shows the hours from 20:00 to 12:00. Sleep onsets and offset are marked with dark and light blue dots respectively, and weekend-nights are shown by a gray shading of the sleep-interval. **E-H:** Examples of *the sleep portraits* for users 1-4, which are 2d-histograms (1/2 hours bins) of bed time on the x-axis (range 22:00-07:00) and wake-time on y-axis (range 03:00-13:00). For each user, the left plot represent weekdays and right plot weekends. The color-bar to the right (of each plot) illustrates the number of nights that take place in each square. The grey step-lines represent the area of recommended sleep duration (7-9 hrs).

shortest recommended nighttime sleep. The horizontal spread of the data-points on weekdays for user 3's sleep portrait in Figure 5.2G, indicates that this is person that has to wake up before certain point of time on weekdays. This constraint likely induces the "explosion of freedom" on weekends, where wake times are advanced to later hours and the individual obtains more sleep. Lastly, the data points span a broad range in the vertical direction, but has a narrow horizontal spread for user 4 in Figure 5.2H. This implies that user 4 tends to fall asleep at approximately the same time every night, while they wake up at variable hours.

5.3 Informative features for analysis of multi-night recordings of sleep

The sleep portraits inspire a set of features which I consider informative and helpful to use when analysing high-resolution sleep activity data-set. As clearly seen in examples of the sleep portraits in Figure 5.2E-H, the shape and location of the 2d-histograms (also referred to as the point-cloud) can tell us a lot about a person's sleep habits.

Typical behavior. To begin with, an estimate which quantifies typical behavior for bed time, wake time and sleep duration is always informative. I use the median, which gives a better estimate than the mean which tends to be influence by extreme values (which I provide more details about later on in section 5.4). Distributions for median sleep onset, offset and duration (separately by day type) is illustrated in Figure 5.3A.

Sleep regularity. Secondly, the width of individual's behavior entails important information and measures sleep regularity. Typically standard deviation (std) is used to quantify sleep regularity[36,86], but I rather suggest a measure based on quantiles since they are less impacted by extreme events, and actually more interpretable. I provide an example to explain; consider having to comprehend either "John has 0.74 hours std in wake-up time on weekdays" or "John wakes up 80 % of the time within a span of 15 minutes on weekdays". Thus, I suggest using a measure called *width* which is the difference between the 90th and 10th percentile, and provides a range in which 80 % of a person's sleep takes place in. The width correlates very well with std (ranging from 0.933 to 0.968, see Figure C.4 in Appendix C) and the distribution of the width for sleep onset, offset and duration separately for weekdays and weekends is presented on Figure 5.3B.

Comparison of bed and wake-time regularity Now, the difference between sleep regularity of bed and wake time can also include valuable information. To compare the two I divide width of sleep onset with the width of sleep offset (separately for weekdays and weekends), defined concretely as:

$$R_{\text{onset \& offset}} = \frac{onset_{width}}{offset_{width}}$$

The distribution of $R_{\text{onset \& offset}}$ can be observed on Figure 5.3C where the average is 1.64 on weekdays and 1.18 on weekends. From the measure I can infer the following:

- $R_{\text{onset \& offset}} << 1$ then there is more variability in wake time than bed-time
- $R_{\text{onset \& offset}} \sim 1$ the width of bed and wake-time is approximately the same
- $R_{\text{onset \& offset}} >> 1$ then there is more variability in bed-time than wake-time

Weekend-weekday sleep behavior differences Previous research has established the importance of measuring sleep separately on weekdays and weekends, and to pay attention to the constraint that weekly social schedules can impose on sleep[44]. To do so, I derive a variant of the measure of social jetlag where I calculate the difference between median estimates of a behavior on weekends and weekdays. The distributions for weekend-weekday median difference for sleep onset, offset and duration are illustrated in Figure 5.3D. Similarly, it is important to understand how the behavior changes on weekends compared to weekdays. For example observe the sleep portrait for user 3 on Figure 5.2G where an important characteristic is the change in the width of sleep offset from weekdays to weekends, or the notorious "explosion of freedom". To capture this, I calculate the ratio between the width of a measure on weekends divided by its width on weekdays – defined concretely as $ratio_{width} = \frac{width_{weekends}}{width_{weekdays}}$. The distributions of these estimates for sleep onset, offset and duration is presented on Figure 5.3E, where we can evidently see the effect of alarms on the distribution weekend-weekday width ratio for sleep offset.

5.4 Skew: The Measure for Direction of Preference

Having discussed shortcomings of traditional sleep epidemiological measures and presented new sleep metrics, I now move on to introduce a novel data-driven metric for direction of preference, discovered by exploration of empirical sleep patterns.

Positive skew as a function of typical sleep duration When observing examples of individual's distribution of sleep duration (for example in Figure 5.1C-F and Figure 5.4B) I observe different shapes and characteristics. I believe that symmetry or asymmetry of a distribution entails a important information about an individual's sleep behavior, and may provide information about direction preference, or sleep need. The first evidence to support that idea is provided in Table 5.3. There I list the percentage of users that have an average larger than the median within a sleep group, but *a sleep group* is defined as the median sleep duration rounded to nearest half hour bin.

Median sleep duration [hour]	**4.5**	**5.0**	**5.5**	**6.0**	**6.5**	**7.0**	**7.5**	**8.0**	**8.5**	**9.0**	**9.5**
% of users with $\mu >$ M	100	95	84	77	64	53	47	37	28	20	14

Table 5.3: %-point of users with an average larger than median within sleep group (defined by rounding median sleep duration to the nearest half hour)

For users who typically obtain short nighttime sleep, it is almost given that their average will be larger than the median sleep duration, meaning there is disproportional tendency for nights with

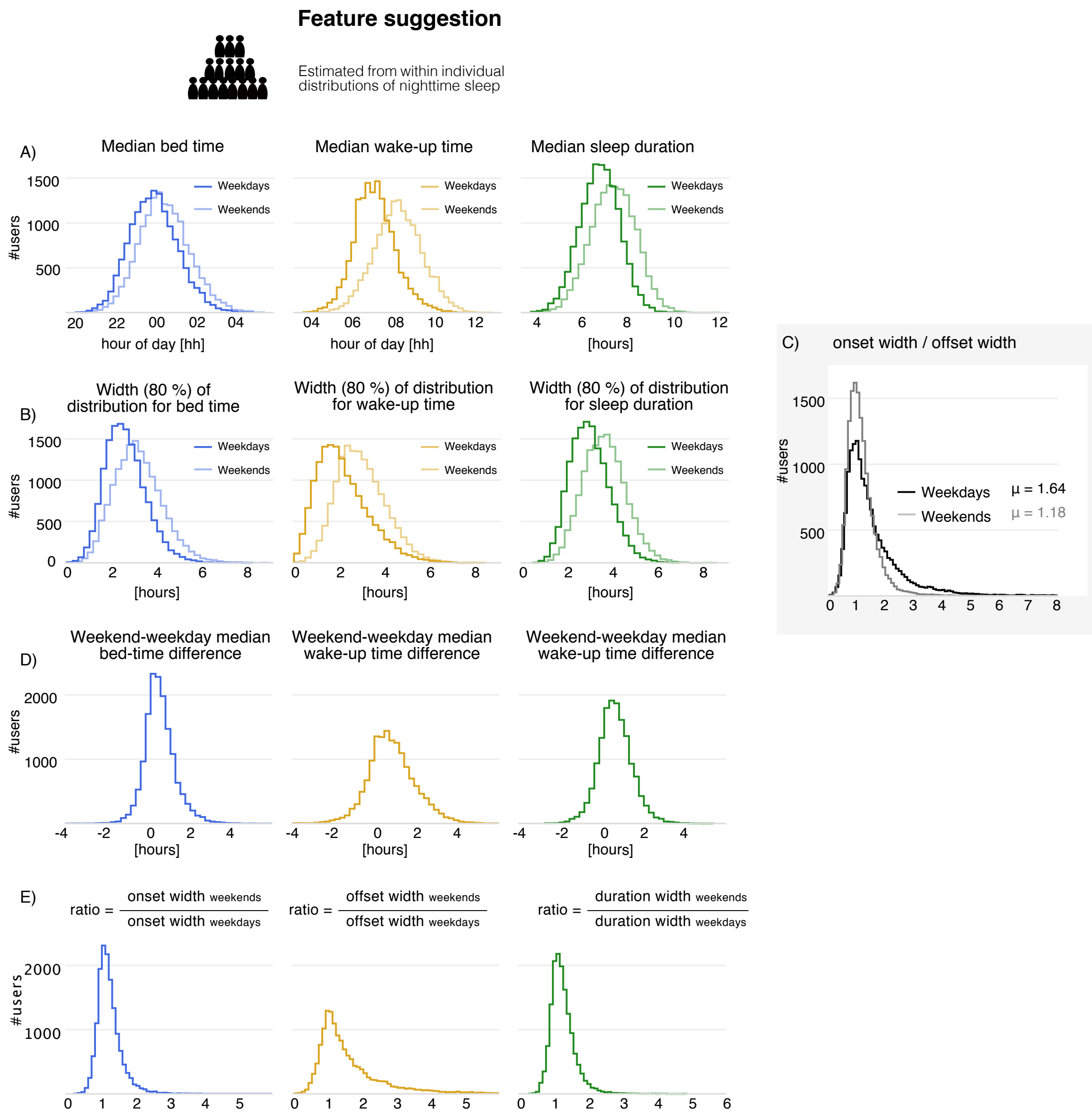


Figure 5.3: Informative features for analysis of multi-night recordings of sleep
Blue, yellow and green colors represent metrics derived from sleep onset, offset and duration respectively. **A:** The distribution of median bed time, wake time and sleep duration separately for weekend and weekday-nights. **B:** The distribution for the width (10th percentile minus the 90th percentile) for bed-time, wake time and sleep duration separately for weekday and weekend-night. **C:** The distribution of ratio between width of sleep onset and sleep offset, separately for weekday (black) and weekend nights (grey). **D:** The distribution for the weekend-weekday median difference for bed-time, wake time and sleep duration. **E:** Weekend-weekday width ratio for bed-time, wake time and sleep duration.

longer duration than typically obtained. The prevalence of this behavior within sleep groups decreases as median sleep duration increases.

The physiological mechanism relating to skew The physiological mechanism that likely contributes to this systematic change across the population, is the process of *sleep-wake homeostasis* – an intrinsic mechanism that maintains sleep pressure and ensures that sleep takes place every night. Individuals accumulate sleep pressure from the end last adequate sleep. Those who consistently sleep less than physiological needed, will build up sleep pressure which can be eliminated with a long night's rest, named "catch-up nights"[13,244]. These "catch-up nights" are more probable to take place for individuals who consistently sleep little (have low median sleep duration), resulting in a right skewed distribution of sleep duration. Similarly, the opposite is expected for those who maintain a longer sleep than they can sustain – leading to disproportionate amount of shorter nights and a left-skewed distribution.

To further examine this, I employ a more sophisticated measure to quantify asymmetry of a distribution – *skewness*. I calculate the moment coefficient of skew for each individual's distribution of sleep duration the following way[245]:

$$skew = \frac{m_3}{s^3} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{\left[\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right]^{\frac{3}{2}}} \tag{5.1}$$

where m_3 is the sample third central moment, s is the sample standard deviation and n is the number of data-points.

Linear dependence of skew and median sleep duration Now skew is estimated for individual's distribution of sleep duration (some examples of these are provided in Figure 5.4B), which is then aggregated into averages for each sleep group and illustrated with the standard error of the mean (SEM). The results are shown in Figure 5.4A, where skew seems to be a function of typical sleep duration and is also a normally distributed trait across the population (see insert in lower left corner in Figure 5.4A). I also observe that the sign of skew changes between 7 and 7.5 hours of sleep, which is within recommended amount of sleep duration for adults [196,197,198,199].

Low skew for extremely short or long typical sleep duration If we assume that the hypothesis that skew renders information about individual sleep need, then skew estimated to be ~ 0 for a distribution of sleep duration should be interpreted as such that the individual is getting sufficient amount of sleep. Therefore it is interesting to consider those that have extremely short median sleep duration, but have no or even negative skew (disproportional tendency for shorter than typical nights). In the same manner, individuals with extremely long median sleep duration with no or positive skew also exhibit 'out-of-ordinary' behavior. The box-plot on Figure 5.4A shows that these individual do exist. I first explore characteristics for those who sleep little (<6.5 hours) and have neutral or negative skewed distribution of sleep duration (skew <0.25). I find that these individuals more likely to be old, male and from the East (Asia), all demographic variables associated with short sleep duration (for further details see section C.3 in Appendix C). Similarly, there are individuals who sleep long (duration> 8 hours) and have positive or no skew (skew >-0.25), but turned out more likelier a part of demographic groups associated with long sleep duration – young, female and Western (residing in N-America or Europe).

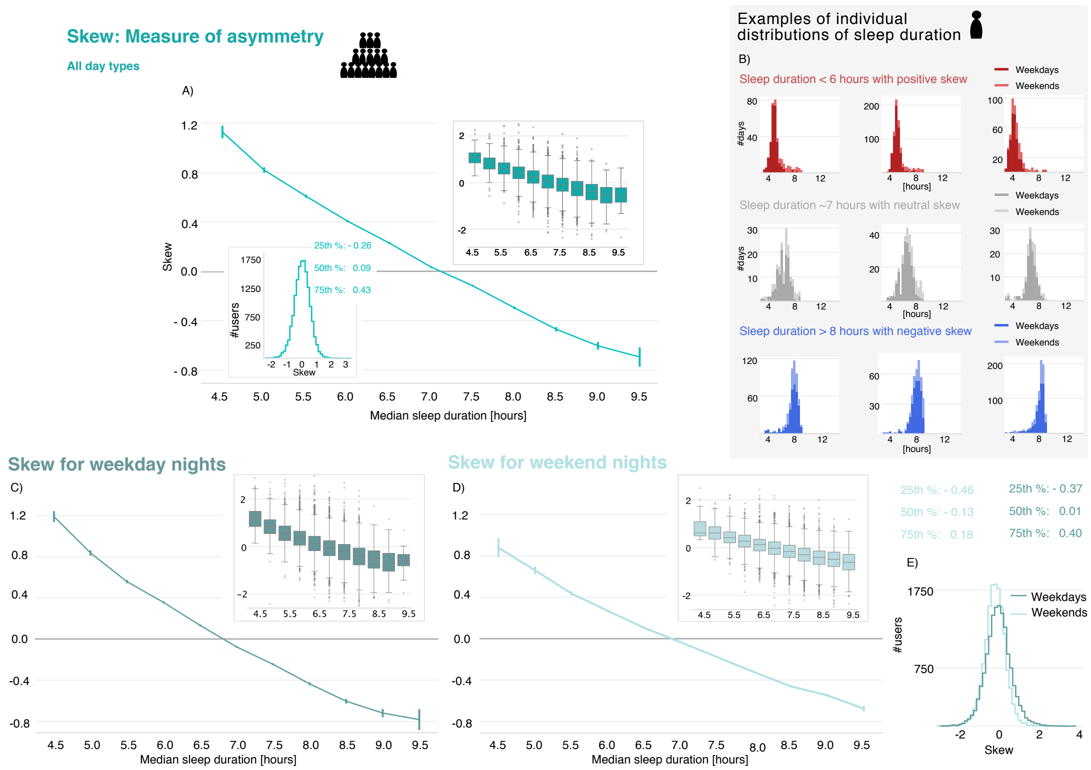


Figure 5.4: Measure of asymmetry of the distribution of sleep duration across the population
A: Aggregated skew, from the individual level for distribution of sleep duration, averaged by sleep group. Sleep groups are defined by rounding median sleep duration to the next half hour bin and error estimates or marked with the standard error of the mean (SEM). The box-plot in the upper right corner shows the underlying distribution behind the average for each sleep group, and the insert in the lower left corner illustrates how skew (for individual level distribution of sleep duration) distributes across the sample. **B:** Examples of individual level distributions of sleep duration for users with positive (red), neutral (grey) and negative (blue) skew. The distributions are stacked, where darker colors represent weekday-nights and lighter color weekend-nights. **C:** Aggregated skew, estimated for individual level for distribution of sleep duration for weekday-nights, averaged by sleep group. Error estimates or marked with the standard error of the mean (SEM). The box-plot in the upper right corner shows the underlying distribution behind the average for each sleep group. **D:** Aggregated skew, estimated for individual level for distribution of sleep duration for weekend-nights, averaged by sleep group. Error estimates or marked with the standard error of the mean (SEM). The box-plot in the upper right corner shows the underlying distribution behind the average for each sleep group. **E:** The distribution for skew (estimated for individual distribution of sleep duration) for weekday and weekend-nights across the population

Skew for weekday and weekend-nights Generally it is consider important to measure and study sleep separately for weekend and weekday nights, since weekly social schedule tend to constrain normal behavioral patterns [184,44]. Skew up until now has been estimated for distribution of sleep duration for all nights, but it can also be quantified by day-type. Figures 5.4C and D respectively show average skew estimated for sleep groups (with SEM) for weekday and weekend nights separately. The same pattern is observed as before; skew decreases as median sleep duration increases, both on weekdays and weekends. However, skew is more conservative on weekends and spans a more narrow range, also clearly observed on Figure 5.4E where the distribution of weekday and weekend skew are compared. This could be explained by the fact that people are more likely constrained by the social clock and early morning work schedules on weekdays. Thus, there is less time and flexibility for sleep and more probability for interruptions to natural behavioral patterns regarding the daily rhythm of rest.

5.4.1 Skew group characteristics

Having introduced the different metrics above to quantify sleep (Figure 5.3), I now explore whether they provide insights to characteristics of different skew groups. I set up a prediction task, where select individuals with either the most positive, neutral or negative skewed distribution of sleep duration (N=2000 for each group) into three separate groups. For each skew group, I randomly select 2000 other individuals and then try to predict whether individuals where originally selected to the group or not. For the prediction task I use features introduced on Figure 5.3 (except $R_{\text{onset \& offset}}$) as well as traditional epidemiological measures: chronotype, social jetlag and midsleep on weekdays. By considering the features importance of the prediction task, I can find what separates skew group from random selection of users and understand what characterises each group. Specifically, I train a decision tree classifier to predict whether an individual belongs to the skew group or not (baseline 50% accuracy), results are summarised in Figure 5.5.

Overall results of the prediction task Using all metrics introduced on Figure 5.3 (except $R_{\text{onset \& offset}}$) and classical sleep epidemiology measures (chronotype, midsleep on weekdays and social jetlag), I am able to predict positive and negative skew with substantially better accuracy (73-79 %) than the baseline (50 %), which is not the case for the neutral group (56-57 % accuracy). The results from the prediction task are summarised in Figure 5.5. The most important features in all cases is median sleep duration, but that makes sense due to the close connection of skew with sleep need. The distinct difference in typical sleep duration between skew groups can be observed for the distributions in Figure 5.5D-F. In fact, this feature is the primary explanation for the prediction of positive skew estimated for the distribution of sleep duration on weekends. Furthermore, those that belong to that skew group, also sleep substantially less than other groups on weekday nights (see Figure C.13 in Appendix C).

Positive skew The group is clearly characterised by shorter sleep duration (see distribution on Figure 5.5D-F). Furthermore, individuals with positively skewed distribution of sleep duration (estimated for all day-types or weekdays) have higher weekend-weekday median duration difference than others (see Figure 5.5G and K, and Figures C.11 and C.12 in Appendix C). Substantial weekend-weekday sleep duration difference is one characteristic of those who are sleep deprived on weekdays (due to constraints of the social clock), and use the weekends to 'catch-up' [184,44].

Negative skew The group is associated with overall longer nighttime sleep, lower weekend-weekday median sleep duration difference and narrower distribution of wake-time (see Figure 5.5G,K,H,L and N. In fact, the distributions on Figures C.11-C.13 in Appendix C, show that individuals with negative skewed distributions of sleep duration generally sustain more regular sleep behavior compared to the two other groups. All of these characteristics are associated with behaviors of steady sleepers who obtain sufficient amount of nighttime sleep.

Neutral skew The group is not characterised by anything specific, but rather falls between the negative and positive skew group in terms of all measured aspects of sleep (see Figure 5.5 and distributions in Figures C.11-C.13 in Appendix C).

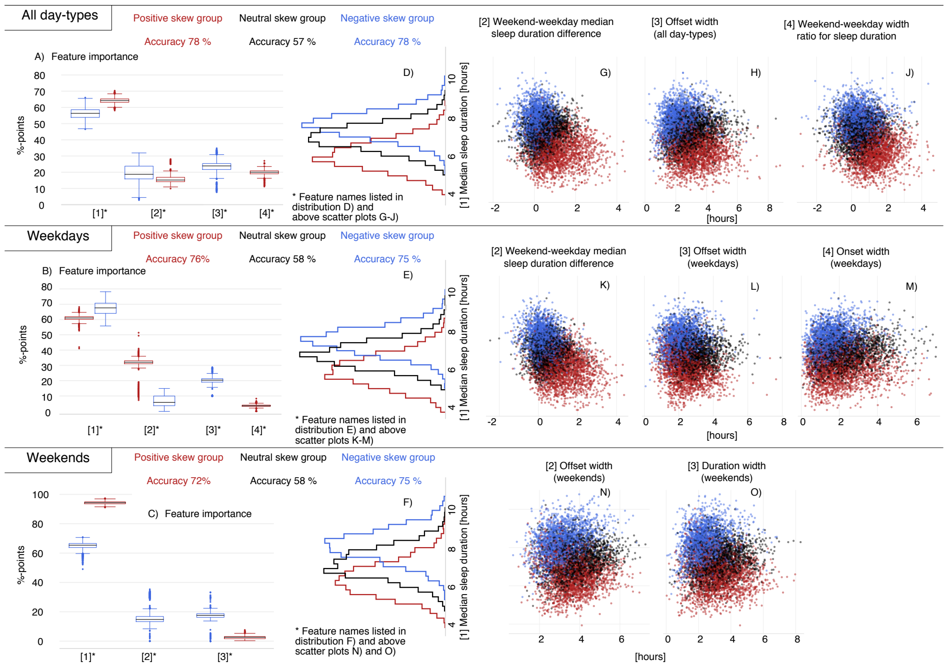


Figure 5.5: Information about skew groups from decision tree classifier
Decision tree classifier is trained to predict whether an individual belongs to skew group or not. There are three groups, and each is comprised of 2000 individuals with either the most positive, neutral and negative skewed distribution of sleep duration, as well as a random selection of 2000 other individuals. **A-C:** Feature importance obtained from decision tree classifier from prediction of positive and negative skew group, where skew is estimated from the distribution of sleep duration for all day-types on **A**, weekday-nights on **B** and weekend-nights on **C**.
D-F: The distribution of median sleep duration for the individuals (N=2000) with the most positive, neutral and negative skewed distribution of sleep duration for all day-types on **D**, weekday-nights on **E** and weekend-nights on **F**.
G-J: Scatter plot of median sleep duration (estimated for all day-types) with weekend-weekday sleep duration difference, offset width (all day-types) and weekend-weekday width ratio for median sleep duration where points are colored by skew group; red, black and blue for positive, neutral and negative skewed distribution of sleep duration respectively.
K-M: Scatter plot of median sleep duration (estimated for weekday nights) with weekend-weekday sleep duration difference, offset width (weekdays) and onset width (weekdays) where points are colored by skew group; red, black and blue for positive, neutral and negative skewed distribution of sleep duration respectively. **N-O:** Scatter plot of median sleep duration (estimated for weekend-nights) with offset width (weekends) and duration width (weekends) where points are colored by skew group; red, black and blue for positive, neutral and negative skewed distribution of sleep duration respectively.

Conclusion

I find high-resolution recordings of sleep measured over long period of time to produce complex and multifaceted patterns across the population. I introduced a new visualisation method, *the sleep portraits*, which illustrate characteristic patterns for individual sleep behavior. Furthermore, I propose a novel data-driven metric to quantify direction of preference which may be indicative of whether people's physiological sleep needs are met or not. The measure is simply the skewness of individual-level distribution of sleep duration and quantifies disproportionate tendency for long or short nights, relative to typical behavior. Currently there exist no methods to measure whether sleep needs are met, except for comparison to the recommended amount of healthy sleep for adults. Thus, skew may be the first metric to provide a way to infer whether sleep needs are met from estimations of high-resolution recordings of sleep duration. However, one of the limitation is that skew has not been validated with subjective nor qualitative estimates of sleep. To mitigate those limitation, I show that skew is: i) linearly dependent with median sleep duration across the population, ii) 5he skew group with an average estimate closest to zero attains typically 7-7.5 hours of sleep per night, which is within range of recommendation[196,197,198,199], iii) individuals who are not skewed but obtain overall long or short nighttime sleep, are likely to have demographic characteristic which are associated with either long or short sleep duration, iv) I find the group who has the most positive skewed individuals (disproportionate tendency for longer nights) to have higher weekend-weekday sleep offset and duration difference than other groups, which is a behavior linked to unhealthy sleep patterns[184,44,189,190,191,192], and v) to the contrary, individuals with negative skew (disproportionate tendency for shorter nights) exhibit more sleep regularity – a behavior associated with good sleep hygiene[31,246,35].

Nú hverfur sól við segulskaut
og signir geisli hæð og laut,
er aftanskinið hverfur hljótt,
það hefur boðið góða nótt.

Árni Þorsteinsson

6

Observed Effects of Biological, Societal and Cultural Differences on Sleep

In Chapter 1, I discussed how biologically based differences such as age and gender can effect sleep, as well as the environmental and contextual aspects of everyday life. Here I explore whether this is observed in the study sample by looking at the differences in sleep patterns between Western and Eastern geographic regions. I investigate whether influences of contextual aspects and behavioral choices, such as country-level policy and having a child, can be detected in the data. I analyse the development of sleep timing and duration across the lifespan and examine life-stage dependent gender differences.

6.1 Sociocultural Variation

I start the examination by observing average sleep onset and duration by age group (for men on weekdays) in Table 6.1 for the top five countries with the most users in. The country of residence seems to influence sleep behaviors, and there is a clear difference between geographic regions in the East (Asia) and West (Europe). I further explore this difference and put all countries into one of the three categories; `East` (Asia), `West` (Europe and N-America) and `Others`. On Figure 6.1A I illustrate the distribution of median sleep onset, offset and duration (by day-type) for users residing either in Eastern or Western geographic regions (there were too few users in the third category for comparison).

	Japan n=17 231 (∼ 25 %)	Germany n=7 140 (∼ 10 %)	Russia n=5 095 (∼ 7 %)	Taiwan n=5 028 (∼ 7 %)	United Kingdom n=3 900 (∼ 6 %)
Age groups	Average sleep onset (hh:mm)				
19-24	00:53	23:55	00:39	01:12	00:24
25-29	00:44	23:52	00:29	00:59	00:07
30-34	00:41	23:40	00:21	00:51	23:52
35-39	00:27	23:36	00:08	00:42	23:46
40-44	00:21	23:37	00:04	00:32	23:41
45-49	00:15	23:30	00:03	00:30	23:45
50-54	00:06	23:31	23:58	00:16	23:36
55-59	23:54	23:26	23:55	23:52	23:45
60-67	23:42	23:26	23:50	23:50	23:42
Age groups	Average sleep duration (hrs)				
19-24	6.6	7.3	7.0	6.7	7.3
25-29	6.4	7.1	7.0	6.7	7.3
30-34	6.4	7.1	7.0	6.6	7.2
35-39	6.3	7.0	7.0	6.4	7.1
40-44	6.3	6.9	7.0	6.5	7.1
45-49	6.2	7.0	7.0	6.4	7.0
50-54	6.2	7.0	7.0	6.6	7.0
55-59	6.3	6.9	7.0	6.6	7.0
60-67	6.4	7.2	7.1	6.5	7.1

Table 6.1: Development of sleep onset and duration by age split up by the top five countries with the most users in the data set

The differences observed in Table 6.1 become more evident on Figure 6.1. Users in the East sleep on average less than those in the West – both on weekdays (6.45 hrs in the East and 7.18 hrs in the West) and weekends (6.90 hrs in the East and 7.81 hrs in the West). On weekdays individuals in the East fall asleep on average later (00:20 East and 23:43 West) while both groups wake up at similar hours. On weekends this is reversed and both groups go to sleep at approximately the same time while people in the West wake up later (on average 07:47 East and 08:25 West). This is in accordance with findings from a recent large scale study exploring geographic disparities with data from wearable devices[46]. In terms of measures for weekend-weekday differences for sleep timing and duration (Figure 6.1B), I identify a slight distinction between Eastern and Western geographic regions. Residence in the West shift their bed marginally more than those in the East and catch overall more sleep on weekends, compared to weekdays, than Eastern residence. The distinction between the two geographic regions is most prominent when observing weekend-weekday difference for sleep offset, but those residing in the West shift wake time substantially more on weekends than those in the East.

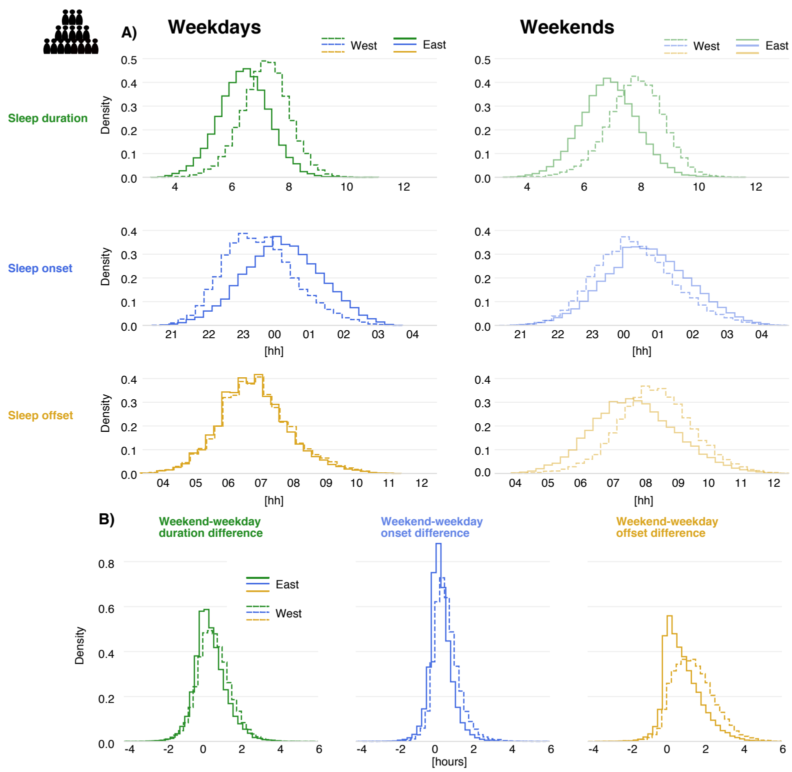


Figure 6.1: Contrast between sleep patterns for residence in Eastern and Western geographic regions
A: The distributions for median sleep onset, offset and duration separately by day-type and region of residence. East (Asia) is represented with normal lines and West (N-America & Europe) with dotted lines. Sleep duration, onset and offset are represented with green, blue and yellow colors respectively. **B:** Weekend-weekday differences for sleep onset, offset and duration for Eastern and Western geographic regions.

6.2 Influence of age and gender on sleep

I explore the effect of age and gender on sleep, in two separate analysis: i) Examining the development across age groups with classical epidemiological measures, chronotype and social jetlag, where I also compare those results with estimates from Roenneberg *et al.*. ii) Additionally, I explore how more relatable measures, such as sleep onset, offset and duration, develop with age and what are the differences between men and women.

6.2.1 Analysis with Classical sleep metrics: Chronotype & Social jetlag

Chronotype (MSF_{sc}) and social jetlag (SJ) are commonly employed in sleep research to quantify phase preference and misalignment[44,71]. I use Roenneberg *et al.* (2003) methodology to derive each individual's phase preference using the methods from the Munich Chronotype questionnaire (defined concretely in Chapter 2). Although generally I do not use these metrics for an in-depth analysis in my work, I facilitate comparison to previous findings and depict the development of these measures with age on Figure 6.2. Looking at the development of chronotype, sleep timing advances to earlier hours as people get older. Men are later chronotypes up until the middle part of the adult lifespan and after that point women have a later phase preference than men. Misalignment between weekend and weekday sleep timing, is most pronounced in the youngest groups of people. Social jetlag decreases over lifespan, but plateaus or slows down from age 35-49, and then there is a relatively rapid decline from age 50 to 67. Gender differences are negligible, but there is a clear separation between men and women from age 40-49 where women have higher social jetlag than men.

Comparison of social jetlag and chonotype to Roenneberg *et al.* study I compare the development of chronotype with age to results from Roenneberg *et al.* (2007), the estimates match pretty well although phase preference is more advanced for young men in Roenneberg's sample compared to mine[15]. However, when comparing social jetlag from my sample to Roenneberg's results I find large discrepancies. For example, average social jetlag for 20 year old men is estimated approximately 3 hours in Roenneberg's study, while it is about 50 minutes in my sample. Overall, I find considerably lower levels of social jetlag across all observed age groups compared with the values reported by Roenneberg *et al.* (2012)[44]. In Roenneberg's study, the sample consisted of questionnaire respondents from four European countries (Germany, Switzerland, the Netherlands and

Austria). The differences may reflect a mismatch between global and regional circadian preferences, recall biases linked to the questionnaire and/or other unobserved differences.

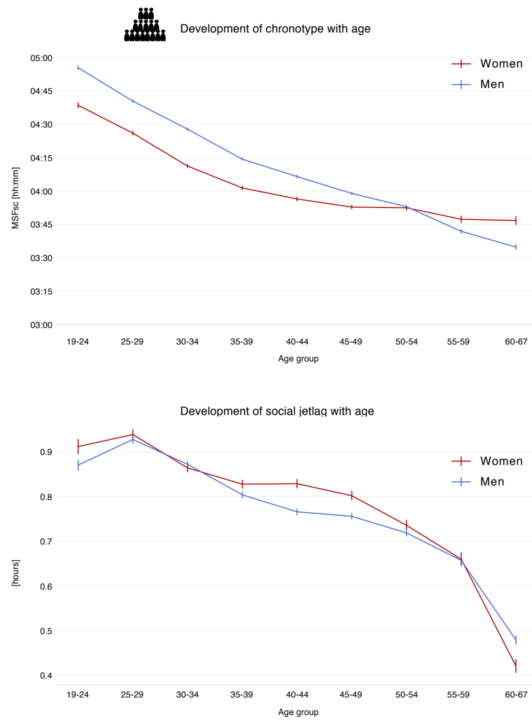


Figure 6.2: Development of Chronotype and Social jetlag with age
The upper panel illustrates development of chronotype with age, estimated for men and women separately. The lower panel shows development of social jetlag with age, estimated for men and women separately. All error bars are estimates of standard error of the mean (SEM).

Regional differences in Social jetlag I further explore regional differences by constraining my sample to only include the same four European countries as Roenneberg et al. and compare estimates to the full sample (see Figure 6.3A & B). The estimates of social jetlag are markedly higher for the four European countries from Roenneberg's sample. Furthermore, I compare social jetlag levels between regional strata of my sample from Asia and the four European countries from Roenneberg's study (Figure 6.3C). The figure suggests that social jetlag for young adults may be over twice as large in the same European region sampled by Roenneberg *et al.* compared to geographic regions in Asia, and ∼1.5 times larger for middle-aged and older adults[44]. I further explore regional differences by depicting estimates of social jetlag by age groups for users residing in three continents; Asia, Europe and N-America (see Figure 6.3D). Social jetlag is most pronounced in young Europeans (age 19-29), but overall there is similar behavior identified across the adult lifespan for Europeans and N-Americans. However, social jetlag measures are slightly higher in Europe compared to N-America, all the while Asia has substantially lower estimates than the other continents for all age groups.

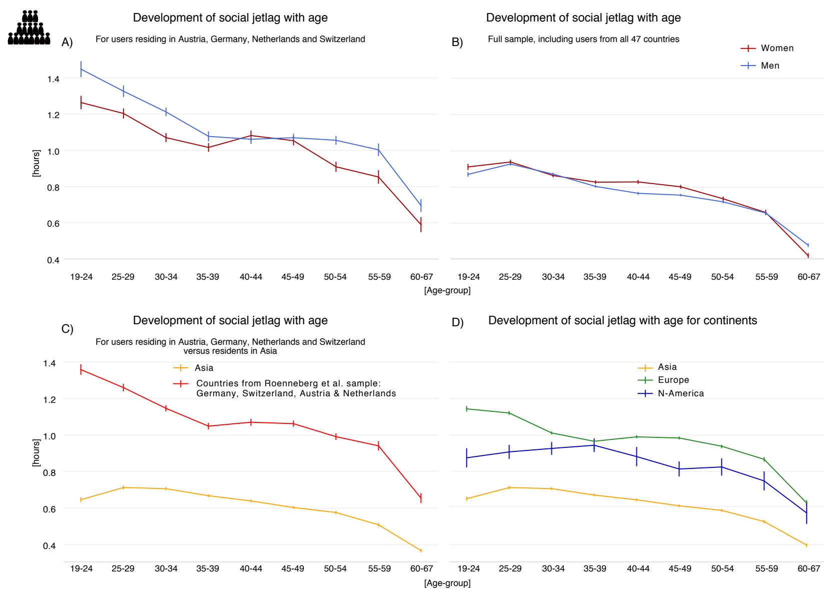


Figure 6.3: Regional differences of Social jetlag

A: Development of average social jetlag with SEM across age groups for the subset of my sample from the primary geographic regions represented in Roenneberg *et al.*'s (2012) questionnaire-based study. **B:** Development of average social jetlag with SEM across age groups for the for the full study sample. **C:** Development of average social jetlag with SEM across age groups for the subset of my sample from the primary geographic regions represented in Roenneberg *et al.*'s (2012) and countries in Asia from my sample as well. **D:** Development of average social jetlag with SEM across age groups, separately for residence in the three continents; Asia, Europe and N-America.

6.2.2 Development of sleep onset, offset and duration with age

In order to summarize the development of sleep onset, offset, and duration across the lifespan, I calculate each user's average value and then aggregate across the study sample by age, gender, and day type (weekday or weekend), illustrated on Figures 6.4-6.6. The relationship between these measures (sleep onset, offset and duration) with age is also explored with mixed effects models to confirm observed trends on Figures 6.4-6.6 with the raw data. Given the longitudinal and hierarchical structure of the data with repeated measurements within users, and users are then nested within their country of residence, observations are likely highly correlated on both levels (country and user). To account for this dependence I adopt a mixed effects modeling framework which controls for user and country-level variation while examining age-related trends in sleep patterns and assessing the influence of demographic factors. Figures 6.4-6.6 are re-used from my paper, *Gender differences in nighttime sleep patterns and variability across the adult lifespan: a global-scale wearables study*[26].

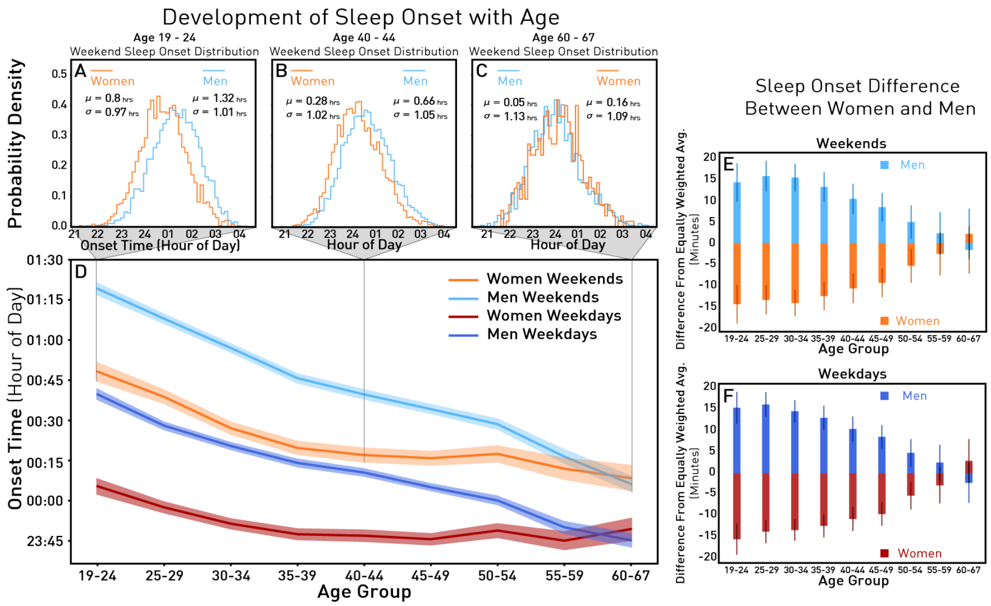


Figure 6.4: Development of sleep onset with age

Distributions for sleep onset on weekends split up by gender for different age groups: **A** age 19–24, **B** age 40–44, and **C** age 60–67. The development of average sleep onset by age group split up by gender and day type (weekend/weekday). The red/orange colors correspond to women, light/dark blue colors correspond to men, darker colors represent weekdays and lighter colors signify weekends. The colored envelopes display 95% CIs around each age group mean **D**. The equally weighted, between gender sleep onset difference by age group with 95% CI on weekends **E** and weekdays **F**.

Sleep onset Figure 6.4D illustrates how sleep onset advances to earlier hours with age, both for men and women. The difference between bed time on weekends and weekdays is roughly constant across the lifespan. There is a large distinction between bed-time for men and women across younger age groups or up until age 40-44, and after that the difference grows smaller with increasing age. Eventually, the statistical difference between the curves is eliminated and the two groups go to bed at approximately the same time from age 55-67. Even though the 95 % confidence intervals on Figure 6.4D are narrow, the underlying distributions are quite broad as exemplified on Figure 6.4A-C, where the distribution for average sleep onset on weekends is illustrated for age group 19-24, 40-44, and 60-67 respectively. I directly visualize the gender differences on Figure 6.4E (weekends) and F (weekdays), which display the difference of average sleep onset for men and women from the weighted average curve of sleep onset by gender. The gender gap in bed-time appears to persist until about 40 years of age, when the two curves begin to converge.

Figure 6.4D illustrate the *aggregated* raw data from the study sample but I confirm those results with mixed effects models. The model estimates an overall 29 ± 0.20 minute difference between sleep onset on weekends and weekdays for women and 28 ± 0.20 minute difference for men (age group 40–44). The model suggests an even steeper rate of decrease of sleep onset for men than the raw data exhibits (see Figure D.2 and Table D.1 in Appendix D). Consequently, the difference between men and women at age 40–44 on weekdays is estimated less by the model (15 ± 1.5 minutes 95% CI) compared to the raw data (24 ± 1.5 minutes 95% CI). Furthermore, the model estimates the onset curves for men and women to intersect slightly earlier or within age range 50–54.

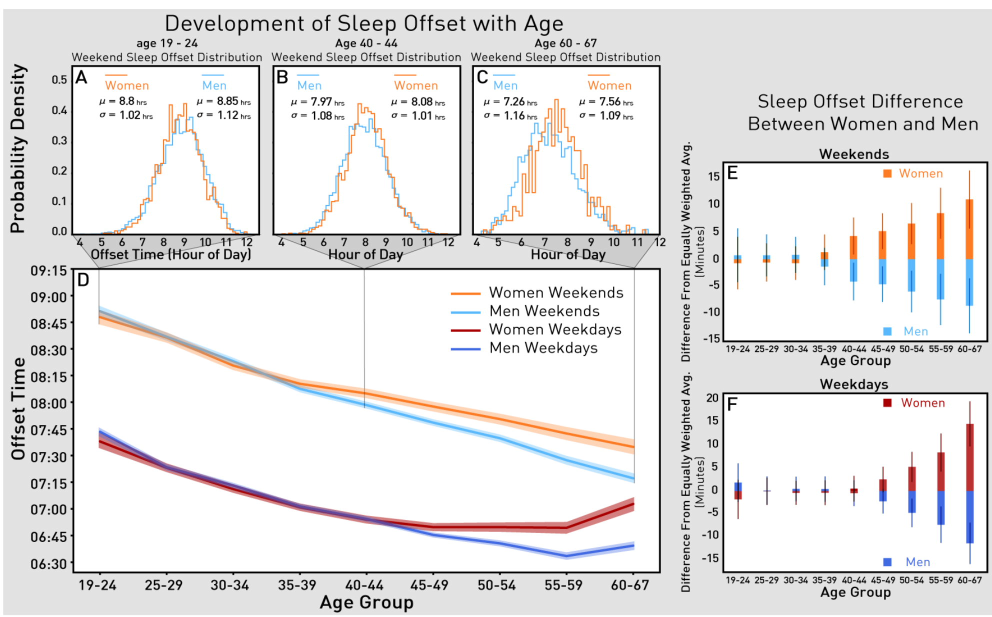


Figure 6.5: Development of sleep offset with age

Distributions for sleep offset on weekends split up by gender for different age groups: **A** age 19–24, **B** age 40–44, and **C** age 60–67. The development of average sleep offset by age group split up by gender and day type (weekend/weekday). The red/orange colors correspond to women, light/dark blue colors correspond to men, darker colors represent weekdays and lighter colors signify weekends. The colored envelopes display 95% CIs around each age group mean **D**. The equally weighted, between gender sleep offset difference by age group with 95% CI on weekends **E** and weekdays **F**.

Sleep offset Figure 6.5D shows that people tend to wake up earlier as they get older, and there is a substantial weekend-weekday difference in wake-time that mostly persist across the lifespan. Interestingly, the patterns identified concerning gender differences in the development of sleep onset with age are now reversed, and the two groups wake-up at the same time up-until the middle of the adult lifespan, and thereafter separate with men rising earlier than women. Thus, to summarise, from age 19 to 39 women and men exhibit an average tendency to go to bed at different times yet wake up at similar hours. The sleep offset curves for men and women diverge earlier on weekends (40-44) compared to weekdays (45-49). Gender differences can be studied more closely on Figures 6.5E (weekends) and F (weekdays) which illustrate the difference of sleep offset by gender and age group from the equally weighted average of the curves for men and women. As before, I illustrate that even though the 95 % confidence interval for average sleep offset is narrow, the underlying distribution can be broad and span a wide range of behavior (see distribution of sleep offset on weekends on Figure 6.5A-C for age group 19-24, 40-44 and 60-67 respectively). I compare the patterns observed for aggregated averages of sleep offset with age on Figure 6.5D with results from mixed effects model, and observe a close agreement between the two (see Table D.2 and Figure D.3 in Appendix D).

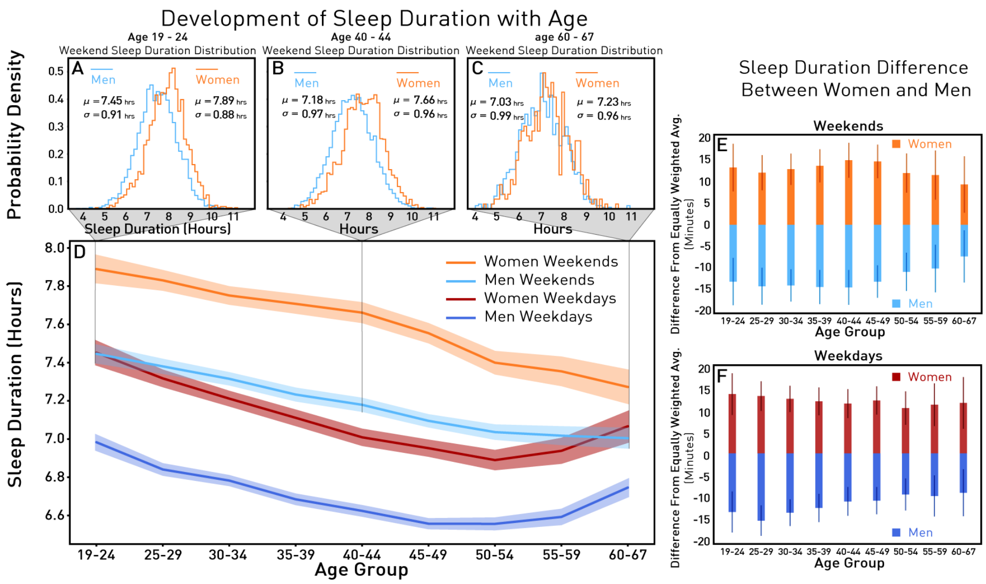


Figure 6.6: Development of sleep duration with age
Distributions for sleep duration on weekends split up by gender for different age groups: **A** age 19–24, **B** age 40–44, and **C** age 60–67. The development of average sleep duration by age group split up by gender and day type (weekend/weekday). The red/orange colors correspond to women, light/dark blue colors correspond to men, darker colors represent weekdays and lighter colors signify weekends. The colored envelopes display 95% CIs around each age group mean **D**. The equally weighted, between gender sleep duration difference by age group with 95% CI on weekends **E** and weekdays **F**.

Sleep duration Now we turn our attention to sleep duration where Figure 6.6D depicts the development of aggregated averages with age. Sleep duration generally decreases with increasing age, except on weekdays from age 55-67 where there is a slight increase. The weekend-weekday contrast nearly persist for all age group, although the differences grow smaller from age 50-54 for women, and 55-59 for men. Women sleep on average always more than men, across the entire lifespan. This is highlighted on Figure 6.6E (weekends) and F(weekdays) which show the difference of sleep duration by gender and age group from the equally weighted average of the curves for men and women. The mixed effects model for sleep duration generally confirms the trends observed for the aggregated raw data visible in Figure 6.6D (see Table D.3 and Figure D.4 in Appendix D). The most prominent discrepancy is the different rate of change in sleep duration with age (between aggregated averages on Figure 6.6 and mixed effects model) resulting in less prominent gender differences than observed on Figure 6.6D. Consequently, the curves for men and women come close to overlap from age 55 to 67 (see Figure D.4 in Appendix D). The aggregated raw data estimates women at age 40–44 to sleep 23 ± 1.7 minutes longer than men, whereas the model estimates the difference to be 11 ± 1.0 (95 % CI).

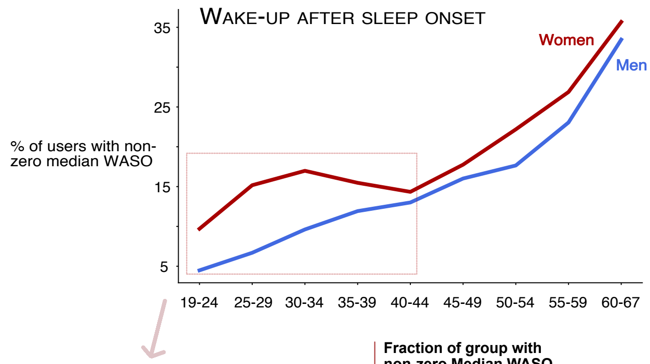
6.3 Can policy and gender roles effect sleep?

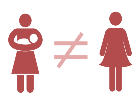
Societal structure such law, policy, culture and other aspects, can have huge effects on human life. Thus far I have identified distinct regional differences in sleep patterns most probably rooted in cultural differences, as well as life-stage dependent gender differences. Now I explicitly examine whether laws regarding retirement and gender roles can effect sleep.

Effective retirement age Here I explore whether law and policies can effect age-related changes in sleep by examining whether I am able to detect signals of effective retirement in the study sample. I use mixed effects model where measures are nested within user (random effect) and age, gender and country are fixed effects, with a three-way interaction term. I conduct the analysis for three of the countries with the most data in our sample; Japan, Germany and the United Kingdom, which are chosen because of the following reasons: i) to make sure each country contained a sufficient number of users and amount data, ii) to ensure there would be at least one country from Asia and one from Europe (to explore regional differences) and iii) to be able to reference accessible, official records on effective retirement age for each country. Figures D.5-D.7 in Appendix D illustrate the estimates of sleep onset, offset and duration with mixed effects model, and Tables D.4 and D.5 list out exact numbers of predicted outcomes for selected ages, separately for men and women, residing in Germany and Japan. A study exploring changes in sleep duration and timing during retirement found the following results, and I quote: "Transitioning to retirement is associated with longer sleep duration, later bedtimes, and later wake times. These changes were detectable about 1 y post-work transition and were persistent up to 3 y later"[247]. I use this example to assess when retirement occurs in the study sample. Specifically, I require increase in sleep duration and delay in both sleep onset and offset to proxy for possible changes in work demand. From the exact estimates in Table D.4, referring to individuals living in Germany, I observe increase in sleep duration, and bed and wake-time become later after 55 years of age for women and 60 years of age for men in Germany. The increments are small and changes occur smoothly towards the age threshold, 67. Thus, decreased workload and more flexible work-schedules may begin to occur sometime after 55 years of age for women and 60 for men living in Germany. OECD reports effective retirement age in Germany at 64 (estimated from numbers from 2013-2018), thus a relatively good match with the study sample[248]. Similarly, Table D.5 provides exact estimates of sleep onset, offset and duration for men and women living in Japan (with mixed effects model). I identify a slight increase in sleep duration for women aged 58 and men at 65, but only a decrease or plateau for sleep onset and offset. Thus, transition toward flexible work schedules or retirement is not apparent in these sleep outcomes before 67 years of age in Japan. This aligns well with official records that indicate the effective retirement age is 71 for men and 67 for women residing in Japan[248].

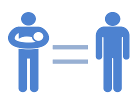
Gender Roles & Parenting Secondly, I explore gender inequality in sleep quality by using information about parent mobile application usage to explore life-stage differences in nighttime awakenings. Sleep was recorded in 1-min epochs, thus only wake time after sleep onset (WASO) greater than 60 s were registered by the wearable devices. The percentage of users with nonzero median WASO was plotted by age group and gender in Figure 6.7 (taken from my paper *Gender differences in nighttime sleep patterns and variability across the adult lifespan: a global-scale wearables study*[26]). There is a clear separation between men and women in terms of nighttime awakenings around age 19-39 (see Figure 6.7), and incidentally, these are the years in which people are most likely to procreate[249]. I investigated the hypothesis that increased prevalence of nighttime awakenings among women during early adulthood could be associated to infant-rearing, young children exhibit irregular sleep patterns for the first two years of their life[250]. I explicitly analyse nighttime awakening for age group 19–39 or where the curves for men and women in Figure 6.7 diverge. To infer parental status and infant-rearing, I use information about aggregated mobile phone app usage. Specifically, I anonymously identify users as probable parents if they have installed and use apps intended for parents with young children, which I call 'parent apps' (see further details about parent app classification section D.4 in Appendix D).

Approximately 13.5 % of women aged 19-39 without parent app usage have a non-zero median WASO while that prevalence is 27.1 % for the age matched group with parent apps installed (absolute difference of 13.6 %). However, I find 8.6 % of men (aged 19-39) without parent app usage to have non-zero median WASO while that prevalence is 11.1 % for men with parent apps in use, thus the absolute difference is 2.5 % and considerably lower than the 13.6 % difference between the female groups.





	Fraction of group with non-zero Median WASO
Women with parent app installed	72.9 %
Women without parent app installed	86.5 %



	Fraction of group with non-zero Median WASO
Men with parent app installed	88.9 %
Men without parent app installed	91.4 %

Figure 6.7: Gender inequality in sleep quality
The plot illustrates percentage of people with non-zero median WASO, by age group and gender. The red color corresponds to women and blue color to men. The table below illustrates the percentage of users with non-zero median WASO within groups of users that do, or do not have parent apps in use, separately for men and women.

Furthermore, I find women (age 19–39) with parent app installed on their devices to have a significantly different distribution of median (denoted M) WASO than age-matched women without these application in use (estimated with two sample Kolmogorov–Smirnov (KS) statistics, $p = 9.66x10^{-21}$), where $\mathrm{M}_{\mathrm{WASO\ women\ w.\ parent\ app}} = 184$ s and $\mathrm{M}_{\mathrm{WASO\ women\ wo.\ parent\ app}} = 65$ s. By comparison, the distribution of median WASO for young adult men with parent app installed does not differ statistically (estimated with two sample KS statistics $p = 0.228$, where $\mathrm{M}_{\mathrm{WASO\ men\ w.\ parent\ app}} = 52$ s and $\mathrm{M}_{\mathrm{WASO\ men\ wo.\ parent\ app}} = 37$ s. Similarly, I do the same comparison but only for users with non-zero median WASO ($M_{WASO} \neq 0$). I find women (aged 19–39) with a proxied parental status to have a significantly different distribution of mean values (denoted μ) for their WASO than similarly aged women without parent apps (estimated with two sample KS statistics, $p = 1.38x10^{-14}$) where $\mu_{\text{WASO women w. parent app}} = 1105$ s and $\mu_{\text{WASO women wo. parent app}} = 874$ s. In contrast, when the same comparison is carried out for men, I find that their distributions not to differ significantly (estimated with two-sample KS statistics, $p = 0.207$) where $\mu_{\text{WASO men w. parent app}} = 905$ s and $\mu_{\text{WASO men wo. parent app}} = 867$ s.

Conclusion

I find distinct regional differences between sleep patterns within Eastern (Asia) and Western (N-America & Europe) countries. Those living in the East go on average later to bed on weekdays, while both groups wake up at the same time. The pattern is reversed on weekends, where both groups seem to fall asleep the same time but residence in Western countries wake up later. These differences in bed and wake-time result in longer average sleep duration in the West, both on weekends and weekdays. This is in accordance with findings from a recent large scale study exploring geographic disparities with data from wearable devices[46].

I study how sleep timing and duration develops with age, but the human phase of preference has been shown to advance to earlier hours with age[15]. I confirm those findings but also expand on them, by documenting the underlying dynamics between sleep onset and offset across age groups and genders. Men tend to have a later sleep onset than women up until 50–54 years of age, while both groups wake-up at similar hours up until age 35-39 but thereafter, women tend to wake up later. The overlap in wake-time between men and women up until middle adulthood may be due to aligned external demands such as attending university, work, raising young children and more. By exploring sleep start and end, rather than just midsleep as commonly used in epidemiological sleep studies, I capture these gender differences which have not been reported before.

I compare my estimates of classical epidemiological measures to other large scale studies, and find considerably lower levels of social jetlag in my sample across all observed age groups compared with the values reported by Roenneberg *et al.*[44]. I find the discrepancies partially due to mismatch between global and regional circadian preferences, but also likely recall biases linked to self-reports via questionnaire and/or other unobserved differences. This highlights the importance of accounting for underlying country-level differences when studying sleep misalignment to prevent biased global estimates.

I perform an exploratory analysis of the possible effect of retirement in age-related sleep changes. Laws and policy regarding retirement varies by country, thus I explore effective retirement in selected countries, and demonstrates that different regional policies appear to affect people's sleep patterns. Furthermore, I identify gender inequality in sleep quality during early adulthood probably driven by infant-rearing, which other studies have found as well[121].

Það er margt sem myrkrið veit,
minn er hugur þungur.
Oft ég svarta sandinn leit
svíða grænan engireit.
Í jöklinum hljóða dauðadjúpar sprungur.

Jóhann Sigurjónsson

7

Travel & Sleep

Interruptions to everyday life can potentially disrupt the human daily rhythm of rest and activity for a period of time. Here I explore how travel and new resting environment can effect sleep. Travel has increased dramatically over the past two decades, with the number of air-travelers nearly tripling but there are good reasons to believe that travel has negative impact on sleep[251]. The *First night effect* (FNE) is characterised by difficulties with falling asleep and prolonged sleep-onset latency. It is found to take place on the first night of sleep in new environment and is a consequence of a single brain-hemisphere displaying elevated alertness in unfamiliar surroundings[125,126,127]. Furthermore, the journey to the destination can also have negative impact on sleep but *travel fatigue* and *jet-lag* are conditions which can cause sleep complications[128,252,253,129,130,131,132]. Lastly, I note that all Figures in this chapter are reused from `paper 2`.

7.1 The change in sleep during travel depends on typical behavior at home

The data In order to classify nights as `home/travel` I transform raw GPS to *stop-locations* using the `infostop` algorithm[254] (with recorded start and end time). A *sleep location* is defined as the stop location with start-time closest to the sleep onset, where the user does not leave the location until after the sleep has ended. The location where most nights take place is defined to be user's *home location*, but I only include users if their percentage of nights-at-home is higher than 70%. From now on, I refer to nights that take place at least 20 km away from home as *travel-nights*. The final data used for analyses consists of 2.4 million weekday nights (6.0% away from home) from about 19 300 users and 0.8 million weekend nights (9.3% away from home) from 13 300 users.

How to measure the change in sleep due to travel Figure 7.1A illustrates an example nights recorded for a single user. To quantify the change in sleep due travel relative to at home behavior, I define a new variable $\Delta_s = \mu_s - M_{\text{home}}$ where s $\in \{\text{home, travel}\}$, M denotes the median and μ the average. The reason why I estimate Δ_s for home nights as well, is due to skewness arising from individual level distribution of sleep duration (introduced in Chapter 5), but this will be explained in more details below.

The linear dependence of Δ_{TRAVEL} and median sleep duration To explore the relationship between the change in sleep duration due to travel, and typical nighttime sleep at home I start out with examining the distribution of Δ_s where s $\in \{\text{home, travel}\}$ for users with different median sleep duration. In fact, I define *sleep groups* by rounding individual median sleep duration to the nearest half hour bin. Figure 7.1B illustrates that Δ_{travel} is positive for individuals who obtain short nighttime sleep at home, but quantity decreases as median sleep duration increases. This is further explored on Figue 7.1C where the average Δ_{travel} is estimated for each sleep group and illustrated with the standard error of the mean (SEM). From this plot, it is evident the the change in sleep duration due to travel depends linearly on median sleep duration at home. Thus, those who habitually attain short nighttime sleep at home (<6.5 hrs) tend to catch more sleep when nights take place away from home while those who typically sleep a lot at home (>8 hrs) are likely to sleep less during travel-nights.

The baseline effect for home nights arising from skewed distribution In Figures 7.1B & C, I also plot $\Delta_{home} = \mu_{home} - M_{home}$ (blue color). This is to illustrate there is a baseline effect, which relates to the observed systematic change in sleep duration away from home. The baseline effect refers to decreasing linear trend of Δ_{home} (blue line in Figure 7.1C), which shows that there is a systematic difference between mean and median sleep duration for home nights. In chapter 5 I presented the skew of individual level distribution of sleep duration, and showed that it was linearly dependent on median sleep duration. I believe this pattern arises because of *sleep-wake homeostasis*, a physiological process which regulates sleep pressure[13]. To conclude, there is a baseline effect for the measure of Δ_s for home nights, and to obtain the absolute effect of travel on sleep, I compare Δ_{travel} with Δ_{home}.

Robust of results despite imbalanced sample sizes of travel and home nights I directly compare Δ_{travel} and Δ_{home} by plotting the distributions in Figure 7.1D. There is a clear distinction between these two distributions, where Δ_{travel} is much broader. To rule out that these huge quantitative differences are occurring due to imbalanced sample size of travel and home nights, I

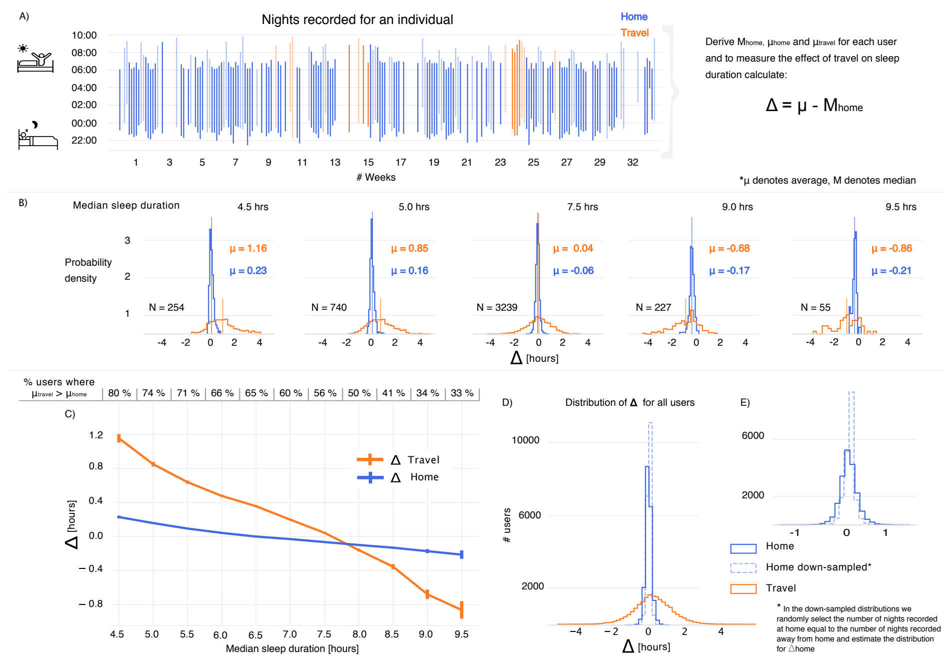


Figure 7.1: Sleep activity patterns and the relative change in sleep duration for travel-nights
A: For every individual we measure sleep onset, offset and duration for each recorded night. From these records we derive three measurements; median sleep duration at home ($\mathcal{M}_{\text{home}}$) and average sleep duration at home and away from home (μ_{home} & μ_{travel}). To measure change in sleep duration due to travel, relative to typical behaviour at home, we derive a new measure $\Delta_{\text{travel}} = \mu_{\text{travel}} - \mathcal{M}_{\text{home}}$. **B:** The distributions of Δ_{travel} (orange) and Δ_{home} (blue) for groups with different median sleep duration where users are grouped together by rounding their median to the nearest half-hour bin (referred to as *sleep groups*). **C:** The average Δ_{travel} for all sleep groups (median duration ranging from 4.5 - 9.5 hours) with the standard error of the mean (SEM). **D:** The distribution for Δ_{home}, $\Delta_{\text{home DS}}$ and Δ_{travel} for all users. **E:** A larger visual representation (more narrow range of the x-axis) for the distribution for Δ_{home} and $\Delta_{\text{home DS}}$ from panel D.

perform a down-sampling of home nights from individual's distribution of sleep duration, where the number of home nights are sampled to be as many as travel nights. The down-sampled distribution ($\Delta_{\text{home DS}}$) is illustrated in Figure 7.1D and with a narrower x-axis in Figure 7.1E. The down-sampled distribution is in fact, slightly narrower than Δ_{home} and remains quite different from the broad range of behaviour observed for the distribution of Δ_{travel} (for a more details see section E.1 in Appendix E). The illustrations in Figure 7.1 only shows behaviour on weekdays since I follow the convention of sleep research and analyse weekdays and weekends separately. In next section, I include data from weekends to understand the effect of travel on weekend nights as well.

7.2 Effects of travel on weekend-nights & the disproportional impact on misaligned individuals

Figure 7.2A depicts how social jetlag distributes across the study sample where 80 % of the users have social jetlag ranging from 9-98 minutes. Social jetlag was introduced in Chapter 2, but it quantifies the difference between weekend and weekday sleep timing and was designed to measure misalignment between the biological and social clock [184]. Figure 7.2B shows that social jetlag depends on sleep duration and individuals with high social jetlag typically sleep little on weekdays (4-5 hours) but a lot on weekends (9-10 hours). Figure 7.2B illustrates the distribution of Δ_{travel} for individuals with different range of social jetlag (defined by percentiles from the overall distribution). It is clear that the higher the social jetlag, the more sleep is gained during travel nights on weekdays, and to the contrary – more sleep is lost with increasing social jetlag for weekend travel nights.

Next I explore how travel effects sleep for weekend-nights. Figure 7.2D depicts the distribution of Δ_{travel} for each sleep group on weekdays (dark orange color) and weekends (lighter orange color), and the dotted horizontal lines mark the distribution quartiles. Overall the same pattern is identified for weekend and weekday travel-nights; the change in sleep duration due travel decreases as median sleep duration increases. However, the relative change is slightly larger in the positive direction (line pushed further up on y-axis) for weekday nights, observed when comparing the distribution quartiles in Figure 7.2D and the average Δ_{travel} by sleep groups on Figure 7.2E. These differences can be explained by the fact that people are usually more constrained by time and alarm clocks on weekdays, consequently sleeping less than they might need and therefore more susceptible to gain sleep - the opposite is expected for weekends; more room to lose sleep [184,44].

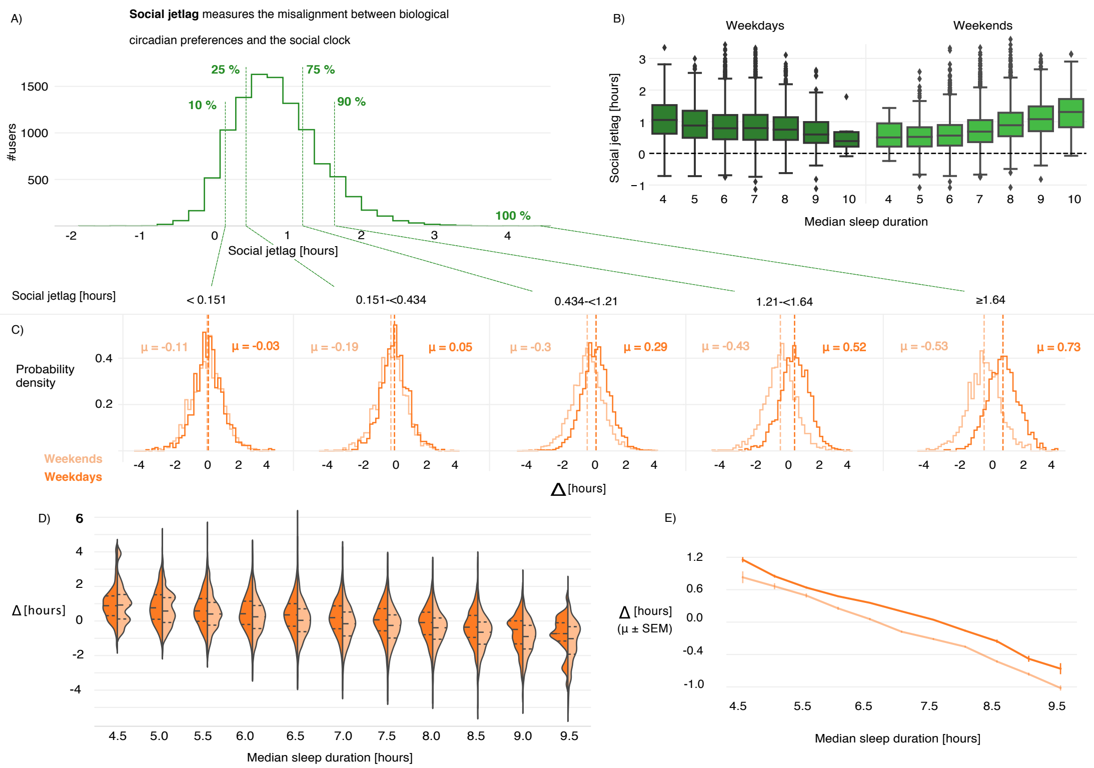


Figure 7.2: Disproportionate effect of travel on individuals with high social jetlag and the connection between weekends and weekdays

A: Defines social jetlag and visualises how it distributes across the user sample. **B:** Box-plots of social jetlag for users with different range of sleep duration. The horizontal lines represents the median, box limits correspond to upper and lower quartiles, whiskers define the 1.5x interquartile range and points are outliers. **C:** Distribution for Δ_{travel} on weekends and weekdays for groups of users with different range of social jetlag (defined by percentiles) **D:** The distributions of Δ_{travel} for sleep groups (half-hour bins for median sleep duration) by day type – weekends (lighter orange color), weekdays (darker orange color) and the dotted lines mark the quartiles of the distributions. **E:** The average Δ_{travel} plotted with the standard error of the mean (SEM) by sleep groups on weekdays (dark orange color) and weekends (light orange color)

7.3 The change in sleep timing for travel-nights

Thus far, I have observed a systematic change in sleep duration due to travel, but one might wonder how that translates over to changes in sleep timing. I explore the effect of travel on bed and wake time in similar manner as I did for sleep duration, where I calculate $\Delta_{\text{onset travel}} = \mu_{\text{onset travel}} -$

$M_{\text{onset home}}$ and $\Delta_{\text{offset travel}}$. These quantities are then aggregated into averages by user groups, defined by percentiles (10th, 25th, 50th, 75th, & 90th) of the distribution of median sleep duration (see Figure 7.3A1 for weekdays and A2 for weekends) and illustrated with SEM in Figure 7.3B1 for weekdays and B2 for weekends.

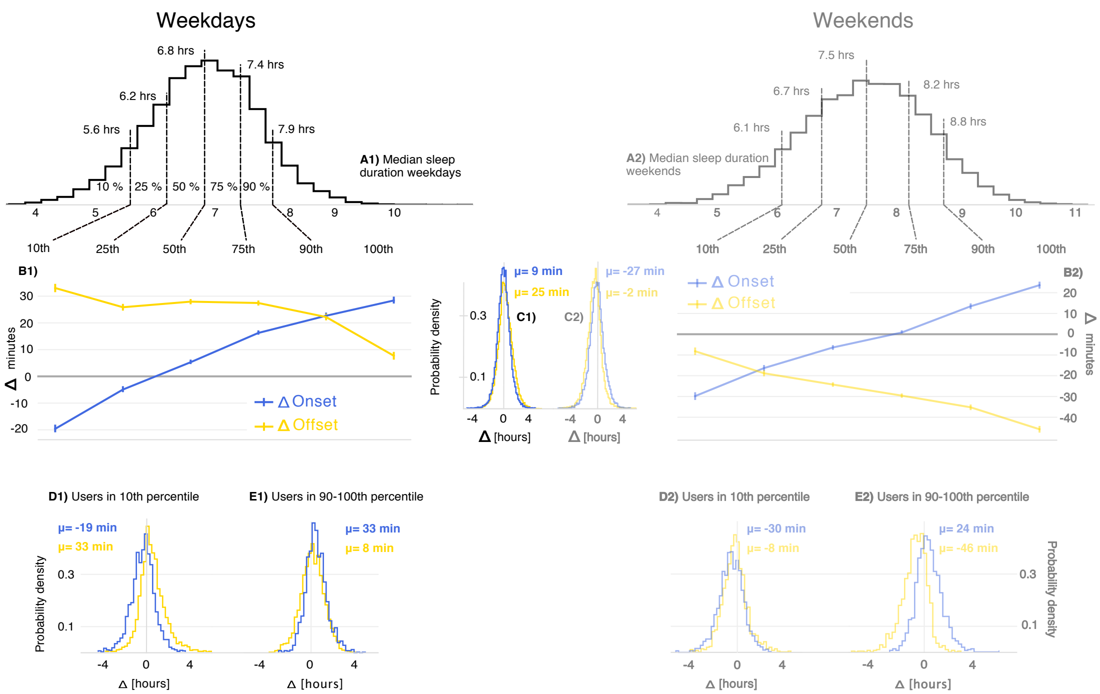


Figure 7.3: Change in sleep onset and offset for travel-nights

A.1 & A.2: The distributions of sleep duration on weekdays (A.1) and weekends (B.1) with the 10th, 25th, 50th, 75th and 90th percentiles marked with dotted lines. **B1 & B2**: The change in sleep onset and offset relative to typical home-sleep ($\Delta_{\text{onset travel}}$ and $\Delta_{\text{offset travel}}$) aggregated into averages (with SEM) by user groups defined by percentiles of sleep duration illustrated on (A1 & A2). **C1 & C2**: The distributions of $\Delta_{\text{onset travel}}$ and $\Delta_{\text{offset travel}}$ for all users in the sample. **D1 & D2**: The distributions of $\Delta_{\text{onset travel}}$ and $\Delta_{\text{offset travel}}$ for users with the lowest sleep duration on weekdays (D1) and weekends (D2) (bottom 10th percentile) **E1 & E2**: The distributions of $\Delta_{\text{onset travel}}$ and $\Delta_{\text{offset travel}}$ for users with the highest sleep duration on weekdays (E1) and weekends (E2) (90-100th percentile)

The relative change in sleep onset due to travel Figure 7.3B1 and B2 reveal that $\Delta_{\text{onset travel}}$ seems to have a linear relationship with typical sleep duration at home and bedtime advances from later hours (relative to typical behaviour at home) as typical home-sleep duration increases. Those who sleep less than 6.2 hours on weekdays (bottom 25th percentile) go to bed earlier on weekday travel-nights and those sleeping 7.5 hours or less (bottom 50th percentile) advance their bed-time to earlier hours on weekends when nights take place away form home.

THE RELATIVE CHANGE IN SLEEP OFFSET DUE TO TRAVEL Wake time during travel tends to be later for all groups on weekdays but earlier on weekends (see yellow curves in Figure 7.3 B1 and B2). The users that change wake-time marginally the most (relative to typical at-home habits), are users in the bottom 10th percentile on weekdays (for the distribution of sleep duration) and top 10th percentile on weekends, waking up 33 ± 2 minutes later on weekdays and 46 ± 2 minutes earlier on weekends during travel-nights. These are the groups most likely to include individuals with high social jetlag (because of their rank in the distribution of sleep duration), which I showed before are disproportionately effected by travel compared to those with low magnitudes of social jetlag (see section 7.2) The top 10th percentile on weekdays and bottom 10th percentile on weekends change their behaviour the least (shift of 8 ± 2 minutes in wake-up time). The middle groups, 10-90th percentile in the distribution of median sleep duration, exhibit more homogeneity on weekdays where the relative change wake-time on weekdays is 22-28 minutes later, whereas the range is broader on weekends and I observe a slight linear dependence with typical sleep duration at home.

The different patterns observed in the change of sleep timing due to travel on weekends and weekdays can be explained by the constraints that the social clock induces. Sleep has a tendency to be occurring at earlier hours than is desired on weekdays due to morning work schedule, while likely occurring at hours closer to biological preferences on weekends [184,44]. Thus, there is more room to shift sleep timing to later hours on weekdays and flexibility to advance bed and wake time to earlier hours on weekends. That is exactly what I observe in Figure 7.3; bed and wake-up times is shifted to later hours for nearly all groups on weekdays, and on the contrary to earlier hours on weekends.

7.3.1 Effect of Demographics & Robustness of results

In the analysis above, I did not control for any heterogeneity in terms of age, gender and country of residence. In order to explore whether the change in sleep due travel depends on demographic covariates, I employed a mixed effects model. Moreover, I do so to validate the patterns identified in Figure 7.1-7.3 with the aggregated raw data. Overall the results were confirmed and the most influential covariate was region of residence (East/West). An important parameter in the analysis is how many nights of travel-nights user must have to be included in the data-set ($N_{\text{travel nights}} = 2$. To confirm the robustness of the results, I explore whether they depend on the minimum number of travel days required per user. Estimates of fixed effects are examined, while the inclusion criteria changes, ranging from minimum 2 to 12 travels days per user. Overall the same results are found when number of travel days required per user is increased, with some indications of a slight change in magnitude (all of these results can be found in `Paper 2`).

Conclusion

I observe a systematic change in sleep duration for travel-nights, relative to typical at-home behaviour. The change due to travel depends linearly on typical sleep quantity at home and decreases as median sleep duration increases. Individuals are inclined to gain sleep during weekday travel-nights, but rather lose sleep during weekend-nights. That can likely be contributed to constraints of the social clock resulting in overall less sleep duration on weekdays at home, and longer nighttime sleep on weekends[44]. That is further supported by the fact that misaligned individuals (individuals who have high social jetlag) are disproportionately effected by travel[184,44]. Wake time during travel is on average advanced to later hours on weekdays compared to typical nights at home, but to earlier hours on weekends. The change in bed time for travel nights is linearly dependent on typical sleep duration at home, and is advanced to later hours as median sleep duration increases. That emphasizes the fact that wake time is a more controllable factor than bed-time when it comes to sleep, since alarms can wake people up at specific hours but cannot impose sleep onset at a predefined point in time. My previous work indicated the same results, where individuals seemed to catch longer nighttime sleep on weekends by shifting their bed-time marginally more than wake time[26].

Generally, travel was believed to have deleterious effects on sleep, but the analysis above reveals it has a more complex impact[128,252,130,132,131,253,129]. The main finding is that sleep during travel tends to have a balancing effect, where those who generally obtain short nighttime sleep at home, sleep more than at home when travelling, while individuals whose overall nighttime is characterized by long duration, tend to sleep less when nights take place away from home.

We shall not cease from exploration

And the end of all our exploring

Will be to arrive where we started

And know the place for the first time.

T.S. Elliot

8 Discussion

Drawing on global and large-scale sleep activity data-set, I find the daily reoccurring state of night-time sleep to produce complex and multidimensional patterns across the population. Research within the field of sleep epidemiology has been limited to data with self-reported estimates of sleep, but today sleep recording technology has been revolutionised. Wearable devices can easily be employed to obtain high-resolution sleep recordings over long period of time. Here I explore whether there are any salient sleep patterns missed, when traditional sleep epidemiological metrics are used to analyse high-resolution sleep recordings. In consequence, I propose new quantitative metrics and methods to study data-sets with multi-night recordings of sleep. Furthermore, I seek to validate and expand upon fundamental knowledge regarding age-related changes for sleep timing and duration, as well as explore life-stage dependent gender differences. I briefly explore how cultural context, gender roles and regional law can influence sleep. Lastly, I investigate how travel and new resting environment effects sleep and present results that have not been documented before.

THE NEED FOR NEW METRICS? I explore behaviors among groups of users with similar estimates of traditional sleep epidemiology measures, such as habitual sleep duration, chronotype and social jetlag[71,184,53]. I find that these users groups can be comprised of behaviors that differ largely with respect to multiple characteristics of sleep. I conclude that these metrics are valuable, but only measure one aspect of the multi-dimensional process of sleep. I present alternative set features to quantify different aspects of sleep, which are partially built upon previous research where I considered **i)** the importance of analysing sleep separately for weekend and weekday nights, **ii)** the misalignment that the social clock might induce and **iii)** regarded qualitative information that sleep regularity entails[44,31,35].

COMPLEX PATTERNS OF SLEEP I find high-resolution recordings of sleep from thousands of users from all over the world, to manifest complex and multidimensional patterns patterns. In chapter 1, I summaries findings from multiple studies, and conclude that sleep is sensitive, or reactive to contextual, environmental and behavioral factors. This reactivity is reasonable in evolutionary context, but humans are in a vulnerable state while sleeping, thus we likely evolved this reactivity to avoid any harmful situation. The theory of 'social acceleration' argues that modern day human life is moving *faster*, our existence are cluttered with more information than ever and time is a scarce commodity[255,256,257,258,259,260]. The compound of the reactivity of sleep to its surroundings and high paced modern life, is likely inducing these complex patterns of sleep.

Empirical measure for direction of preference I present a novel data-driven metric, *skew*, which quantifies individual direction of preference and is proposed to measure whether physiological sleep need is met or not. This measure is simply skewness of individual distribution of sleep duration, and quantifies disproportional tendency for long or short nights, relative to typical behavior. The only existing way to quantify whether an individual is obtaining sufficient amount of sleep is by comparing their habitual sleep duration to the recommended amount (7-9 hrs for adults)[196,197,198,199]. These recommendation are inferred from epidemiological studies by examining health outcomes for different range of self-reported habitual sleep duration. Recently, the US National Sleep Foundation conducted a scientifically rigorous review on the matter with an expert committee, which concluded that medium sleep duration from 4 up to 7 hours may be appropriate for some people[197]. Guidelines regarding healthy amount of sleep should be held to high standards, but arguably may not apply to all and the metric of skew may be the first to identify individuals with sleep needs outside the range of 7-9 hrs.

Development of sleep duration with age & life-stage dependent gender differences I find average sleep duration to decrease over the lifespan and women sleep on average more than men for all age groups. These dynamics have been documented before and believed to have both biological and social basis[44,46,261,247]. I find the difference in average sleep duration between men and women the largest during young to middle adulthood. That observation coincides with part of the lifespan when sleep interruptions are considerably more common among women than men, likely due to the different burden of infant-rearing. Thus, imbalanced care giving demands might contribute to the gender differences in sleep duration during young to middle adulthood.

Development of sleep timing with age & life-stage dependent gender differences Previous epidemiological studies have demonstrated that the phase preference advances to earlier hours with age, and men on average are later chronotypes than women up until 40–50 years of age. After that point, men's chronotype overlap or become earlier than women's [15]. My exploration of changes in sleep timing over the lifespan confirm these findings but also expand upon them, by documenting underlying dynamics between sleep onset and offset across age groups for men and women separately. I find that men tend to have a later bed-time than women up until 50–54 years of age, while up until the age range of 35–39 there is no significant difference in wake time. Thereafter or from middle to late adulthood, women tend to wake up earlier than men. The inversion in gender differences in bed and wake-time may be indicative of gender-gaps in both domestic and labor demands.

Observed effects of cultural context & regional policy on sleep Today there are documented evidence regarding large disparities in sleep patterns across cultures, with the most prominent contrast between Eastern (Asia) and Western (Europe and North America) geographic regions [30,86,46,89]. I confirm these findings with a brief exploration where residence in the East overall sleep on less than those in the West. On weekdays the difference is due later sleep onset for Eastern residence (same sleep offset), while the pattern is reversed on weekends where both groups go to bed at similar hours but Western residence wake-up later. Furthermore, explore whether regional policy effects age-related changes in sleep patterns in three countries, Japan, Germany and United Kingdom. I find that sleep onset and offset begin to advance to later hours in late adulthood for residence Germany and the United Kingdom, but not evident for Japan within the age range of the sample. These results coincide relatively well with official records about effective retirement age [248]. Thus these regional difference in transitioning out of the labor force may influence age-related changes in sleep behavior within the country-level sphere.

Travel serves to balance skewed sleep patterns The change in sleep duration due to travel depends linearly on median sleep at home, a pattern identified both for weekdays and weekends. The main finding is that sleep during travel serves to balance at-home sleep behavior. Those who typically obtain short nighttime sleep at home, sleep more when travelling relative to home behavior, while those who maintain overall long median sleep duration, tend to sleep less than at home during travel-nights. The effect of travel on sleep has not been studied extensively nor for a cohort of this size before [262,263,264,265,128,266,267,268,269]. Interestingly, one of these studies identified the same pattern as I do – travel was negatively correlated with sleep duration on weekdays among kite surfers (N=94) [267]. Generally, travel is believed to have adverse effects on sleep, but my results imply a more complex impact [128,252,130,132,131,253,129].

Future outlook Finally, I hope that this study provides evidence and spikes interest for the richness of data stemming from consumer wearable devices. Hopefully the landscape of sleep research will transition to fully accept the use of wearable technology to explore and understand complex patterns of human sleep in modern society. I am excited for new paradigm of discoveries, and hopeful for changes.

A
Chapter 3

A.1 Parameter estimation in Mixed Effects Models

Parameter estimations for mixed effects model can not be written exactly, and therefore a maximum likelihood estimation (MLE) used. MLE is a method to estimate parameters of appropriate probability distribution for observed data by maximizing a likelihood function[221]. In mixed effects models, MLE is a function of the observations and the model parameters, which returns probability of observing a particular observation $\mathbf{y}$, given a set of model parameters. The mixed effects model parameters are a vector of the fixed effect, β, and γ is the vector of parameters used in the two covariance matrices $\mathbf{G}$ and $\mathbf{R}$. The covariance matrix $\mathbf{V}$, which describes the covariance between any two observations in the data set, can be calculated directly from the matrix representation of the model:

$$\begin{aligned}
\mathbf{V} = Var(\mathbf{y}) = Var(\mathbf{X}\beta + \mathbf{Z}u + \varepsilon) = \\
Var(\mathbf{X}\beta) + Var(\mathbf{Z}u) + Var(\varepsilon) = \\
Var(\mathbf{Z}u) + Var(\varepsilon) = \\
\mathbf{Z}\,Var(u)\,\mathbf{Z}' + Var(\varepsilon) = \\
\mathbf{Z}\,Var(u)\,\mathbf{Z}' + \mathbf{R}
\end{aligned}$$

variance of fixed effects are zero and $\mathbf{Z}$ is constant.

Other matrix reduction methods applied are based on rules of calculus for stochastic variables. The negative log likelihood function is given by:

$$\begin{aligned}
\ell(\mathbf{y}, \beta, \gamma) = \\
\frac{1}{2}\{nlog(\pi) + log|\mathbf{V}(\gamma)| + (\mathbf{y} - \mathbf{X}\beta)'\mathbf{V}(\gamma)^{-1}(\mathbf{y} - \mathbf{X}\beta)\} \\
\propto \frac{1}{2}\{log|\mathbf{V}(\gamma)| + (\mathbf{y} - \mathbf{X}\beta)'\mathbf{V}(\gamma)^{-1}(\mathbf{y} - \mathbf{X}\beta)\}
\end{aligned}$$

and the parameters are actually found by minimizing the negative log likelihood function.

$$(\hat{\beta}, \hat{\gamma}) = argmin_{(\beta,\gamma)}\, \ell(\mathbf{y}, \beta, \gamma)$$

The minimum is found in three steps:

1. Regardless of how the likelihood function estimates the optimal parameters to be, the estimates of fixed effects β can always be expressed as a function of random effects parameters where $\hat{B}(\hat{\gamma}) = (\mathbf{X}'(\mathbf{V}(\gamma))^{-1}\mathbf{X})^{-1}\mathbf{X}'(\mathbf{V}(\gamma))^{-1}\mathbf{y}$

2. The estimates of random effects are found by minimizing $\ell(\mathbf{y}, \hat{\beta}(\hat{\gamma}), \gamma)$

3. Fixed effects are estimated by $\hat{\beta} = \hat{\beta}(\hat{\gamma})$

However, the maximum likelihood method tends to underestimate the random effects and therefore an alternative criterion is used; the restricted (or residual) maximum likelihood (REML), which is considered the gold standard of parameter optimization in mixed effects models[215,221,218,228]. The REML methodology optimizes the full residuals rather than the observations ($\mathbf{y}$), which can be justified by the fact that the residuals contain a lot of information about the variance parameters ($\mathbf{V}$). In traditional linear models, (without fixed or random effects), the error term does not depend on the variance structure and residuals are independent from the mean parameter estimates. The residuals in mixed effects models is defined as[215]

$$\mathbf{e} = \mathbf{y} - \hat{\beta}\mathbf{X}$$

and depend on the variance since $\hat{\beta}$ is defined as

$$\hat{\beta} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$$

If we consider residuals independent from the fixed effect estimates, then the likelihood function of the original data can be expressed as product of the likelihoods for residuals and parameter estimates:

$$L(\mathbf{y}, \beta, \gamma) = L(\mathbf{e}, \beta, \gamma) \times L(\hat{\beta}, \beta, \gamma)$$

Applying negative log likelihood function, ℓ, on L and isolating the expression with residuals on one side gives:

$$\ell(\mathbf{e},\beta,\gamma) = \ell(\mathbf{y},\beta,\gamma) - \ell(\hat{\beta},\beta,\gamma)$$

Thus, the likelihood of residuals is the original likelihood subtracted the likelihood of parameter estimates. The original likelihood is expressed with the 'true' fixed effects, β, whereas the likelihood of parameters is expressed with β's estimated from the data and the distribution expressed as

$$\hat{\beta} \sim N(\beta, (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1})$$

If p denotes the number of parameters estimated in the mixed effects model, the multivariate normal negative log-likelihood is defined as:

$$\ell(\hat{\beta},\beta,\gamma) = \frac{1}{2}\{p \times log(2\pi) \ + log|(\mathbf{X}'\mathbf{V}(\gamma)^{-1}\mathbf{X})^{-1}| + (\beta - \hat{\beta})'\mathbf{X}'\mathbf{V}(\gamma)^{-1}\mathbf{X})^{-1}(\beta - \hat{\beta})\}$$

Since the β estimate is unbiased, it will always be estimated as $\hat{\beta}$ and the last term does not contribute to the likelihood and therefore the residual maximum likelihood is formulated as[215]

$$\begin{aligned} \ell_{re}(\beta, \gamma) = \\ \frac{1}{2}\{nlog(2\pi) + log|\mathbf{V}(\gamma)| + (\mathbf{y} - \mathbf{X}\beta)'(\mathbf{V}(\gamma))^{-1}(\mathbf{y} - \mathbf{X}\beta) - p \times log(2\pi) - log|(\mathbf{X}'(\mathbf{V}(\gamma))^{-1}\mathbf{X})^{-1}|\} = \\ \frac{1}{2}\{(n-p) \times log(2\pi) + log|\mathbf{V}(\gamma)| + (\mathbf{y} - \mathbf{X}\beta)'(\mathbf{V}(\gamma))^{-1}(\mathbf{y} - \mathbf{X}\beta)| + log|(\mathbf{X}'(\mathbf{V}(\gamma))^{-1}\mathbf{X})|\} \\ \propto \frac{1}{2}\{log|\mathbf{V}(\gamma)| + (\mathbf{y} - \mathbf{X}\beta)'(\mathbf{V}(\gamma))^{-1}(\mathbf{y} - \mathbf{X}\beta) + log|(\mathbf{X}'(\mathbf{V}(\gamma))^{-1}\mathbf{X})|\} \end{aligned}$$

The random effects are not statistical parameters in the model, and therefore not typically estimated. In some instances they contain desirable information and are then 'estimated' with empirical Bayes method, the matrix notation is formulated as:

$$\hat{\mathbf{u}} = \mathbf{G}\,\mathbf{Z}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\beta) \tag{A.1}$$

These 'estimates' of random effects are also called the posterior means and in statistics, referred to as prediction[225,221,220].

Significance of fixed effects

Typically in mixed effects model the point of interest are the estimates of fixed effects rather than random effects. Now I will examine how the significance of these are estimated. A linear combination of fixed effects model parameters $\mathbf{L}\beta'$ can be estimated if and only if there is vector λ such that $\lambda\mathbf{X} = \mathbf{L}'$. Estimate of fixed effects $\hat{\beta}$ will then be:[215]

$$\mathbf{L}\hat{\beta} = \mathbf{L}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$$

which can reduced to $\mathbf{L}'(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{L}$ applying the rule of $cov(\mathbf{A}\mathbf{x}) = \mathbf{A}cov(\mathbf{x})\mathbf{A}'$ and other matrix calculations. Considering that $\mathbf{L}\beta' = c$ is true, then:

$$\mathbf{L}\beta' - c \sim N(0, \mathbf{L}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y})$$

The Wald test can be applied and is formulated as[215]:

$$W = (\mathbf{L}'\beta - c)'\mathbf{L}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}(\mathbf{L}'\beta - c) \tag{A.2}$$

W has an approximately $\chi^2_{df_1}$ distribution where he degrees of freedom is the number of elements eliminated by the hypothesis. A better approximation can be attained with the Wald F-test where

$F = \frac{W}{df_1}$, combined with Satterthwaite approximation which supplies the effective degrees of freedom df_2 where the assumption is that F is F_{df_1,df_2} distributed. The p-value for the hypothesis $\mathbf{L}'\beta = c$ will then be[215]:

$$P_{\mathbf{L}'\beta=c} = P(F_{df_1,df_2} > F)$$

The Satterthwaite approximation can both be applied in the case of parameter estimation with maximum likelihood or residual maximum likelihood methodology[231]. Simulation indicate a Type 1 error is close to 0.05 when models are fitted using REML and p-values estimated with Satterhwaite approximation. These are considered anti-conservative but the error rates were sensitive to number of levels and items which became more acceptable with higher number of groups and items[231].

B

Chapter 4

B.1 Comparison of study sample to Walch *et al.* data-set

Country	Average sleep duration [hrs] Walch *et al.*	Average sleep duration [hrs] Study sample	Number of users in study sample
Netherlands	8.14	7.49	586
Belgium	8.10	7.37	202
France	8.10	7.44	2409
Australia	8.10	7.34	677
Canada	8.03	7.22	173
Italy	7.94	7.18	898
United Kingdom	7.94	7.41	3900
United States	7.92	7.08	941
Switzerland	7.90	7.33	518
China	7.89	6.96	2359
Denmark	7.87	7.32	473
Spain	7.86	7.09	3394
Mexico	7.81	6.87	461
Germany	7.74	7.37	7140
Brazil	7.62	6.91	201
Japan	7.50	6.47	17231
Singapore	7.48	6.72	275

Table B.1: Comparison of average sleep duration from the study sample to statistics from the data-set in Walch *et al.* study

Country	Average sleep onset [hh:mm] Walch *et al.*	Average sleep onset [hh:mm] Study sample	Number of users in study sample
Australia	22:49	23:43	677
Belgium	22:53	00:03	202
United States	22:57	00:11	941
Canada	23:03	00:19	173
Denmark	23:05	23:42	473
Switzerland	23:09	23:49	518
Netherlands	23:09	00:03	586
United Kingdom	23:11	23:53	3900
France	23:20	00:03	2409
Germany	23:21	23:42	7140
Japan	23:29	00:23	17231
Mexico	23:32	00:28	461
China	23:36	00:42	2359
Brazil	23:36	00:25	201
Italy	23:45	00:17	898
Singapore	23:48	00:37	275
Spain	23:51	00:43	3394

Table B.2: Comparison of average sleep onset from the study sample to statistics from the data-set in Walch *et al.* study

Country	Average sleep onset [hh:mm] Walch *et al.*	Average sleep onset [hh:mm] Study sample	Number of users in study sample
Australia	06:52	07:10	677
United States	06:52	07:23	941
Denmark	06:53	07:09	473
Belgium	06:56	07:32	202
Japan	06:58	06:58	17231
Switzerland	06:59	07:16	518
Canada	07:03	07:39	173
Germany	07:07	07:11	7140
UK	07:07	07:25	3900
Brazil	07:11	07:26	201
Singapore	07:15	07:26	275
Netherlands	07:15	07:40	586
Mexico	07:19	07:27	461
France	07:24	07:36	2409
China	07:27	07:45	2359
Italy	07:39	07:36	898
Spain	07:40	07:56	3394

Table B.3: Comparison of average sleep offset from the study sample to statistics from the data-set in Walch *et al.* study

B.2 Demographic representation

I compare age statistics (median age) in the study sample to information provided by the United Nation Population Division (UN) for the five countries with the most users in the data-set in Table B.4 [238]. I also compare age standardized BMI statistics of the study sample to population estimates provided by the World Health Organization (WHO) in Table B.5 [239,240].

Country (# users)	Study sample (median age)	UN data-set (median age)
Japan (N=17231)	45	46
Germany (N=7140)	39	46
Russia (N=5095)	34	39
Taiwan (5028)	40	40
United Kingdom (N=3900)	40	40

Table B.4: Comparison of median age of users from the five countries with the most data in the study sample to population statistics provided by the United Nations

Country	Study sample		WHO data set	
	Male	Female	Male	Female
Japan (N=17231)	24.1	22.8	23.1-24	21.3-22.3
China with Taiwan included (N=7387)	25.2	23.4	23.6-24.7	22.9-24.1
Germany (N=7140)	27.1	26.3	26.5-28.1	24.9-26.8
Russia (N=5095)	26.9	25.4	25.1-26.8	25.5-27.3
United Kingdom (N=3900)	26.5	27.9	26.9-27.7	26.6-27.4

Table B.5: Comparison of age standardized BMI values for users from the five countries with the most data in the study sample to population statistics provided by the World Health Organization (WHO)3. The range of values displayed for the WHO data set represents the 95 % confidence interval.

C
Chapter 5

C.1 Underlying differences across users with same estimates of chronotype, social jetlag and median sleep duration

On Figure C.1-C.3 we illustrate how different metrics of sleep (sleep duration, bed and wake-time, weekend-weekday differences, width of sleep onset, offset and duration, and the weekend-weekday width differences) distributes across users who have one sleep epidemiological metric approximately the same.

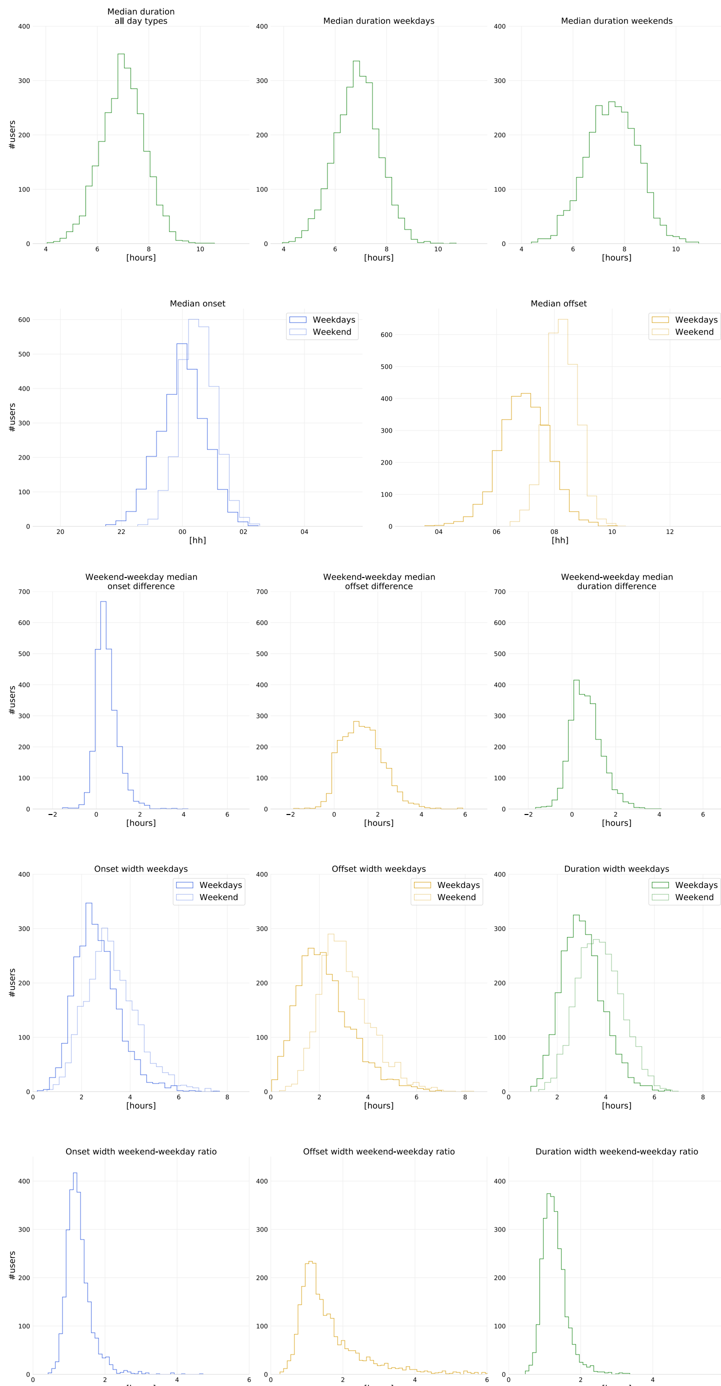


Figure C.1: Distribution of different metrics of sleep for users with approximately the same chronotype (∼04:30 and $N_{group} = 2719$)

Figure C.2: Distribution of different metrics of sleep for users with approximately the same social jetlag (∼0.75 hours and $N_{group} = 2479$)

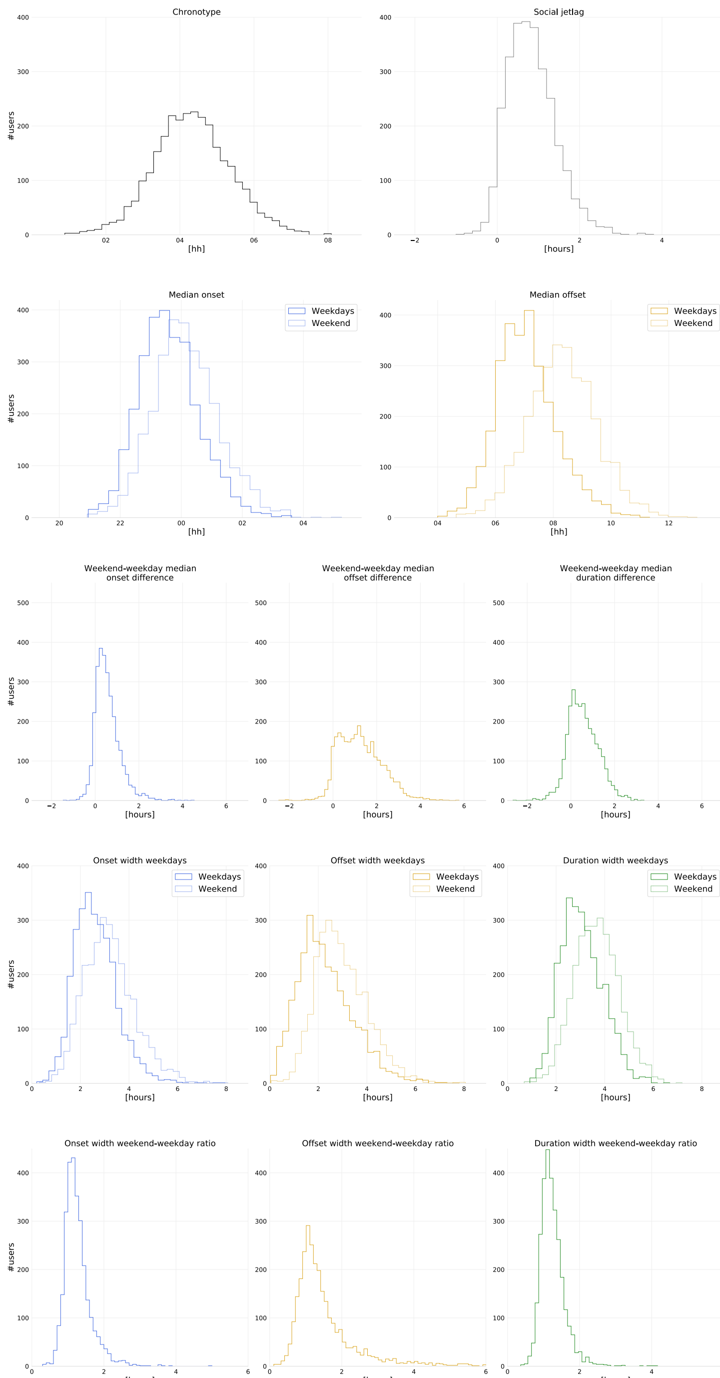


Figure C.3: Distribution of different metrics of sleep for users with approximately the same overall median sleep duration (∼7.5 hours and $N_{group} = 2860$)

C.2 New measure for variability

I introduce a new measure to quantify sleep variability. Typically variability is measured as standard deviation (std) of the distribution for measure, but I rather suggest using the difference between the 90th and 10th percentile. I consider it more intuitive and provide an example to demonstrate; imagine the scenario where one has to comprehend either "John has 0.74 hours std in wake-up time on weekdays" or "John wakes up 80 % of the time within a span of 15 minutes on weekdays". The two measures correlate nearly perfectly as demonstrated on Figure C.4.

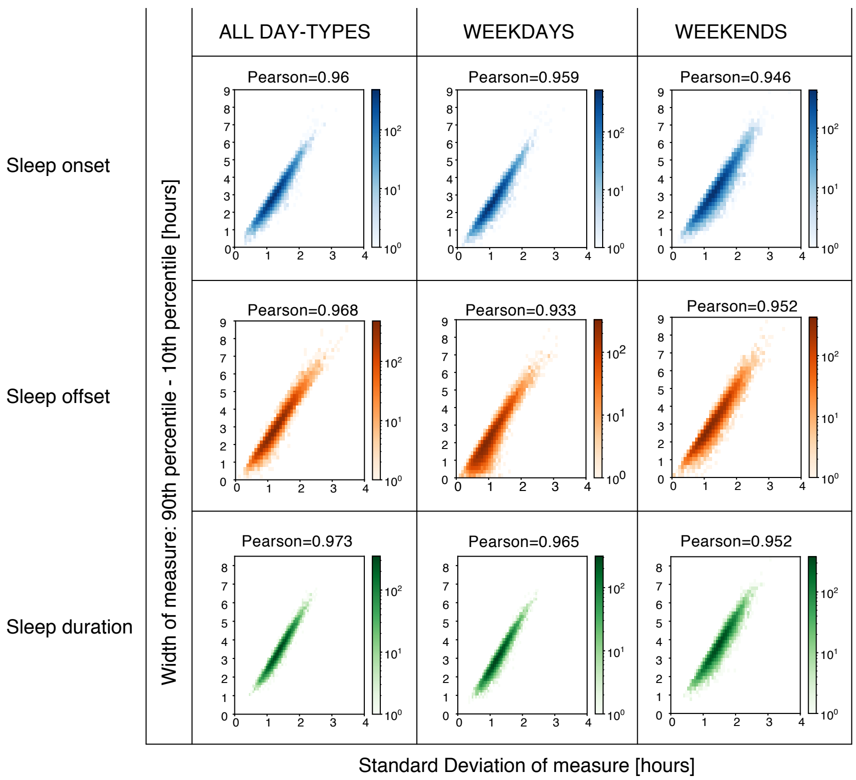


Figure C.4: Illustrates how width of a distribution (calculated as the 90th percentile minus 10th percentile) correlates with standard deviation for sleep onset, offset and duration by all day-types, weekdays and weekends separately

C.3 Short or long sleep duration but no skew?

I estimate the probability for a randomly sampled individual to belong to different demographic groups (`gender` [female/male], `age group` [19-24/25-29/30-34/35-39/40-44/45-49/50-54/55-59/60-67] and `region of residence` (east=Asia/west=Europe & N-America) in the full data-set. Then I filter the data by certain criteria to obtain two dataframes; i) short median sleep duration (<6.5 hrs) with no or negative skew (skew<0.25) and ii) long median sleep duration (>8 hrs) with positive or no skew (>-0.25). For these two different dataframes we look again at the probability to belong to demographic groups (listed above) and obtain the relative probability by dividing with the probability of being part of that demographic group in the full data-set.

C.3.1 Overall

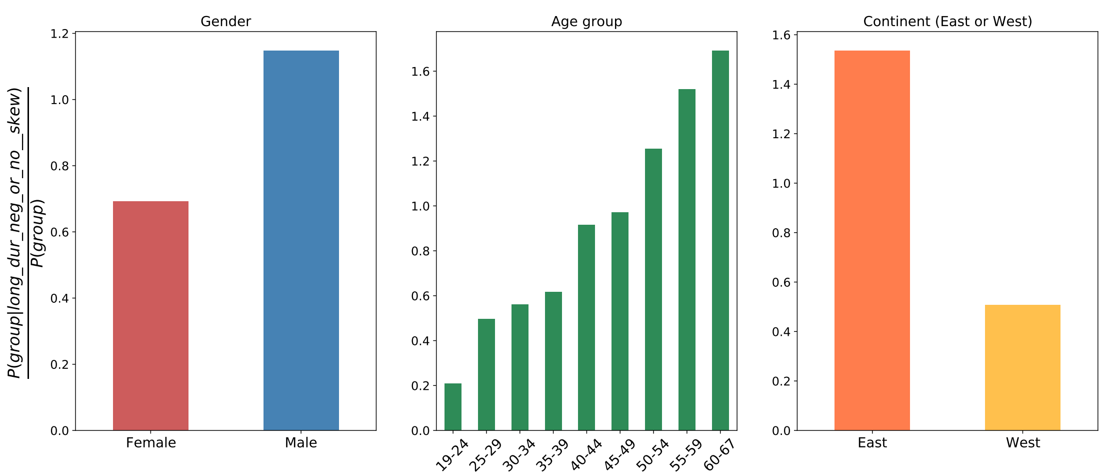


Figure C.5: The relative probability that individual belongs to a demographic group in the filtered data with short sleep duration (<6.5 hrs) and no or negative skew (skew<0.25). Skew is estimated for the distribution of sleep duration for all day-types.

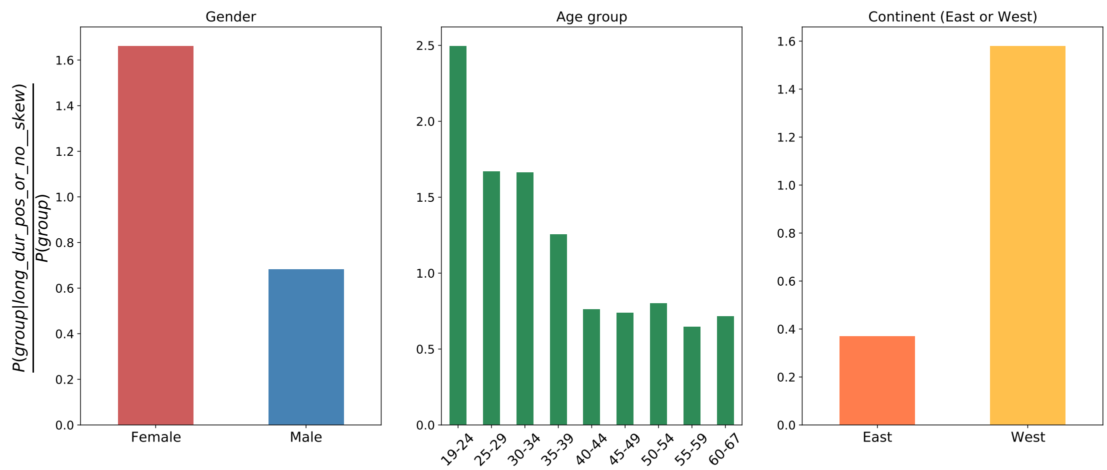


Figure C.6: The relative probability that individual belongs to a demographic group in the filtered data with long sleep duration (>8.0 hrs) and no or positive skew (skew>-0.25). Skew is estimated for the distribution of sleep duration for all day-types.

C.3.2 Weekdays

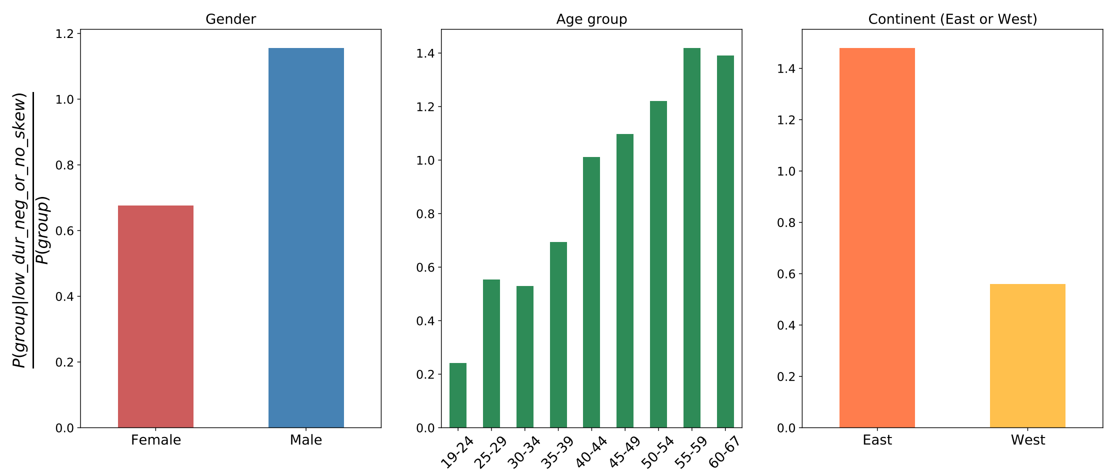


Figure C.7: The relative probability that individual belongs to a demographic group in the filtered data with short sleep duration (<6.5 hrs) and no or negative skew (skew<0.25). Skew is estimated for the distribution of sleep duration for weekday-nights.

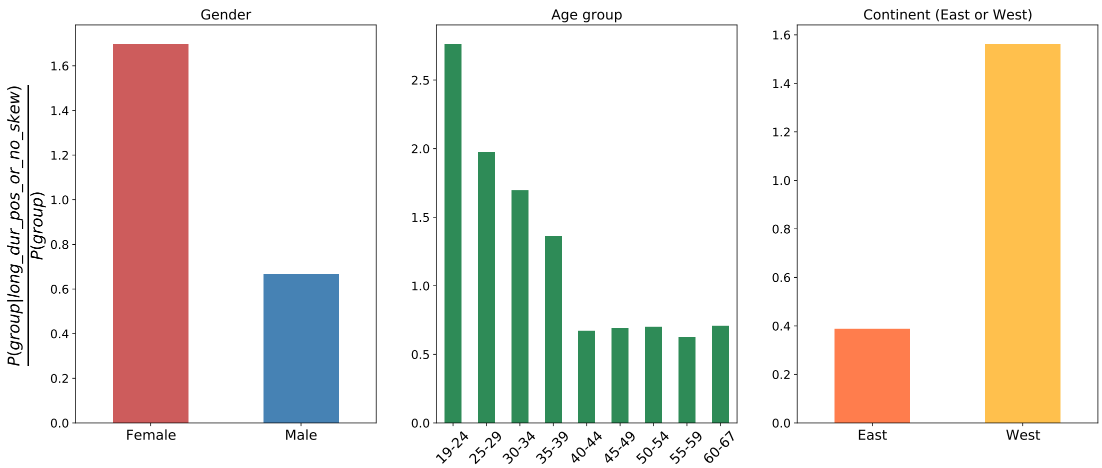


Figure C.8: The relative probability that an individual belongs to a demographic group in the filtered data with long sleep duration (>8.0 hrs) and positive or no skew (skew>-0.25). Skew is estimated for the distribution of sleep duration for weekday-nights.

C.3.3 Weekends

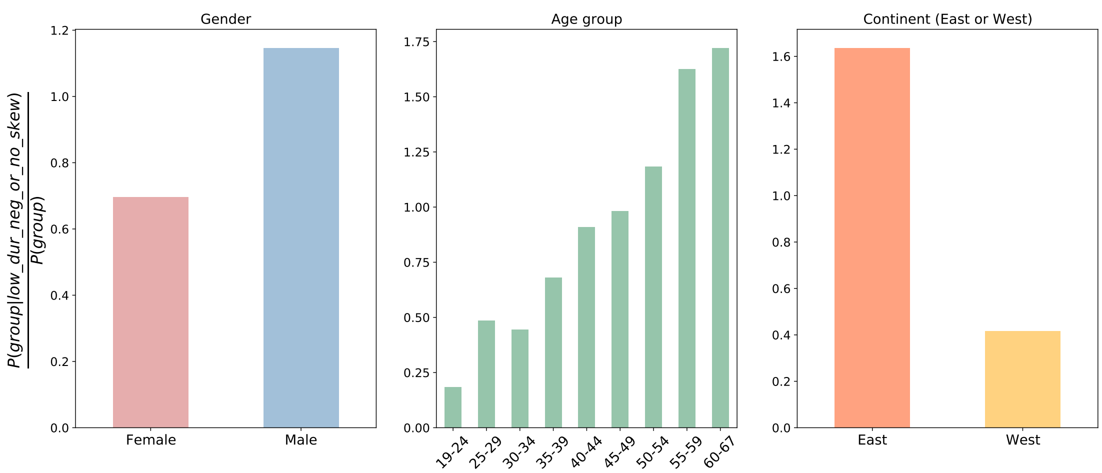


Figure C.9: The relative probability that an individual belongs to a demographic group in the filtered data with short sleep duration (<6.5 hrs) and no or negative skew (skew<0.25). Skew is estimated for the distribution of sleep duration for weekend-nights

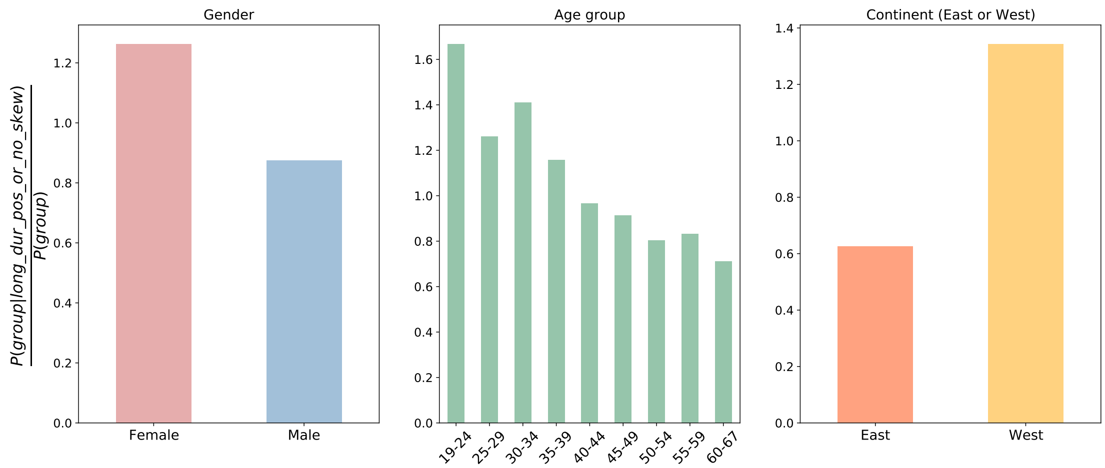


Figure C.10: The relative probability that an individual belongs to a demographic group in the filtered data with long sleep duration (>8.0 hrs) and positive or no skew (skew>-0.25). Skew is estimated for the distribution of sleep duration for weekend-nights

C.4 Skew group characteristics

I plot the distribution of all sleep metrics, by the three skew groups; the 2000 individuals with either the most positive, neutral and negative skewed distribution of sleep duration for all day-types, weekday-nights or weekend-nights respectively Figures C.11–C.13.

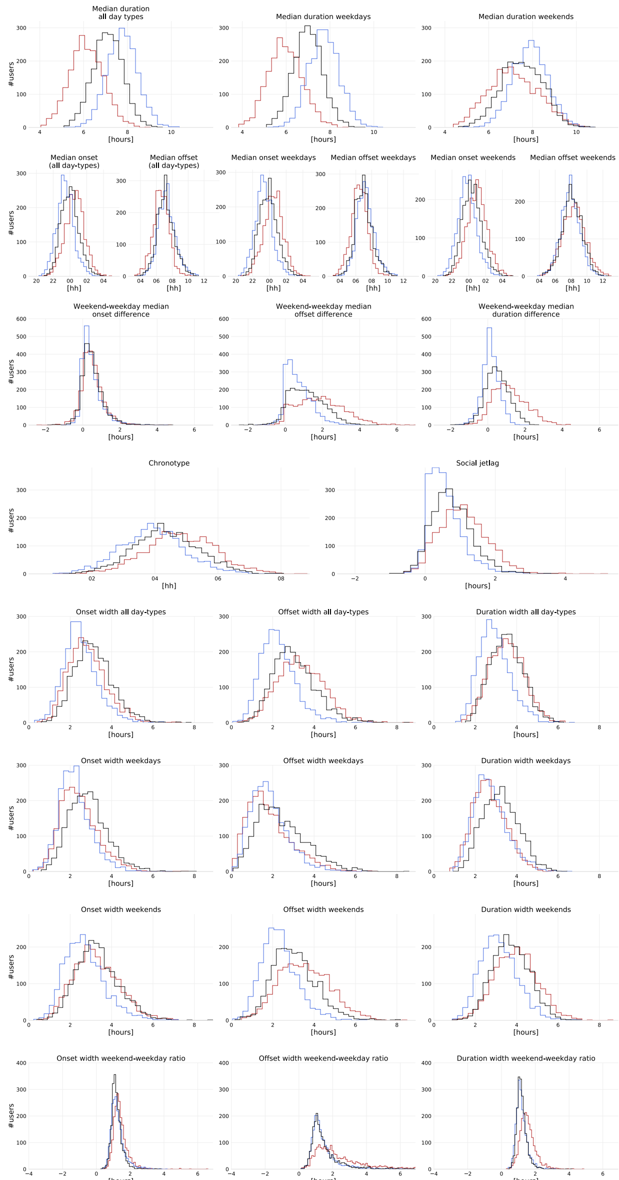


Figure C.11: Distributions of multiple features for the 2000 most positively (red), neutrally (black) and negatively (blue) skewed individuals. Skew is estimated for the distribution of sleep duration for all day-types.

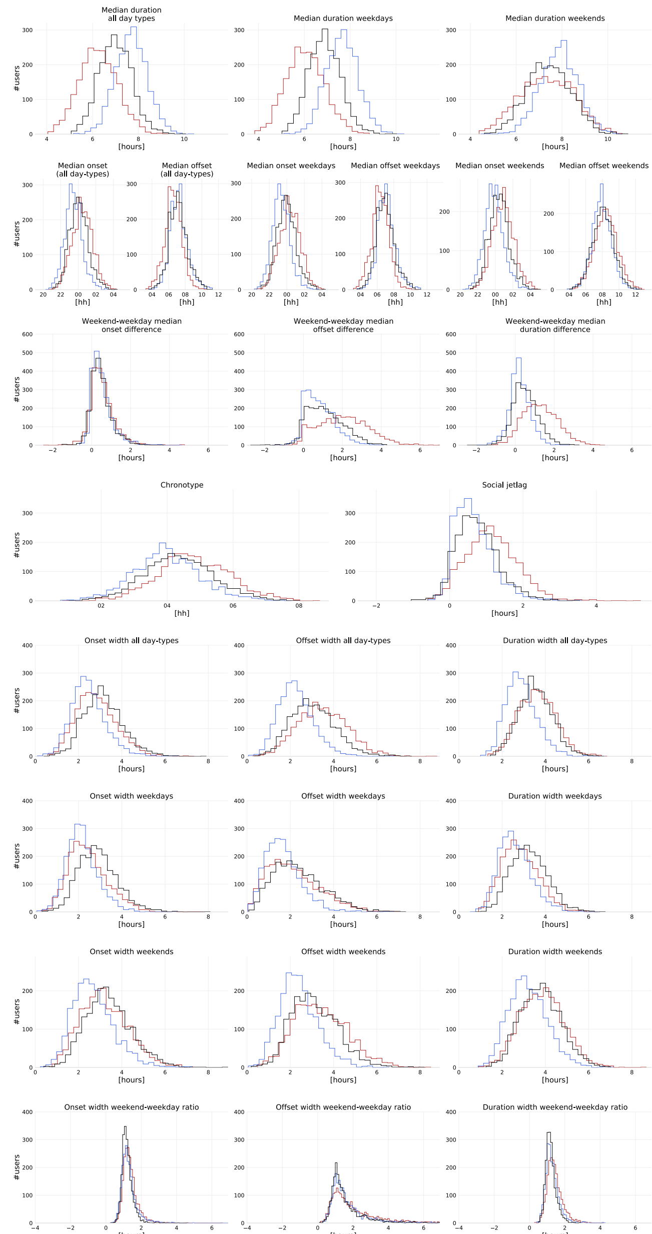


Figure C.12: Distributions of multiple features for the 2000 most positively (red), neutrally (black) and negatively (blue) skewed individuals. Skew is estimated for the distribution of sleep duration on weekdays.

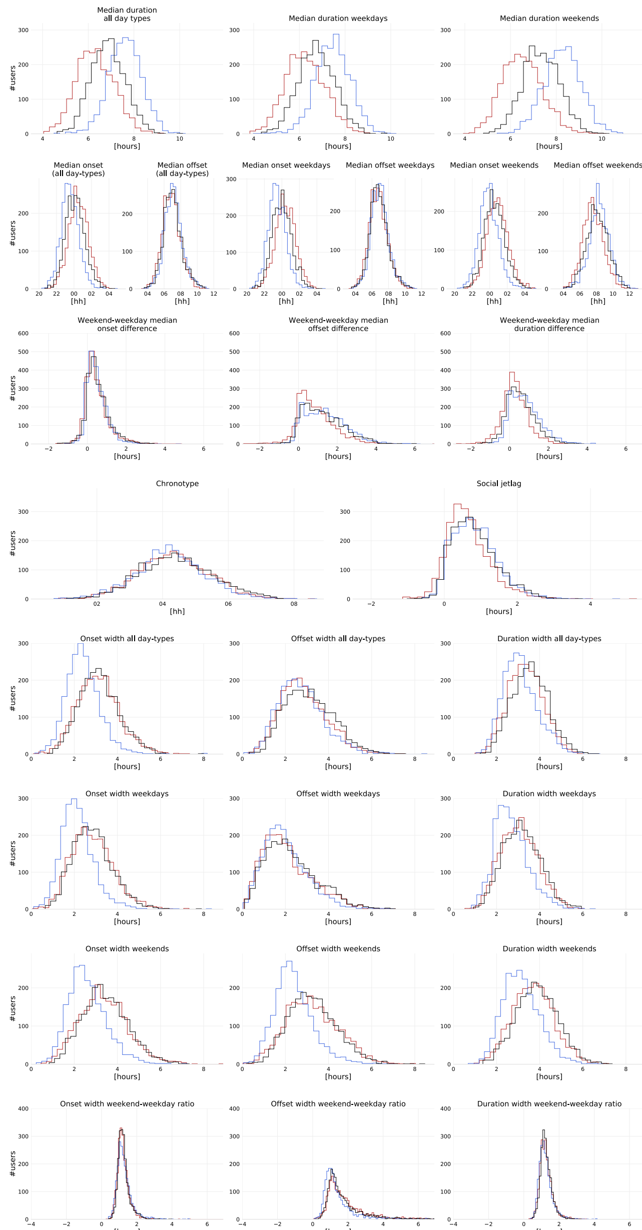


Figure C.13: Distributions of multiple features for the 2000 most positively (red), neutrally (black) and negatively (blue) skewed individuals. Skew is estimated for the distribution of sleep duration on weekends.

D

Chapter 6

D.1 Regional differences

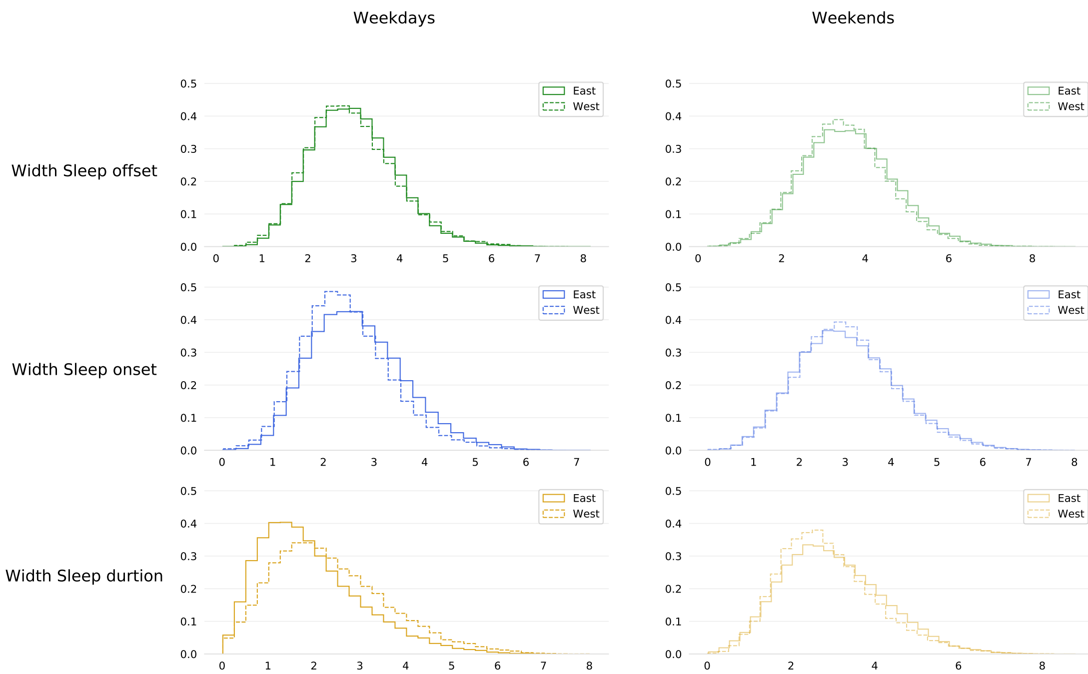


Figure D.1: The distribution of width (10th percentile subtracted from the 90th percentile) for sleep onset (blue), offset (yellow) and duration (green) separately by day-type and residents in either the East (full line) or West (dotted line)

D.2 Results from Mixed effects models

D.2.1 Sleep onset

Fixed effects	Estimate	Std.Error	Df	t-value	P(>|t|)
(Intercept)	-2.439e-01	6.141e-02	4.753e+01	-3.972	0.000241 ***
age_centered_squared	4.857e-04	4.970e-05	6.969e+04	9.772	<2e-16 ***
weekendTrue	4.887e-01	1.741e-03	1.117e+07	280.716	<2e-16 ***
BMIcategory0	1.605e-01	3.567e-02	6.982e+04	4.499	6.83e-06 ***
BMIcategory2	9.389e-03	1.652e-02	6.967e+04	0.568	0.569778
BMIcategory3	1.039e-01	2.035e-02	7.001e+04	5.106	3.29e-07 ***
genderMALE	2.546e-01	1.291e-02	6.967e+04	19.724	<2e-16 ***
age_centered	-1.092e-02	6.427e-04	6.964e+04	-16.983	<2e-16 ***
age_centered_squared:weekendTrue	5.754e-05	5.481e-06	1.118e+07	10.498	<2e-16 ***
age_centered_squared:BMIcategory0	-2.386e-04	1.467e-04	6.947e+04	-1.626	0.103910
age_centered_squared:BMIcategory2	6.793e-05	5.467e-05	6.967e+04	1.243	0.214054
age_centered_squared:BMIcategory3	1.563e-04	7.465e-05	7.000e+04	2.094	0.036235 *
age_centered_squared:genderMALE	-2.973e-04	5.302e-05	6.961e+04	-5.608	2.06e-08 ***
weekendTrue:genderMALE	-2.341e-02	1.718e-03	1.117e+07	-13.629	<2e-16 ***
weekendTrue:BMIcategory0	-5.131e-02	4.616e-03	1.118e+07	-11.115	<2e-16 ***
weekendTrue:BMIcategory2	6.232e-02	1.787e-03	1.118e+07	34.876	<2e-16 ***
weekendTrue:BMIcategory3	9.841e-02	2.441e-03	1.118e+07	40.318	<2e-16 ***
BMIcategory0:genderMALE	-4.704e-02	4.048e-02	6.988e+04	-1.162	0.245187
BMIcategory2:genderMALE	2.895e-02	1.756e-02	6.959e+04	1.648	0.099284 .
BMIcategory3:genderMALE	5.550e-02	2.189e-02	6.988e+04	2.535	0.011236 *
genderMALE:age_centered	-1.427e-02	6.810e-04	6.958e+04	-20.947	<2e-16 ***
weekendTrue:age_centered	-8.261e-03	7.138e-05	1.118e+07	-115.736	<2e-16 ***
BMIcategory0:age_centered	-4.530e-04	2.112e-03	6.886e+04	-0.214	0.830191
BMIcategory2:age_centered	6.269e-03	7.025e-04	6.969e+04	8.923	<2e-16 ***
BMIcategory3:age_centered	9.354e-03	9.484e-04	7.006e+04	9.863	<2e-16 ***

Table D.1: Estimates of fixed effects from mixed effects model for sleep onset

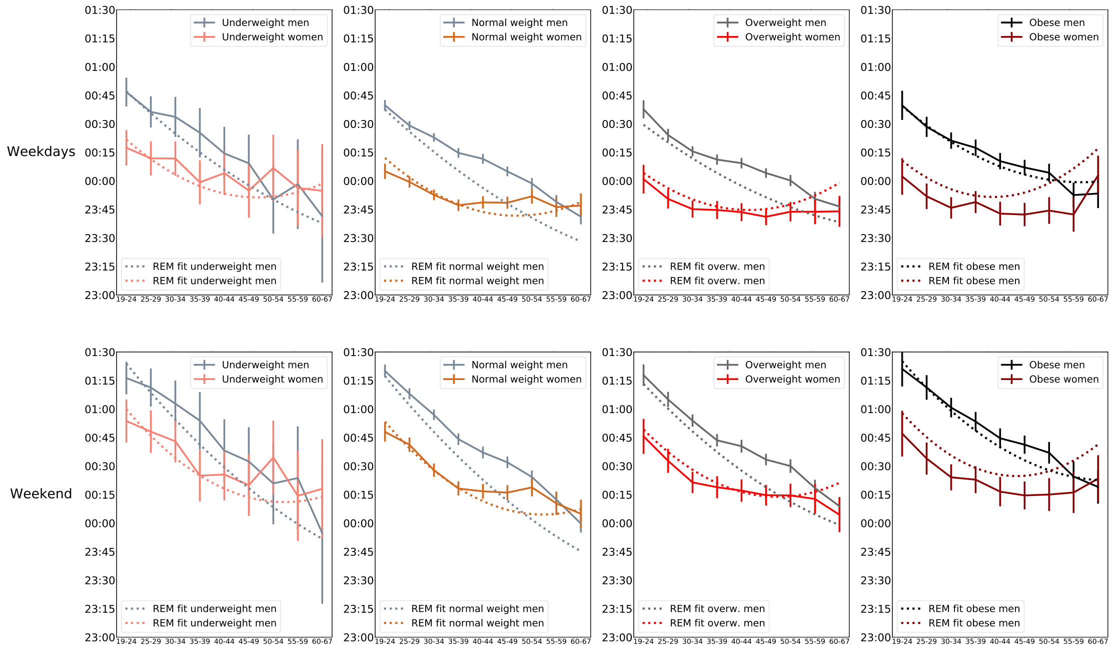


Figure D.2: Aggregate averages of the raw data for development of sleep onset by age group with 95 % CI split up by gender, BMI category and day type (weekend/weekday) raw data compared to estimates of mixed effects model fit

D.2.2 Sleep offset

Fixed effects	Estimate	Std.Error	Df	t-value	P(>|t|)
(Intercept)	6.970e+00	4.925e-02	4.402e+01	141.529	<2e-16 ***
age_centered_squared	7.999e-04	4.345e-05	6.878e+04	18.407	<2e-16 ***
weekendTrue	1.151e+00	1.721e-03	1.111e+07	669.021	<2e-16 ***
BMIcategory0	1.212e-01	3.324e-02	6.886e+04	3.646	0.000266 ***
BMIcategory2	-5.333e-02	1.476e-02	6.874e+04	-3.612	0.000304 ***
BMIcategory3	-6.564e-02	1.850e-02	6.913e+04	-3.549	0.000387 ***
age_centered	-1.786e-02	6.019e-04	6.870e+04	-29.681	<2e-16 ***
age_centered_squared:weekendTrue	-5.365e-04	5.426e-06	1.109e+07	-98.873	<2e-16 ***
age_centered_squared:BMIcategory0	-4.940e-04	1.382e-04	6.848e+04	-3.574	0.000352 ***
age_centered_squared:BMIcategory2	1.125e-04	5.134e-05	6.872e+04	2.192	0.028382 *
age_centered_squared:BMIcategory3	3.690e-04	7.027e-05	6.910e+04	5.251	1.51e-07 ***
age_centered_squared:genderMALE	-1.257e-04	4.185e-05	6.876e+04	-3.003	0.002672 **
weekendTrue:genderMALE	-4.450e-02	1.699e-03	1.113e+07	-26.193	<2e-16 ***
weekendTrue:BMIcategory0	-7.068e-03	4.584e-03	1.109e+07	-1.542	0.123068
weekendTrue:BMIcategory2	6.044e-02	1.768e-03	1.109e+07	34.180	<2e-16 ***
weekendTrue:BMIcategory3	5.904e-02	2.422e-03	1.109e+07	24.380	<2e-16 ***
BMIcategory0:genderMALE	9.549e-02	3.779e-02	6.892e+04	2.527	0.011519 *
BMIcategory2:genderMALE	-9.800e-03	1.421e-02	6.873e+04	-0.690	0.490280
BMIcategory3:genderMALE	-2.170e-02	1.878e-02	6.909e+04	-1.155	0.248033
genderMALE:age_centered	-8.940e-03	6.354e-04	6.864e+04	-14.070	<2e-16 ***
weekendTrue:age_centered	-1.221e-02	7.074e-05	1.109e+07	-172.609	<2e-16 ***
BMIcategory0:age_centered	-5.052e-03	1.989e-03	6.781e+04	-2.540	0.011098 *
BMIcategory2:age_centered	5.795e-03	6.612e-04	6.874e+04	8.763	<2e-16 ***
BMIcategory3:age_centered	7.412e-03	8.934e-04	6.916e+04	8.296	<2e-16 ***

Table D.2: Estimates of fixed effects from mixed effects model for sleep offset

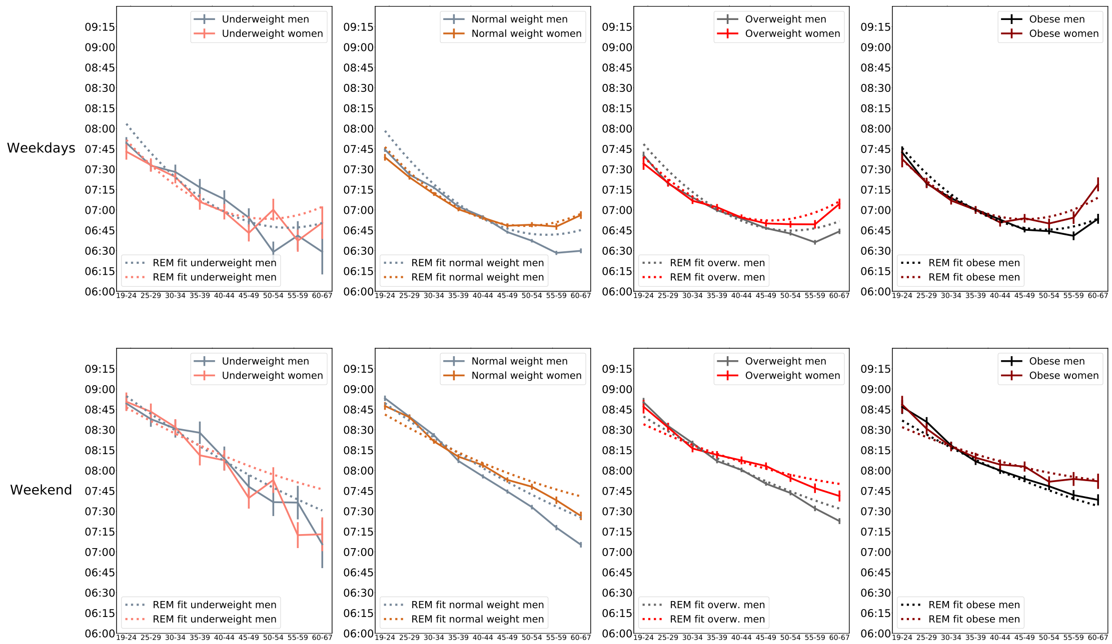


Figure D.3: Aggregate averages of the raw data for development of sleep offset by age group with 95 % CI split up by gender, BMI category and day type (weekend/weekday) raw data compared to estimates of mixed effects model fit

D.2.3 Sleep duration

Fixed effects	Estimate	Std.Error	Df	t-value	P(>|t|)
Fixed effects	7.031e+00	4.784e-02	4.927e+01	146.986	<2e-16 ***
(Intercept)	4.060e-04	2.671e-05	6.965e+04	15.202	<2e-16 ***
age_centered_squared	6.355e-01	1.986e-03	1.118e+07	320.009	<2e-16 ***
weekendTrue	-2.156e-02	2.806e-02	6.959e+04	-0.768	0.442334
BMIcategory0	-5.557e-02	1.298e-02	6.931e+04	-4.280	1.87e-05 ***
BMIcategory2	-1.705e-01	1.602e-02	6.997e+04	-10.645	<2e-16 ***
BMIcategory3	-1.882e-01	8.525e-03	6.948e+04	-22.072	<2e-16 ***
genderMALE	-7.872e-03	5.034e-04	6.935e+04	-15.639	<2e-16 ***
age_centered	-5.691e-04	6.252e-06	1.118e+07	-91.025	<2e-16 ***
age_centered_squared:weekendTrue	-3.042e-04	1.152e-04	6.884e+04	-2.642	0.008255 **
age_centered_squared:BMIcategory0	6.892e-05	4.287e-05	6.928e+04	1.608	0.107912
age_centered_squared:BMIcategory2	1.890e-04	5.876e-05	6.994e+04	3.217	0.001295 **
age_centered_squared:BMIcategory3	-2.490e-02	1.960e-03	1.118e+07	-12.706	<2e-16 ***
weekendTrue:genderMALE	4.808e-02	5.265e-03	1.118e+07	9.132	<2e-16 ***
weekendTrue:BMIcategory0	-3.909e-03	2.038e-03	1.118e+07	-1.918	0.055158 .
weekendTrue:BMIcategory2	-4.545e-02	2.784e-03	1.118e+07	-16.323	<2e-16 ***
weekendTrue:BMIcategory3	1.102e-01	3.180e-02	6.967e+04	3.465	0.000531 ***
BMIcategory0:genderMALE	-6.515e-02	1.380e-02	6.912e+04	-4.721	2.35e-06 ***
BMIcategory2:genderMALE	-1.116e-01	1.722e-02	6.970e+04	-6.484	9.00e-11 ***
BMIcategory3:genderMALE	4.861e-03	5.333e-04	6.928e+04	9.116	<2e-16 ***
genderMALE:age_centered	-4.022e-03	8.142e-05	1.118e+07	-49.399	<2e-16 ***
weekendTrue:age_centered	-5.180e-03	1.655e-03	6.767e+04	-3.130	0.001750 **
BMIcategory0:age_centered	-5.370e-04	5.524e-04	6.932e+04	-0.972	0.330993
BMIcategory2:age_centered	-2.629e-03	7.468e-04	7.006e+04	-3.521	0.000430 ***
BMIcategory3:age_centered					

Table D.3: Estimates of fixed effects from mixed effects model for sleep duration

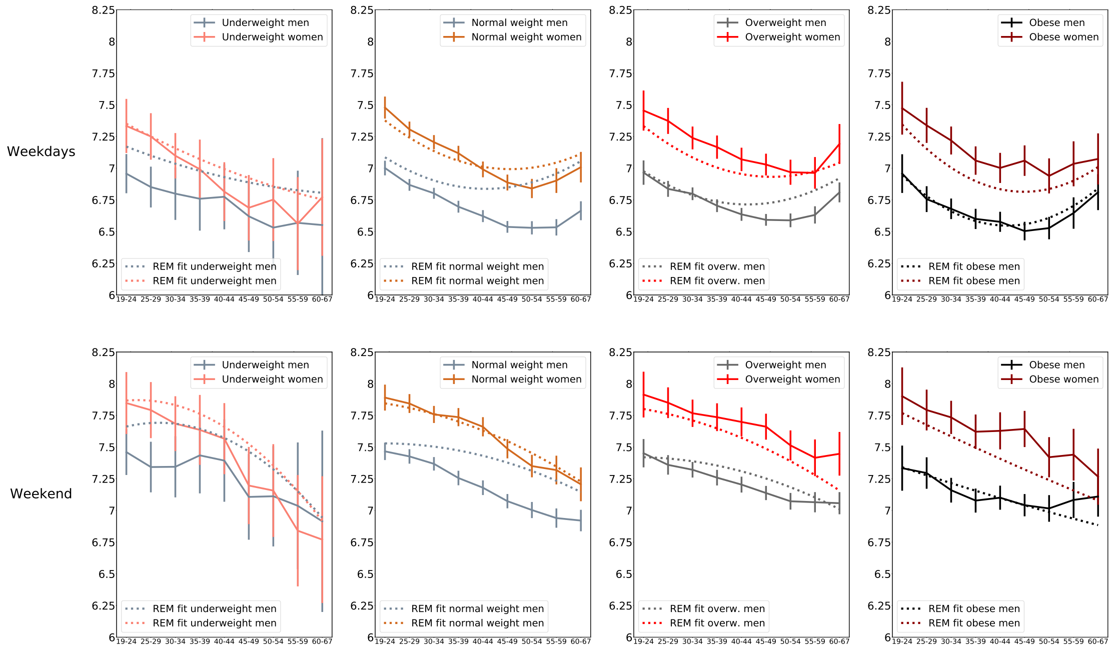


Figure D.4: Aggregate averages of the raw data for development of sleep duration by age group with 95 % CI split up by gender, BMI category and day type (weekend/weekday) raw data compared to estimates of mixed effects model fit

D.3 Effective retirement age

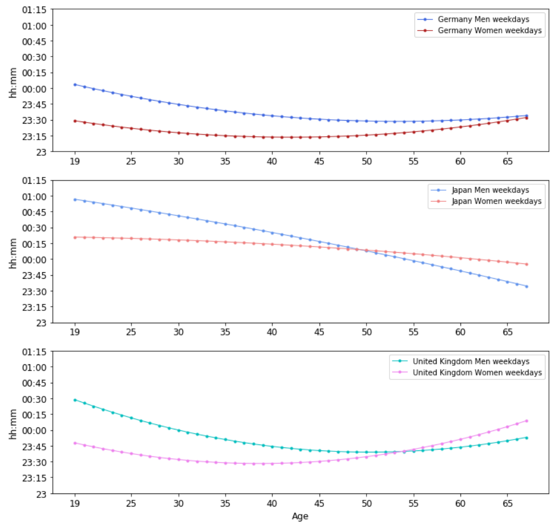


Figure D.5: Visualization of the mixed effects model fit for development of sleep onset with age, separately for Japan, Germany and UK

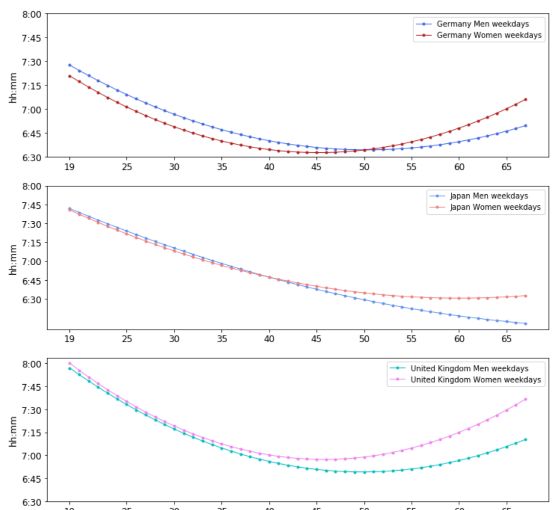


Figure D.6: Visualization of the mixed effects model fit for development of sleep offset with age, separately for Japan, Germany and UK

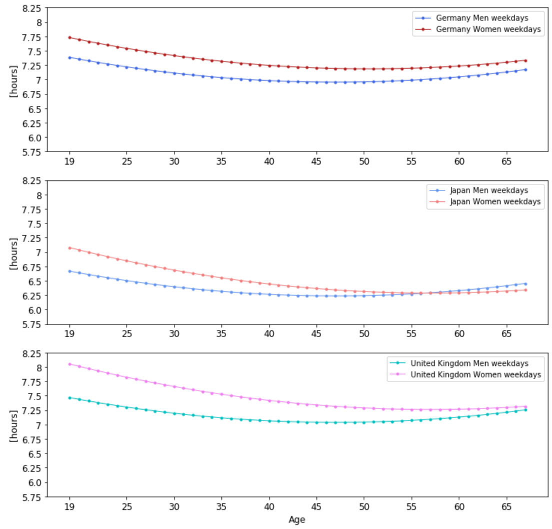


Figure D.7: Visualization of the mixed effects model fit for development of sleep duration with age, separately for Japan, Germany and UK

Age	Sleep Onset [hh:mm]		Sleep Offset [hh:mm]		Sleep duration [hours]	
	Women	Men	Women	Men	Women	Men
19	23:29	00:04	07:21	07:28	7.73	7.38
25	23:22	23:53	07:01	07:09	7.54	7.22
30	23:18	23:44	06:49	06:57	7.42	7.11
35	23:15	23:38	06:40	06:47	7.32	7.03
40	23:14	23:34	06:35	06:40	7.25	6.98
45	23:14	23:31	06:33	06:36	7.20	6.96
50	23:16	23:29	06:34	06:34	7.19	6.96
51	23:16	23:29	06:35	06:34	7.19	6.96
52	23:17	23:29	06:36	06:35	7.19	6.97
53	23:17	23:29	06:37	06:35	7.19	6.97
54	23:18	23:29	06:38	06:35	7.19	6.98
55	23:19	23:29	06:40	06:35	7.20	6.99
56	23:19	23:29	06:41	06:36	7.20	7.00
57	23:20	23:29	06:43	06:37	7.21	7.01
58	23:21	23:29	06:44	06:38	7.22	7.02
59	23:21	23:29	06:46	06:38	7.22	7.03
60	23:22	23:30	06:48	06:40	7.23	7.05
61	23:24	23:30	06:50	06:41	7.25	7.06
62	23:25	23:31	06:53	06:42	7.26	7.08
63	23:26	23:31	06:55	06:43	7.27	7.09
64	23:28	23:32	06:58	06:44	7.28	7.11
65	23:29	23:32	07:00	06:46	7.30	7.13
66	23:31	23:33	07:04	06:48	7.32	7.14
67	23:32	23:34	07:06	06:50	7.33	7.17

Table D.4: Exact estimate of sleep onset offset and duration from mixed effects model for men and women residing in Germany

Age	Sleep Onset [hh:mm]		Sleep Offset [hh:mm]		Sleep duration [hours]	
	Women	Men	Women	Men	Women	Men
19	00:21	00:56	07:41	07:42	7.08	6.67
25	00:20	00:59	07:22	07:24	6.85	6.5
30	00:18	00:41	07:08	07:10,	6.69	6.5
35	00:16	00:33	06:57	06:58	6.55	6.32
40	00:14	00:25	06:47	06:47	6.44	6.26
45	00:11	00:17	06:40	06:38	6.37	6.24
50	00:08	00:08	06:35	06:29	6.31	6.24
51	00:08	00:06	06:34	06:28	6.31	6.25
52	00:07	00:04	06:34	06:26	6.30	6.25
53	00:07	00:02	06:32	06:25	6.29	6.26
54	00:06	00:01	06:32	06:23	6.29	6.26
55	00:05	23:58	06:32	06:22	6.29	6.27
56	00:04	23:56	06:31	06:21	6.29	6.28
57	00:03	23:55	06:31	06:20	6.29	6.29
58	00:02	23:53	06:31	06:19	6.29	6.30
59	00:01	23:51	06:31	06:17	6.29	6.31
60	00:01	23:49	06:31	06:17	6.29	6.33
61	23:59	23:47	06:31	06:16	6.29	6.34
62	23:59	23:45	06:31	06:14	6.30	6.36
63	23:58	23:43	06:31	06:14	6.31	6.38
64	23:57	23:41	06:31	06:13	6.31	6.39
65	23:57	23:38	06:32	06:12	6.32	6.41
66	23:56	23:37	06:32	06:11	6.33	6.43
67	23:55	23:35	06:32	06:11	6.34	6.45

Table D.5: Exact estimate of sleep onset offset and duration from mixed effects model (with country as fixed effects) for men and women residing in Japan

D.4 Identifying "parent apps"

I identify mobile applications as "parent apps" if they are intended for parents with young children. I conducted an online search to create a list of apps that are popular and intended to assist parents with newborns and young children. Those are; Nurture, Mush, Hoop, Ask the Midwife, The Wonder Weeks, Pabobo, Tiny Beans, Breast Feeding Friend, Annabel Karmel's Recipes, Fisher-Price Apps, Talkspace, Peanut, Today's Parent My Family, MyMedela app, Sound Sleeper, Milk Maid, What to Expect, The Wonder Weeks Milestone Memories, Mom Maps, Mama Papa Map, Iku Log, Baby Manager, Wonder Weeks, Famm, Naki Pita! The second filtering method I used was to search for keywords in app names and if there was a match, the app was categorized as "parent app". The key words used were:

- In English; baby, infant, breastfeeding, parental control, parenting, pregnancy, pregnant
- In Spanish, Japanese, Chinese and Arabic; I searched for baby

E

Chapter 7

E.1 Down-sampling nights at home

One of the limitations when studying effect of travel on sleep, is the disproportionate amount travel-nights compared to nights at home in the data-set (6 % of weekdays and 9.3 % of weekends are travel-nights).

To contest to that presumption that imbalance sample sizes influence the results, I perform down-sampling where I randomly select nights at home to be equal to the number of travel-nights (for each user) and compare the sample distributions (both visually and by percentiles) for $\Delta_{home\ DS}$ and Δ_{home}. The process is described step-by-step;

- Repeat 50 times;
  - For each user I randomly choose N_{travel} nights recorded at home
  - For those randomly drawn nights, I estimate Δ_{home} and store it for each user
- Estimate $\Delta_{home\ DS}$ for each user from the 50 trials
- Estimate the quartiles for the sample distribution of $\Delta_{home\ DS}$

Results are listed in Tables E.1 and E.2 and distributions also visualised on Figure E.1. The distribution for down-sampled home-sleep ($\Delta_{home\ DS}$) is actually narrower than for the full sample. That can be rationalized by the fact that 70 % of users have 5 or less days recorded travel-nights, but when I examined the development of the standard deviation by number of data-points, the standard deviation increases as the number of data-points increases, and did not stabilized until there are about 10 recorded nights. TBD cite paper.

Iteration	1	2	3	4	5	**Full sample – Home**	**Full sample – Travel**
Minimum	-0.565	-0.588	-0.596	-0.619	-0.617	-1.39	-5.25
Lower quartile	-0.0532	-0.0515	-0.0533	-0.0524	-0.0534	-0.110	-0.417
Median	0	0	0	0	0	-0.0140	0.239
Upper quartile	0.0342	0.0340	0.0340	0.0363	0.0346	0.086	0.933
Maximum	0.726	0.754	0.711	0.735	0.766	1.16	5.98

Table E.1: Sample quartiles of $\Delta_{home\ DS}$ [hours] home-nights are randomly selected and equal to the number of travel-nights on weekdays

Iteration	1	2	3	4	5	Full sample – Home	Full sample – Travel
Minimum	-0.692	-0.841	-0.680	-0.752	-0.746	-1.39	-5.25
Lower quartile	-0.0763	-0.0783	-0.0777	-0.0785	-0.0789	-0.110	-0.417
Median	0	0	0	0	0	-0.0140	0.239
Upper quartile	0.0114	0.0110	0.0117	0.0112	0.0134	0.086	0.933
Maximum	0.712	0.604	0.586	0.6454	0.6322	1.16	5.98

Table E.2: Sample quartiles of $\Delta_{home\ DS}$ [hours] home-nights are randomly selected and equal to the number of travel-nights on weekends

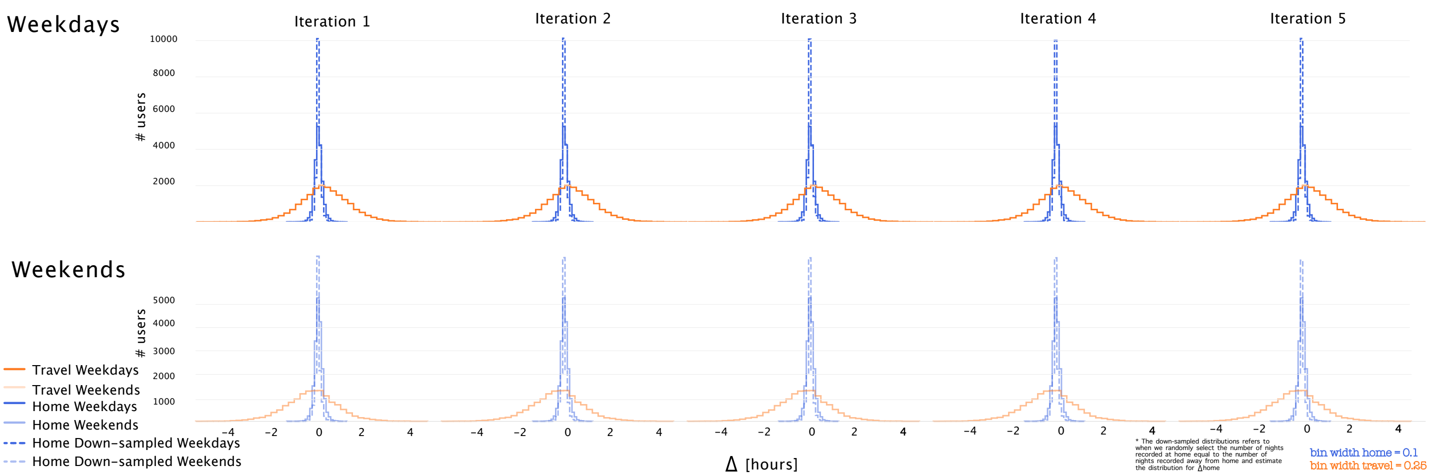


Figure E.1: Distributions of $\Delta_{home\ DS}$ on weekdays and weekends (from the five iterations described above) with Δ_{home} and Δ_{travel}

References

[1] Vijay Kumar Chattu, MD Manzar, Soosanna Kumary, Deepa Burman, David Warren Spence, Seithikurippu R Pandi-Perumal, et al. The global problem of insufficient sleep and its serious public health implications. In *Healthcare*, volume 7, page 1. Multidisciplinary Digital Publishing Institute, 2019.

[2] Neil J Douglas, Stephen Thomas, and Mohammed A Jan. Clinical value of polysomnography. *The Lancet*, 339(8789):347–350, 1992.

[3] Eugene Aserinsky, Nathaniel Kleitman, et al. Regularly occurring periods of eye motility, and concomitant phenomena, during sleep. *Science*, 118(3062):273–274, 1953.

[4] Eugene Aserinsky and Nathaniel Kleitman. Two types of ocular motility occurring in sleep. *Journal of applied physiology*, 8(1):1–10, 1955.

[5] Andrew S Tubbs, Hannah K Dollish, Fabian Fernandez, and Michael A Grandner. The basics of sleep physiology and behavior. In *Sleep and health*, pages 3–10. Elsevier, 2019.

[6] SLEEP COMPUTING COMMITTEE OF THE JAPANESE SOCIETY OF SLEEP RESEARCH SOCIETY (JSSR):, Tadao Hori, Yoshio Sugita, Einosuke Koga, Shuichiro Shirakawa, Katuhiro Inoue, Sunao Uchida, Hiroo Kuwahara, Masako Kousaka, Toshinori Kobayashi, et al. Proposed supplements and amendments to 'a manual of standardized terminology, techniques and scoring system for sleep stages of human subjects', the rechtschaffen & kales (1968) standard. *Psychiatry and clinical neurosciences*, 55(3):305–310, 2001.

[7] Richard B Berry, Rita Brooks, Charlene Gamaldo, Susan M Harding, Robin M Lloyd, Stuart F Quan, Matthew T Troester, and Bradley V Vaughn. Aasm scoring manual updates for 2017 (version 2.4), 2017.

[8] Tarja Porkka-Heiskanen, Robert E Strecker, Mahesh Thakkar, Alvhild A Bjørkum, Robert W Greene, and Robert W McCarley. Adenosine: a mediator of the sleep-inducing effects of prolonged wakefulness. *Science*, 276(5316):1265–1268, 1997.

[9] Tarja Porkka-Heiskanen, Lauri Alanko, Anna Kalinchuk, and Dag Stenberg. Adenosine and sleep. *Sleep medicine reviews*, 6(4):321–332, 2002.

[10] Derk-Jan Dijk. Regulation and functional correlates of slow wave sleep. *Journal of Clinical Sleep Medicine*, 5(2 suppl):S6–S15, 2009.

[11] Pierre Maquet, Christian Degueldre, Guy Delfiore, Joël Aerts, Jean-Marie Péters, André Luxen, and Georges Franck. Functional neuroanatomy of human slow wave sleep. *Journal of Neuroscience*, 17(8):2807–2812, 1997.

[12] Matthew Walker. *Why we sleep: The new science of sleep and dreams*. Penguin UK, 2017.

[13] Alexander A Borbély. A two process model of sleep regulation. *Hum neurobiol*, 1(3):195–204, 1982.

[14] Alexander A Borb and Peter Achermann. Sleep homeostasis and models of sleep regulation. *Journal of biological rhythms*, 14(6):559–570, 1999.

[15] Till Roenneberg, Tim Kuehnle, Myriam Juda, Thomas Kantermann, Karla Allebrandt, Marijke Gordijn, and Martha Merrow. Epidemiology of the human circadian clock. *Sleep medicine reviews*, 11(6):429–438, 2007.

[16] Till Roenneberg, Anna Wirz-Justice, and Martha Merrow. Life between clocks: daily temporal patterns of human chronotypes. *Journal of biological rhythms*, 18(1):80–90, 2003.

[17] Rütger A Wever. *The circadian system of man: results of experiments under temporal isolation*. Springer Science & Business Media, 2013.

[18] Judith A Floyd, S Marshall Medler, Joel W Ager, and James J Janisse. Age-related changes in initiation and maintenance of sleep: a meta-analysis. *Research in nursing & health*, 23(2):106–117, 2000.

[19] Maurice M Ohayon, Mary A Carskadon, Christian Guilleminault, and Michael V Vitiello. Meta-analysis of quantitative sleep parameters from childhood to old age in healthy individuals: developing normative sleep values across the human lifespan. *Sleep*, 27(7):1255–1273, 2004.

[20] Judith A Floyd, James J Janisse, Elizabeth S Jenuwine, and Joel W Ager. Changes in rem-sleep percentage over the adult lifespan. *Sleep*, 30(7):829–836, 2007.

[21] Conor J Wild, Emily S Nichols, Michael E Battista, Bobby Stojanoski, and Adrian M Owen. Dissociable effects of self-reported daily sleep duration on high-level cognitive abilities. *Sleep*, 41(12):zsy182, 2018.

[22] Andrew Gadie, Meredith Shafto, Yue Leng, Rogier A Kievit, et al. How are age-related differences in sleep quality associated with health outcomes? an epidemiological investigation in a uk cohort of 2406 adults. *BMJ open*, 7(7):e014920, 2017.

[23] Bryce A Mander, Joseph R Winer, and Matthew P Walker. Sleep and human aging. *Neuron*, 94(1):19–36, 2017.

[24] Edward F Pace-Schott and Rebecca MC Spencer. Age-related changes in the cognitive function of sleep. In *Progress in brain research*, volume 191, pages 75–89. Elsevier, 2011.

[25] Anna A Kondratova and Roman V Kondratov. The circadian clock and pathology of the ageing brain. *Nature Reviews Neuroscience*, 13(5):325–335, 2012.

[26] Sigga Svala Jonasdottir, Kelton Minor, and Sune Lehmann. Gender differences in nighttime sleep patterns and variability across the adult lifespan: a global-scale wearables study. *Sleep*, 2020.

[27] Ana Adan, Simon N Archer, Maria Paz Hidalgo, Lee Di Milia, Vincenzo Natale, and Christoph Randler. Circadian typology: a comprehensive review. *Chronobiology international*, 29(9):1153–1175, 2012.

[28] Thomas Kantermann, Haein Sung, and Helen J Burgess. Comparing the morningness-eveningness questionnaire and munich chronotype questionnaire to the dim light melatonin onset. *Journal of biological rhythms*, 30(5):449–453, 2015.

[29] Gianina Luca, José Haba Rubio, Daniela Andries, Nadia Tobback, Peter Vollenweider, Gérard Waeber, Pedro Marques Vidal, Martin Preisig, Raphaël Heinzer, and Mehdi Tafti. Age and gender variations of sleep in subjects without sleep disorders. *Annals of medicine*, 47(6):482–491, 2015.

[30] L Kuula, M Gradisar, K Martinmäki, C Richardson, D Bonnar, K Bartel, C Lang, L Leinonen, and AK Pesonen. Using big data to explore worldwide trends in objective sleep in the transition to adulthood. *Sleep Medicine*, 62:69–76, 2019.

[31] Bei Bei, Joshua F Wiley, John Trinder, and Rachel Manber. Beyond the mean: a systematic review on the correlates of daily intraindividual variability of sleep/wake patterns. *Sleep medicine reviews*, 28:108–124, 2016.

[32] Joshua J Gooley. How much day-to-day variability in sleep timing is unhealthy? *Sleep*, 39(2):269–270, 2016.

[33] Sakari Lemola, Thomas Ledermann, and Elliot M Friedman. Variability of sleep duration is related to subjective sleep quality and subjective well-being: an actigraphy study. *PloS one*, 8(8):e71292, 2013.

[34] Sanjay R Patel, Amanda L Hayes, Terri Blackwell, Daniel S Evans, Sonia Ancoli-Israel, Yun K Wing, and Katie L Stone. The association between sleep patterns and obesity in older adults. *International journal of obesity*, 38(9):1159–1164, 2014.

[35] Danica C Slavish, Daniel J Taylor, and Kenneth L Lichstein. Intraindividual variability in sleep and comorbid medical and mental health conditions. *Sleep*, 42(6):zsz052, 2019.

[36] Dorothee Fischer, David A Lombardi, Helen Marucci-Wellman, and Till Roenneberg. Chronotypes in the us–influence of age and sex. *PloS one*, 12(6):e0178782, 2017.

[37] Haley R Dillon, Kenneth L Lichstein, Natalie D Dautovich, Daniel J Taylor, Brant W Riedel, and Andrew J Bush. Variability in self-reported normal sleep across the adult age span. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 70(1):46–56, 2015.

[38] Kristen L Knutson, Paul J Rathouz, Lijing L Yan, Kiang Liu, and Diane S Lauderdale. Intra-individual daily and yearly variability in actigraphically recorded sleep measures: the cardia study. *Sleep*, 30(6):793–796, 2007.

[39] Till Roenneberg, Tim Kuehnle, Peter P Pramstaller, Jan Ricken, Miriam Havel, Angelika Guth, and Martha Merrow. A marker for the end of adolescence. *Current biology*, 14(24):R1038–R1039, 2004.

[40] Jeanne F Duffy, Sean W Cain, Anne-Marie Chang, Andrew JK Phillips, Mirjam Y Münch, Claude Gronfier, James K Wyatt, Derk-Jan Dijk, Kenneth P Wright, and Charles A Czeisler. Sex difference in the near-24-hour intrinsic period of the human circadian timing system. *Proceedings of the National Academy of Sciences*, 108(Supplement 3):15602–15608, 2011.

[41] Sean W Cain, Christopher F Dennison, Jamie M Zeitzer, Aaron M Guzik, Sat Bir S Khalsa, Nayantara Santhi, Martin W Schoen, Charles A Czeisler, and Jeanne F Duffy. Sex differences in phase angle of entrainment and melatonin amplitude in humans. *Journal of biological rhythms*, 25(4):288–296, 2010.

[42] Till Roenneberg, Luísa K Pilz, Giulia Zerbini, and Eva C Winnebeck. Chronotype and social jetlag: a (self-) critical review. *Biology*, 8(3):54, 2019.

[43] Ana Adan and Vincenzo Natale. Gender differences in morningness–eveningness preference. *Chronobiology international*, 19(4):709–720, 2002.

[44] Till Roenneberg, Karla V Allebrandt, Martha Merrow, and Céline Vetter. Social jetlag and obesity. *Current Biology*, 22(10):939–943, 2012.

[45] Lorenzo Tonetti, Marco Fabbri, and Vincenzo Natale. Sex difference in sleep-time preference and sleep need: A cross-sectional survey among italian pre-adolescents, adolescents, and adults. *Chronobiology international*, 25(5):745–759, 2008.

[46] Ju Lynn Ong, Jesisca Tandi, Amiya Patanaik, June C Lo, and Michael WL Chee. Large-scale data from wearables reveal regional disparities in sleep patterns that persist across age and sex. *Scientific reports*, 9(1):1–8, 2019.

[47] Joyce T Bromberger, Karen A Matthews, Lewis H Kuller, Rena R Wing, Elaine N Meilahn, and Pam Plantinga. Prospective study of the determinants of age at menopause. *American journal of epidemiology*, 145(2):124–133, 1997.

[48] World Health Organization. Obesity: preventing and managing the global epidemic. 2000.

[49] Katherine M Flegal, Margaret D Carroll, Robert J Kuczmarski, and Clifford L Johnson. Overweight and obesity in the united states: prevalence and trends, 1960–1994. *International journal of obesity*, 22(1):39–47, 1998.

[50] Anne Berghöfer, Tobias Pischon, Thomas Reinhold, Caroline M Apovian, Arya M Sharma, and Stefan N Willich. Obesity prevalence from a european perspective: a systematic review. *BMC public health*, 8(1):1–10, 2008.

[51] Philip T James, Rachel Leach, Eleni Kalamara, and Maryam Shayeghi. The worldwide obesity epidemic. *Obesity research*, 9(S11):228S–233S, 2001.

[52] Rajeev Ahirwar and Prakash Ranjan Mondal. Prevalence of obesity in india: A systematic review. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 13(1):318–321, 2019.

[53] Earl S Ford, Timothy J Cunningham, and Janet B Croft. Trends in self-reported sleep duration among us adults from 1985 to 2012. *Sleep*, 38(5):829–832, 2015.

[54] Matthew P Walker. A societal sleep prescription. *Neuron*, 103(4):559–562, 2019.

[55] Charles A Czeisler. Perspective: casting light on sleep deficiency. *Nature*, 497(7450):S13–S13, 2013.

[56] Saverio Stranges, William Tigbe, Francesc Xavier Gómez-Olivé, Margaret Thorogood, and Ngianga-Bakwin Kandala. 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*, 35(8):1173–1181, 2012.

[57] Erkki Kronholm, Timo Partonen, Tiina Laatikainen, Markku Peltonen, Mikko Härmä, Christer Hublin, Jaako Kaprio, Arja R Aro, Markku Partinen, Mikael Fogelholm, et al. Trends in self-reported sleep duration and insomnia-related symptoms in finland from 1972 to 2005: a comparative review and re-analysis of finnish population samples. *Journal of sleep research*, 17(1):54–62, 2008.

[58] Kristen L Knutson, Eve Van Cauter, Paul J Rathouz, Thomas DeLeire, and Diane S Lauderdale. Trends in the prevalence of short sleepers in the usa: 1975–2006. *Sleep*, 33(1):37–45, 2010.

[59] Francesco P Cappuccio, Frances M Taggart, Ngianga-Bakwin Kandala, Andrew Currie, Ed Peile, Saverio Stranges, and Michelle A Miller. Meta-analysis of short sleep duration and obesity in children and adults. *Sleep*, 31(5):619–626, 2008.

[60] LS Nielsen, KV Danielsen, and TIA Sørensen. Short sleep duration as a possible cause of obesity: critical analysis of the epidemiological evidence. *Obesity Reviews*, 12(2):78–92, 2011.

[61] Xiaoli Chen, May A Beydoun, and Youfa Wang. Is sleep duration associated with childhood obesity? a systematic review and meta-analysis. *Obesity*, 16(2):265, 2008.

[62] Karine Spiegel, Esra Tasali, Rachel Leproult, and Eve Van Cauter. Effects of poor and short sleep on glucose metabolism and obesity risk. *Nature Reviews Endocrinology*, 5(5):253, 2009.

[63] Sanjay R Patel and Frank B Hu. Short sleep duration and weight gain: a systematic review. *Obesity*, 16(3):643–653, 2008.

[64] Shahrad Taheri, Ling Lin, Diane Austin, Terry Young, and Emmanuel Mignot. Short sleep duration is associated with reduced leptin, elevated ghrelin, and increased body mass index. *PLoS Med*, 1(3):e62, 2004.

[65] Kristen L Knutson, Karine Spiegel, Plamen Penev, and Eve Van Cauter. The metabolic consequences of sleep deprivation. *Sleep medicine reviews*, 11(3):163–178, 2007.

[66] Nathaniel F Watson, Dedra Buchwald, Michael V Vitiello, Carolyn Noonan, and Jack Goldberg. A twin study of sleep duration and body mass index. *Journal of Clinical Sleep Medicine*, 6(1):11–17, 2010.

[67] Patrice L Capers, Aaron D Fobian, Kathryn A Kaiser, Rohit Borah, and David B Allison. A systematic review and meta-analysis of randomized controlled trials of the impact of sleep duration on adiposity and components of energy balance. *Obesity reviews*, 16(9):771–782, 2015.

[68] M-P St-Onge. Sleep–obesity relation: Underlying mechanisms and consequences for treatment. *Obesity Reviews*, 18:34–39, 2017.

[69] Rütger A Wever. Light effects on human circadian rhythms: a review of recent andechs experiments. *Journal of Biological Rhythms*, 4(2):49–73, 1989.

[70] Charles A Czeisler, Gary S Richardson, Janet C Zimmerman, Martin C Moore-Ede, and Elliot D Weitzman. Entrainment of human orcadian rhythms by light-dark cycles: A reassessment. *Photochemistry and photobiology*, 34(2):239–247, 1981.

[71] Till Roenneberg, Serge Daan, and Martha Merrow. The art of entrainment. *Journal of biological rhythms*, 18(3):183–194, 2003.

[72] Patricia L Turner, Eus JW Van Someren, and Martin A Mainster. The role of environmental light in sleep and health: effects of ocular aging and cataract surgery. *Sleep medicine reviews*, 14(4):269–280, 2010.

[73] David R Weaver. The suprachiasmatic nucleus: a 25-year retrospective. *Journal of biological rhythms*, 13(2):100–112, 1998.

[74] Melanie S Freedman, Robert J Lucas, Bobby Soni, Malcolm von Schantz, Marta Muñoz, Zoë David-Gray, and Russell Foster. Regulation of mammalian circadian behavior by non-rod, non-cone, ocular photoreceptors. *Science*, 284(5413):502–504, 1999.

[75] Iain M Mclntyre, Trevor R Norman, Graham D Burrows, and Stuart M Armstrong. Human melatonin suppression by light is intensity dependent. *Journal of pineal research*, 6(2):149–156, 1989.

[76] Charles A Czeisler, Theresa L Shanahan, Elizabeth B Klerman, Heinz Martens, Daniel J Brotman, Jonathan S Emens, Torsten Klein, and Joseph F Rizzo. Suppression of melatonin secretion in some blind patients by exposure to bright light. *New England Journal of Medicine*, 332(1):6–11, 1995.

[77] Anne-Marie Chang, Daniel Aeschbach, Jeanne F Duffy, and Charles A Czeisler. Evening use of light-emitting ereaders negatively affects sleep, circadian timing, and next-morning alertness. *Proceedings of the National Academy of Sciences*, 112(4):1232–1237, 2015.

[78] Brittany Wood, Mark S Rea, Barbara Plitnick, and Mariana G Figueiro. Light level and duration of exposure determine the impact of self-luminous tablets on melatonin suppression. *Applied ergonomics*, 44(2):237–240, 2013.

[79] Michael A. Grandner. *Sleep and Health*. 2019.

[80] Scott S Campbell and Roger J Broughton. Rapid decline in body temperature before sleep: fluffing the physiological pillow? *Chronobiology international*, 11(2):126–131, 1994.

[81] Kelton Minor, Andreas Bjerre-Nielsen, Sigga Svala Jonasdottir, Sune Lehmann, and Nick Obradovich. Ambient heat and human sleep. *arXiv preprint arXiv:2011.07161*, 2020.

[82] Nick Obradovich, Robyn Migliorini, Sara C Mednick, and James H Fowler. Nighttime temperature and human sleep loss in a changing climate. *Science advances*, 3(5):e1601555, 2017.

[83] Kazue Okamoto-Mizuno and Koh Mizuno. Effects of thermal environment on sleep and circadian rhythm. *Journal of physiological anthropology*, 31(1):1–9, 2012.

[84] Daniel Monsivais, Kunal Bhattacharya, Asim Ghosh, Robin IM Dunbar, and Kimmo Kaski. Seasonal and geographical impact on human resting periods. *Scientific reports*, 7(1):1–10, 2017.

[85] Gandhi Yetish, Hillard Kaplan, Michael Gurven, Brian Wood, Herman Pontzer, Paul R Manger, Charles Wilson, Ronald McGregor, and Jerome M Siegel. Natural sleep and its

seasonal variations in three pre-industrial societies. *Current Biology*, 25(21):2862–2868, 2015.

[86] Olivia J Walch, Amy Cochran, and Daniel B Forger. A global quantification of "normal" sleep schedules using smartphone data. *Science advances*, 2(5):e1501705, 2016.

[87] Pablo E Brockmann, David Gozal, Luis Villarroel, Felipe Damiani, Felipe Nuñez, and Christian Cajochen. Geographic latitude and sleep duration: A population-based survey from the tropic of capricorn to the antarctic circle. *Chronobiology international*, 34(3):373–381, 2017.

[88] Christoph Randler and Arash Rahafar. Latitude affects morningness-eveningness: evidence for the environment hypothesis based on a systematic review. *Scientific Reports*, 7(1):1–6, 2017.

[89] June C Lo, Ruth LF Leong, Kep-Kee Loh, Derk-Jan Dijk, and Michael WL Chee. Young adults' sleep duration on work days: differences between east and west. *Frontiers in neurology*, 5:81, 2014.

[90] Yoko Komada, Isa Okajima, Shingo Kitamura, and Yuichi Inoue. A survey on social jetlag in japan: a nationwide, cross-sectional internet survey. *Sleep and Biological Rhythms*, 17(4):417–422, 2019.

[91] Zhongxing Zhang, Christian Cajochen, and Ramin Khatami. Social jetlag and chronotypes in the chinese population: analysis of data recorded by wearable devices. *Journal of medical internet research*, 21(6):e13482, 2019.

[92] Brigitte Steger. Getting away with sleep—social and cultural aspects of dozing in parliament. *Social Science Japan Journal*, 6(2):181–197, 2003.

[93] Michael A Grandner, Natasha J Williams, Kristen L Knutson, Dorothy Roberts, and Girardin Jean-Louis. Sleep disparity, race/ethnicity, and socioeconomic position. *Sleep medicine*, 18:7–18, 2016.

[94] Les A Gellis, Kenneth L Lichstein, Isabel C Scarinci, H Heith Durrence, Daniel J Taylor, Andrew J Bush, and Brant W Riedel. Socioeconomic status and insomnia. *Journal of abnormal psychology*, 114(1):111, 2005.

[95] Lianne M Tomfohr, Sonia Ancoli-Israel, and Joel E Dimsdale. Childhood socioeconomic status and race are associated with adult sleep. *Behavioral sleep medicine*, 8(4):219–230, 2010.

[96] Lauren Hale and D Phuong Do. Racial differences in self-reports of sleep duration in a population-based study. *Sleep*, 30(9):1096–1103, 2007.

[97] Sherman A James. John henryism and the health of african-americans. 1994.

[98] Shawn D Youngstedt and Christopher E Kline. Epidemiology of exercise and sleep. *Sleep and biological rhythms*, 4(3):215–221, 2006.

[99] Duane L Sherrill, Kimberly Kotchou, and Stuart F Quan. Association of physical activity and human sleep disorders. *Archives of internal medicine*, 158(17):1894–1898, 1998.

[100] Bang Zheng, Canqing Yu, Liling Lin, Huaidong Du, Jun Lv, Yu Guo, Zheng Bian, Yiping Chen, Min Yu, Jianguo Li, et al. Associations of domain-specific physical activities with insomnia symptoms among 0.5 million chinese adults. *Journal of sleep research*, 26(3):330–337, 2017.

[101] M Alexandra Kredlow, Michelle C Capozzoli, Bridget A Hearon, Amanda W Calkins, and Michael W Otto. The effects of physical activity on sleep: a meta-analytic review. *Journal of behavioral medicine*, 38(3):427–449, 2015.

[102] Eizaburo Tanaka, Hiroshi Yatsuya, Mayu Uemura, Chiyoe Murata, Rei Otsuka, Hideaki Toyoshima, Koji Tamakoshi, Satoshi Sasaki, Leo Kawaguchi, and Atsuko Aoyama. Associations of protein, fat, and carbohydrate intakes with insomnia symptoms among middle-aged japanese workers. *Journal of epidemiology*, 23(2):132–138, 2013.

[103] Xiao Tan, Markku Alén, Shu Mei Cheng, Tuija M Mikkola, Jarkko Tenhunen, Arja Lyytikäinen, Petri Wiklund, Fengyu Cong, Antti Saarinen, Ina Tarkka, et al. Associations of disordered sleep with body fat distribution, physical activity and diet among overweight middle-aged men. *Journal of sleep research*, 24(4):414–424, 2015.

[104] Ryoko Katagiri, Keiko Asakura, Satomi Kobayashi, Hitomi Suga, and Satoshi Sasaki. Low intake of vegetables, high intake of confectionary, and unhealthy eating habits are associated with poor sleep quality among middle-aged female japanese workers. *Journal of occupational health*, pages 14–0051, 2014.

[105] Kelly G Baron, Kathryn J Reid, Andrew S Kern, and Phyllis C Zee. Role of sleep timing in caloric intake and bmi. *Obesity*, 19(7):1374–1381, 2011.

[106] Marta Garaulet, Purificación Gómez-Abellán, Juan J Alburquerque-Béjar, Yu-Chi Lee, Jose M Ordovás, and Frank AJL Scheer. Timing of food intake predicts weight loss effectiveness. *International journal of obesity*, 37(4):604–611, 2013.

[107] Andrea M Spaeth, David F Dinges, and Namni Goel. Effects of experimental sleep restriction on weight gain, caloric intake, and meal timing in healthy adults. *Sleep*, 36(7):981–990, 2013.

[108] ITALO Biaggioni, SUBI Paul, ANDR Puckett, and CARM Arzubiaga. Caffeine and theophylline as adenosine receptor antagonists in humans. *Journal of Pharmacology and Experimental Therapeutics*, 258(2):588–593, 1991.

[109] Joaquim A Ribeiro and Ana M Sebastiao. Caffeine and adenosine. *Journal of Alzheimer's Disease*, 20(s1):S3–S15, 2010.

[110] Hans-Peter Landolt, Esther Werth, Alexander A Borbély, and Derk-Jan Dijk. Caffeine intake (200 mg) in the morning affects human sleep and eeg power spectra at night. *Brain research*, 675(1-2):67–74, 1995.

[111] Michael H Bonnet and DL Arand. Caffeine use as a model of acute and chronic insomnia. *Sleep*, 15(6):526–536, 1992.

[112] Timothy Roehrs and Thomas Roth. Caffeine: sleep and daytime sleepiness. *Sleep medicine reviews*, 12(2):153–162, 2008.

[113] Christopher Drake, Timothy Roehrs, John Shambroom, and Thomas Roth. Caffeine effects on sleep taken 0, 3, or 6 hours before going to bed. *Journal of Clinical Sleep Medicine*, 9(11):1195–1200, 2013.

[114] Matthew A Christensen, Laura Bettencourt, Leanne Kaye, Sai T Moturu, Kaylin T Nguyen, Jeffrey E Olgin, Mark J Pletcher, and Gregory M Marcus. Direct measurements of smartphone screen-time: relationships with demographics and sleep. *PloS one*, 11(11):e0165331, 2016.

[115] Vivek Arun Kumar, Vigneshvar Chandrasekaran, and Hema Brahadeeswari. Prevalence of smartphone addiction and its effects on sleep quality: A cross-sectional study among medical students. *Industrial psychiatry journal*, 28(1):82, 2019.

[116] Klodiana Lanaj, Russell E Johnson, and Christopher M Barnes. Beginning the workday yet already depleted? consequences of late-night smartphone use and sleep. *Organizational Behavior and Human Decision Processes*, 124(1):11–23, 2014.

[117] Yolanda Linda Reid Chassiakos, Jenny Radesky, Dimitri Christakis, Megan A Moreno, Corinn Cross, et al. Children and adolescents and digital media. *Pediatrics*, 138(5), 2016.

[118] Christian Cajochen, Sylvia Frey, Doreen Anders, Jakub Späti, Matthias Bues, Achim Pross, Ralph Mager, Anna Wirz-Justice, and Oliver Stefani. Evening exposure to a light-emitting diodes (led)-backlit computer screen affects circadian physiology and cognitive performance. *Journal of applied physiology*, 2011.

[119] Monica P Mallampalli and Christine L Carter. Exploring sex and gender differences in sleep health: a society for women's health research report. *Journal of women's health*, 23(7):553–562, 2014.

[120] Martino F Pengo, Christine H Won, and Ghada Bourjeily. Sleep in women across the life span. *Chest*, 154(1):196–206, 2018.

[121] David Richter, Michael D Krämer, Nicole KY Tang, Hawley E Montgomery-Downs, and Sakari Lemola. Long-term effects of pregnancy and childbirth on sleep satisfaction and duration of first-time and experienced mothers and fathers. *Sleep*, 42(4), 2019.

[122] Gunn Marit Aasvang, Britt Øverland, Reidun Ursin, and Torbjørn Moum. A field study of effects of road traffic and railway noise on polysomnographic sleep parameters. *The Journal of the Acoustical Society of America*, 129(6):3716–3726, 2011.

[123] Mathias Basner, Uwe Müller, and Eva-Maria Elmenhorst. Single and combined effects of air, road, and rail traffic noise on sleep and recuperation. *Sleep*, 34(1):11–23, 2011.

[124] Eva-Maria Elmenhorst, Sibylle Pennig, Vinzent Rolny, Julia Quehl, Uwe Mueller, Hartmut Maaß, and Mathias Basner. Examining nocturnal railway noise and aircraft noise in the field: Sleep, psychomotor performance, and annoyance. *Science of the Total Environment*, 424:48–56, 2012.

[125] HW Agnew Jr, Wilse B Webb, and Robert L Williams. The first night effect: an eeg studyof sleep. *Psychophysiology*, 2(3):263–266, 1966.

[126] Masako Tamaki, Hiroshi Nittono, Mitsuo Hayashi, and Tadao Hori. Examination of the first-night effect during the sleep-onset period. *Sleep*, 28(2):195–202, 2005.

[127] Masako Tamaki, Ji Won Bang, Takeo Watanabe, and Yuka Sasaki. Night watch in one brain hemisphere during sleep associated with the first-night effect in humans. *Current biology*, 26(9):1190–1194, 2016.

[128] J Waterhouse, T Reilly, and B Edwards. The stress of travel. *Journal of sports sciences*, 22(10):946–966, 2004.

[129] Michele Lastella, Gregory D Roach, and Charli Sargent. Travel fatigue and sleep/wake behaviors of professional soccer players during international competition. *Sleep health*, 5(2):141–147, 2019.

[130] Jim Waterhouse, Thomas Reilly, Greg Atkinson, and Ben Edwards. Jet lag: trends and coping strategies. *The lancet*, 369(9567):1117–1129, 2007.

[131] Venkataramanujan Srinivasan, D Warren Spence, Seithikurippu R Pandi-Perumal, Ilya Trakht, and Daniel P Cardinali. Jet lag: therapeutic use of melatonin and possible application of melatonin analogs. *Travel Medicine and Infectious Disease*, 6(1-2):17–28, 2008.

[132] Timothy I Morgenthaler, Teofilo Lee-Chiong, Cathy Alessi, Leah Friedman, R Nisha Aurora, Brian Boehlecke, Terry Brown, Andrew L Chesson Jr, Vishesh Kapur, Rama Maganti, et al. Practice parameters for the clinical evaluation and treatment of circadian rhythm sleep disorders. *Sleep*, 30(11):1445–1459, 2007.

[133] Michael R Irwin. Why sleep is important for health: a psychoneuroimmunology perspective. *Annual review of psychology*, 66:143–172, 2015.

[134] Andrea N Goldstein and Matthew P Walker. The role of sleep in emotional brain function. *Annual review of clinical psychology*, 10:679–708, 2014.

[135] Rebecca A Bernert, Joanne S Kim, Naomi G Iwata, and Michael L Perlis. Sleep disturbances as an evidence-based suicide risk factor. *Current psychiatry reports*, 17(3):15, 2015.

[136] Faith S Luyster, Patrick J Strollo, Phyllis C Zee, and James K Walsh. Sleep: a health imperative. *Sleep*, 35(6):727–734, 2012.

[137] Christopher Ryan King, Kristen L Knutson, Paul J Rathouz, Steve Sidney, Kiang Liu, and Diane S Lauderdale. Short sleep duration and incident coronary artery calcification. *Jama*, 300(24):2859–2866, 2008.

[138] Francesco P Cappuccio, Daniel Cooper, Lanfranco D'elia, Pasquale Strazzullo, and Michelle A Miller. Sleep duration predicts cardiovascular outcomes: a systematic review and meta-analysis of prospective studies. *European heart journal*, 32(12):1484–1492, 2011.

[139] June J Pilcher and Allen I Huffcutt. Effects of sleep deprivation on performance: a meta-analysis. *Sleep*, 19(4):318–326, 1996.

[140] Edward O Bixler, Anthony Kales, Constantine R Soldatos, Joyce D Kales, and S Healey. Prevalence of sleep disorders in the los angeles metropolitan area. *The American journal of psychiatry*, 1979.

[141] Peretz Lavie. Sleep habits and sleep disturbances in industrial workers in israel: main findings and some characteristics of workers complaining of excessive daytime sleepiness. *Sleep*, 4(2):147–158, 1981.

[142] M Karvonen, MI Mikheev, World Health Organization, et al. *Epidemiology of occupational health*. World Health Organization. Regional Office for Europe, 1986.

[143] Dustin T Duncan, Ichiro Kawachi, and Susan Redline. *The Social Epidemiology of Sleep*. Oxford University Press, 2019.

[144] Diane S Lauderdale, Kristen L Knutson, Lijing L Yan, Paul J Rathouz, Stephen B Hulley, Steve Sidney, and Kiang Liu. Objectively measured sleep characteristics among early-middle-aged adults: the cardia study. *American journal of epidemiology*, 164(1):5–16, 2006.

[145] Yeonsu Song, Sonia Ancoli-Israel, Cora E Lewis, Susan Redline, Stephanie L Harrison, and Katie L Stone. The association of race/ethnicity with objectively measured sleep characteristics in older men. *Behavioral Sleep Medicine*, 10(1):54–69, 2012.

[146] Katie L Stone, Sonia Ancoli-Israel, Terri Blackwell, Kristine E Ensrud, Jane A Cauley, Susan Redline, Teresa A Hillier, Jennifer Schneider, David Claman, and Steven R Cummings. Actigraphy-measured sleep characteristics and risk of falls in older women. *Archives of internal medicine*, 168(16):1768–1775, 2008.

[147] Diane S Lauderdale, Kristen L Knutson, Lijing L Yan, Kiang Liu, and Paul J Rathouz. Sleep duration: how well do self-reports reflect objective measures? the cardia sleep study. *Epidemiology (Cambridge, Mass.)*, 19(6):838, 2008.

[148] Chandra L Jackson, Sanjay R Patel, W Braxton Jackson, Pamela L Lutsey, and Susan Redline. Agreement between self-reported and objectively measured sleep duration among white, black, hispanic, and chinese adults in the united states: Multi-ethnic study of atherosclerosis. *Sleep*, 41(6):zsy057, 2018.

[149] Elizabeth M Cespedes, Frank B Hu, Susan Redline, Bernard Rosner, Carmela Alcantara, Jianwen Cai, Martica H Hall, Jose S Loredo, Yasmin Mossavar-Rahmani, Alberto R Ramos, et al. Comparison of self-reported sleep duration with actigraphy: results from the hispanic community health study/study of latinos sueño ancillary study. *American journal of epidemiology*, 183(6):561–573, 2016.

[150] Katarina Aili, Sofia Åström-Paulsson, Ulrich Stoetzer, Magnus Svartengren, and Lena Hillert. Reliability of actigraphy and subjective sleep measurements in adults: the design of sleep assessments. *Journal of Clinical Sleep Medicine*, 13(1):39–47, 2017.

[151] Marcela Zambrim Campanini, Esther Lopez-Garcia, Fernando Rodríguez-Artalejo, Alberto Durán González, Selma Maffei Andrade, and Arthur Eumann Mesas. Agreement between sleep diary and actigraphy in a highly educated brazilian population. *Sleep medicine*, 35:27–34, 2017.

[152] James E Gangwisch, Steven B Heymsfield, Bernadette Boden-Albala, Ruud M Buijs, Felix Kreier, Thomas G Pickering, Andrew G Rundle, Gary K Zammit, and Dolores Malaspina. Short sleep duration as a risk factor for hypertension: analyses of the first national health and nutrition examination survey. *hypertension*, 47(5):833–839, 2006.

[153] Jennifer Girschik, Lin Fritschi, Jane Heyworth, and Flavie Waters. Validation of self-reported sleep against actigraphy. *Journal of epidemiology*, 22(5):462–468, 2012.

[154] Erin O'Brien, Chantelle Hart, and Rena R Wing. Discrepancies between self-reported usual sleep duration and objective measures of total sleep time in treatment-seeking overweight and obese individuals. *Behavioral sleep medicine*, 14(5):539–549, 2016.

[155] Christopher B Miller, Christopher J Gordon, Leanne Toubia, Delwyn J Bartlett, Ronald R Grunstein, Angela L D'Rozario, and Nathaniel S Marshall. Agreement between simple questions about sleep duration and sleep diaries in a large online survey. *Sleep Health*, 1(2):133–137, 2015.

[156] Lianne M Kurina, Martha K McClintock, Jen-Hao Chen, Linda J Waite, Ronald A Thisted, and Diane S Lauderdale. Sleep duration and all-cause mortality: a critical review of measurement and associations. *Annals of epidemiology*, 23(6):361–370, 2013.

[157] Daniel J Buysse, Charles F Reynolds III, Timothy H Monk, Susan R Berman, and David J Kupfer. The pittsburgh sleep quality index: a new instrument for psychiatric practice and research. *Psychiatry research*, 28(2):193–213, 1989.

[158] Jim A Horne and Olov Östberg. A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International journal of chronobiology*, 1976.

[159] Jason C Cole, Sarosh J Motivala, Daniel J Buysse, Michael N Oxman, Myron J Levin, and Michael R Irwin. Validation of a 3-factor scoring model for the pittsburgh sleep quality index in older adults. *Sleep*, 29(1):112–116, 2006.

[160] Sherry A Beaudreau, Adam P Spira, Anita Stewart, Eric J Kezirian, Li-Yung Lui, Kristine Ensrud, Susan Redline, Sonia Ancoli-Israel, Katie L Stone, Study of Osteoporotic Fractures, et al. Validation of the pittsburgh sleep quality index and the epworth sleepiness scale in older black and white women. *Sleep medicine*, 13(1):36–42, 2012.

[161] Michael A Grandner, Daniel F Kripke, In-Young Yoon, and Shawn D Youngstedt. Criterion validity of the pittsburgh sleep quality index: Investigation in a non-clinical sample. *Sleep and biological rhythms*, 4(2):129–136, 2006.

[162] Jacques Taillard, Pierre Philip, Jean-François Chastang, and Bernard Bioulac. Validation of horne and ostberg morningness-eveningness questionnaire in a middle-aged population of french workers. *Journal of biological rhythms*, 19(1):76–86, 2004.

[163] Eirunn Thun, Bjørn Bjorvatn, Theresa Osland, Vidar Martin Steen, Børge Sivertsen, Torill Johansen, Tom Halvor Lilleholt, Idalill Udnes, Inger Hilde Nordhus, and Ståle Pallesen. An actigraphic validation study of seven morningness-eveningness inventories. *European Psychologist*, 2012.

[164] Andrei Zavada, Marijke CM Gordijn, Domien GM Beersma, Serge Daan, and Till Roenneberg. Comparison of the munich chronotype questionnaire with the horne-östberg's morningness-eveningness score. *Chronobiology international*, 22(2):267–278, 2005.

[165] Nathaniel Kleitman. *Sleep and wakefulness*. University of Chicago Press, 1987.

[166] Michael W Young and Steve A Kay. Time zones: a comparative genetics of circadian clocks. *Nature Reviews Genetics*, 2(9):702–715, 2001.

[167] P Franken and D-J Dijk. Circadian clock genes and sleep homeostasis. *European Journal of Neuroscience*, 29(9):1820–1829, 2009.

[168] Takashi Ebisawa, Makoto Uchiyama, Naofumi Kajimura, Kazuo Mishima, Yuichi Kamei, Masaaki Katoh, Tsuyoshi Watanabe, Masanori Sekimoto, Kayo Shibui, Keiko Kim, et al. Association of structural polymorphisms in the human period3 gene with delayed sleep phase syndrome. *EMBO reports*, 2(4):342–346, 2001.

[169] Kong L Toh, Christopher R Jones, Yan He, Erik J Eide, William A Hinz, David M Virshup, Louis J Ptáček, and Ying-Hui Fu. An hper2 phosphorylation site mutation in familial advanced sleep phase syndrome. *Science*, 291(5506):1040–1043, 2001.

[170] David R Samson, Alyssa N Crittenden, Ibrahim A Mabulla, Audax ZP Mabulla, and Charles L Nunn. Chronotype variation drives night-time sentinel-like behaviour in hunter–gatherers. *Proceedings of the Royal Society B: Biological Sciences*, 284(1858):20170967, 2017.

[171] Rosa Levandovski, Etianne Sasso, and Maria Paz Hidalgo. Chronotype: a review of the advances, limits and applicability of the main instruments used in the literature to assess human phenotype. *Trends in psychiatry and psychotherapy*, 35(1):3–11, 2013.

[172] Charles A Czeisler and JJ Gooley. Sleep and circadian rhythms in humans. In *Cold Spring Harbor symposia on quantitative biology*, volume 72, pages 579–597. Cold Spring Harbor Laboratory Press, 2007.

[173] Kenneth P Wright Jr, Andrew W McHill, Brian R Birks, Brandon R Griffin, Thomas Rusterholz, and Evan D Chinoy. Entrainment of the human circadian clock to the natural light-dark cycle. *Current Biology*, 23(16):1554–1558, 2013.

[174] Andrew D Beale, Mario Pedrazzoli, B Gonçalves Bruno da Silva, Felipe Beijamini, Núbia E Duarte, Kieren J Egan, Kristen L Knutson, Malcolm Von Schantz, and Laura C Roden. Comparison between an african town and a neighbouring village shows delayed, but not decreased, sleep during the early stages of urbanisation. *Scientific reports*, 7(1):1–10, 2017.

[175] Claudia RC Moreno, S Vasconcelos, EC Marqueze, Arne Lowden, B Middleton, FM Fischer, FM Louzada, and DJ Skene. Sleep patterns in amazon rubber tappers with and without electric light at home. *Scientific reports*, 5:14074, 2015.

[176] Horacio O De La Iglesia, Eduardo Fernández-Duque, Diego A Golombek, Norberto Lanza, Jeanne F Duffy, Charles A Czeisler, and Claudia R Valeggia. Access to electric light is associated with shorter sleep duration in a traditionally hunter-gatherer community. *Journal of biological rhythms*, 30(4):342–350, 2015.

[177] Mikhail F Borisenkov, Elena V Perminova, and Anna L Kosova. Chronotype, sleep length, and school achievement of 11-to 23-year-old students in northern european russia. *Chronobiology international*, 27(6):1259–1270, 2010.

[178] Felipe Gutiérrez Carvalho, Maria Paz Hidalgo, and Rosa Levandovski. Differences in circadian patterns between rural and urban populations: an epidemiological study in countryside. *Chronobiology international*, 31(3):442–449, 2014.

[179] Neda Ghotbi, Luísa K Pilz, Eva C Winnebeck, Céline Vetter, Giulia Zerbini, David Lenssen, Giovanni Frighetto, Marco Salamanca, Rodolfo Costa, Sara Montagnese, et al. The μmctq: An ultra-short version of the munich chronotype questionnaire. *Journal of Biological Rhythms*, 35(1):98–110, 2020.

[180] Shingo Kitamura, Akiko Hida, Sayaka Aritake, Shigekazu Higuchi, Minori Enomoto, Mie Kato, Céline Vetter, Till Roenneberg, and Kazuo Mishima. Validity of the japanese version of the munich chronotype questionnaire. *Chronobiology international*, 31(7):845–850, 2014.

[181] Elise R Facer-Childs, Brunno M Campos, Benita Middleton, Debra J Skene, and Andrew P Bagshaw. Circadian phenotype impacts the brain's resting-state functional connectivity, attentional performance, and sleepiness. *Sleep*, 42(5):zsz033, 2019.

[182] Jose Arturo Santisteban, Thomas G Brown, and Reut Gruber. Association between the munich chronotype questionnaire and wrist actigraphy. *Sleep Disorders*, 2018, 2018.

[183] Réka Ágnes Haraszti, Krisztina Ella, Norbert Gyöngyösi, Till Roenneberg, and Krisztina Káldi. Social jetlag negatively correlates with academic performance in undergraduates. *Chronobiology international*, 31(5):603–612, 2014.

[184] Marc Wittmann, Jenny Dinich, Martha Merrow, and Till Roenneberg. Social jetlag: misalignment of biological and social time. *Chronobiology international*, 23(1-2):497–509, 2006.

[185] Eurofund. Fifth european working conditions survey, 2012.

[186] U.S. Bureau of Labor Statistics. Job flexibilities and work schedules summary—2017–2018, 2019.

[187] Suzana Almoosawi, Luigi Palla, Ian Walshe, Snieguole Vingeliene, and Jason G Ellis. Long sleep duration and social jetlag are associated inversely with a healthy dietary pattern in adults: Results from the uk national diet and nutrition survey rolling programme y1–4. *Nutrients*, 10(9):1131, 2018.

[188] Juan F Díaz-Morales and Cristina Escribano. Social jetlag, academic achievement and cognitive performance: Understanding gender/sex differences. *Chronobiology international*, 32(6):822–831, 2015.

[189] Anitra DM Koopman, Simone P Rauh, Esther van 't Riet, Lenka Groeneveld, Amber A Van Der Heijden, Petra J Elders, Jacqueline M Dekker, Giel Nijpels, Joline W Beulens, and Femke Rutters. The association between social jetlag, the metabolic syndrome, and type

2 diabetes mellitus in the general population: the new hoorn study. *Journal of biological rhythms*, 32(4):359–368, 2017.

[190] Susan Kohl Malone, Babette Zemel, Charlene Compher, Margaret Souders, Jesse Chittams, Aleda Leis Thompson, Allan Pack, and Terri H Lipman. Social jet lag, chronotype and body mass index in 14–17-year-old adolescents. *Chronobiology international*, 33(9):1255–1266, 2016.

[191] Michael J Parsons, Terrie E Moffitt, Alice M Gregory, Sidra Goldman-Mellor, Patrick M Nolan, Richie Poulton, and Avshalom Caspi. Social jetlag, obesity and metabolic disorder: investigation in a cohort study. *International Journal of Obesity*, 39(5):842–848, 2015.

[192] Femke Rutters, Sofie G Lemmens, Tanja C Adam, Marijke A Bremmer, Petra J Elders, Giel Nijpels, and Jacqueline M Dekker. Is social jetlag associated with an adverse endocrine, behavioral, and cardiovascular risk profile? *Journal of biological rhythms*, 29(5):377–383, 2014.

[193] Rosa Levandovski, Giovana Dantas, Luciana Carvalho Fernandes, Wolnei Caumo, Iraci Torres, Till Roenneberg, Maria Paz Loayza Hidalgo, and Karla Viviani Allebrandt. Depression scores associate with chronotype and social jetlag in a rural population. *Chronobiology international*, 28(9):771–778, 2011.

[194] Mikhail F Borisenkov, Natalia B Petrova, Vladimir D Timonin, Lyudmila I Fradkova, Sergey N Kolomeichuk, Anna L Kosova, and Olga N Kasyanova. Sleep characteristics, chronotype and winter depression in 10–20-year-olds in northern european russia. *Journal of Sleep Research*, 24(3):288–295, 2015.

[195] Michael A Grandner. Epidemiology of insufficient sleep and poor sleep quality. In *Sleep and health*, pages 11–20. Elsevier, 2019.

[196] Consensus Conference Panel, Nathaniel F Watson, M Safwan Badr, Gregory Belenky, Donald L Bliwise, Orfeu M Buxton, Daniel Buysse, David F Dinges, James Gangwisch, Michael A Grandner, 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. *Journal of Clinical Sleep Medicine*, 11(6):591–592, 2015.

[197] Max Hirshkowitz, Kaitlyn Whiton, Steven M Albert, Cathy Alessi, Oliviero Bruni, Lydia DonCarlos, Nancy Hazen, John Herman, Eliot S Katz, Leila Kheirandish-Gozal, et al. National sleep foundation's sleep time duration recommendations: methodology and results summary. *Sleep health*, 1(1):40–43, 2015.

[198] Sutapa Mukherjee, Sanjay R Patel, Stefanos N Kales, Najib T Ayas, Kingman P Strohl, David Gozal, and Atul Malhotra. An official american thoracic society statement: the importance of healthy sleep. recommendations and future priorities. *American journal of respiratory and critical care medicine*, 191(12):1450–1458, 2015.

[199] Marie-Pierre St-Onge, Michael A Grandner, Devin Brown, Molly B Conroy, Girardin Jean-Louis, Michael Coons, and Deepak L Bhatt. Sleep duration and quality: impact on lifestyle behaviors and cardiometabolic health: a scientific statement from the american heart association. *Circulation*, 134(18):e367–e386, 2016.

[200] Edward J Stepanski. The effect of sleep fragmentation on daytime function. *Sleep*, 25(3):268–276, 2002.

[201] Sascha E Martin, Heather M Engleman, Ian J Deary, and Neil J Douglas. The effect of sleep fragmentation on daytime function. *American journal of respiratory and critical care medicine*, 153(4):1328–1332, 1996.

[202] Deepak Shrivastava, Syung Jung, Mohsen Saadat, Roopa Sirohi, and Keri Crewson. How to interpret the results of a sleep study. *Journal of community hospital internal medicine perspectives*, 4(5):24983, 2014.

[203] Peter Hauri and Joan Fisher. Persistent psychophysiologic (learned) insomnia. *Sleep*, 9(1):38–53, 1986.

[204] ML Perlis, DE Giles, WB Mendelson, Richard R Bootzin, and JK Wyatt. Psychophysiological insomnia: the behavioural model and a neurocognitive perspective. *Journal of sleep research*, 6(3):179–188, 1997.

[205] Richard R Bootzin. Stimulus control treatment for insomnia. *Proceedings of the American Psychological Association*, 7:395–396, 1972.

[206] Allison G Harvey. A cognitive model of insomnia. *Behaviour research and therapy*, 40(8):869–893, 2002.

[207] Kelly Glazer Baron, Kathryn J Reid, Roneil G Malkani, Joseph Kang, and Phyllis C Zee. Sleep variability among older adults with insomnia: associations with sleep quality and cardiometabolic disease risk. *Behavioral sleep medicine*, 15(2):144–157, 2017.

[208] Michele L Okun, Charles F Reynolds III, Daniel J Buysse, Timothy H Monk, Sati Mazumdar, Amy Begley, and Martica Hall. Sleep variability, health-related practices and inflammatory markers in a community dwelling sample of older adults. *Psychosomatic medicine*, 73(2):142, 2011.

[209] Andrew JK Phillips, William M Clerx, Conor S O'Brien, Akane Sano, Laura K Barger, Rosalind W Picard, Steven W Lockley, Elizabeth B Klerman, and Charles A Czeisler. Irregular sleep/wake patterns are associated with poorer academic performance and delayed circadian and sleep/wake timing. *Scientific Reports*, 7(1):1–13, 2017.

[210] Eva H Telzer, Diane Goldenberg, Andrew J Fuligni, Matthew D Lieberman, and Adriana Gálvan. Sleep variability in adolescence is associated with altered brain development. *Developmental cognitive neuroscience*, 14:16–22, 2015.

[211] Jeffrey M. Wooldridge. *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, 2010.

[212] Cheng Hsiao. Panel data analysis—advantages and challenges. *Test*, 16(1):1–22, 2007.

[213] Andrew Bell and Kelvyn Jones. Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data. *Political Science Research and Methods*, 3(1):133–153, 2015.

[214] John R Taylor. Error analysis. *Univ. Science Books, Sausalito, California*, 1997.

[215] Per Brockhoff. 02429 analysis of correlated data: Mixed linear models.

[216] Walter W. Stroup. *Generalized Linear Mixed Models*. CRC Press. Taylor Francis Group, 2013.

[217] David Harville. Extension of the gauss-markov theorem to include the estimation of random effects. *The Annals of Statistics*, pages 384–395, 1976.

[218] Nan M Laird and James H Ware. Random-effects models for longitudinal data. *Biometrics*, pages 963–974, 1982.

[219] Colin R Blyth. On simpson's paradox and the sure-thing principle. *Journal of the American Statistical Association*, 67(338):364–366, 1972.

[220] Douglas M Bates. lme4: Mixed-effects modeling with r, 2010.

[221] Walter W. Stroup. *Generalized Linear Mixed Models: Modern Concepts, Methods and Applications*. Texts in Statistical Science. CRC Press, 2013 edition, 2013.

[222] Nathaniel Beck and Jonathan N Katz. Throwing out the baby with the bath water: A comment on green, yoon and kim. 2000.

[223] Nathaniel Beck and Jonathan N Katz. Random coefficient models for time-series—cross-section data: Monte carlo experiments. *Political Analysis*, pages 182–195, 2007.

[224] Sven E Wilson and Daniel M Butler. A lot more to do: The sensitivity of time-series cross-section analyses to simple alternative specifications. *Political Analysis*, pages 101–123, 2007.

[225] Tom A. B. Snijders. *Multilevel Analysis: An Introduction To Basic And Advanced Multilevel Modeling*. 2 edition, 2011.

[226] Ida P Kristensen and Gregory Wawro. Lagging the dog? the robustness of panel corrected standard errors in the presence of serial correlation and observation specific effects. In *annual meeting of the Society for Political Methodology, University of Minnesota*, 2003.

[227] Stefanie Schurer and Jongsay Yong. Personality, well-being and the marginal utility of income: What can we learn from random coefficient models? 2012.

[228] Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*, 2014.

[229] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2018.

[230] Alexandra Kuznetsova, Per B Brockhoff, and Rune Haubo Bojesen Christensen. lmertest package: tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 2017.

[231] Steven G Luke. Evaluating significance in linear mixed-effects models in r. *Behavior research methods*, 49(4):1494–1502, 2017.

[232] Avi Sadeh. The role and validity of actigraphy in sleep medicine: an update. *Sleep medicine reviews*, 15(4):259–267, 2011.

[233] Bhanu Prakash Kolla, Subir Mansukhani, and Meghna P Mansukhani. Consumer sleep tracking devices: a review of mechanisms, validity and utility. *Expert review of medical devices*, 13(5):497–506, 2016.

[234] Hannah Scott, Leon Lack, and Nicole Lovato. A systematic review of the accuracy of sleep wearable devices for estimating sleep onset. *Sleep medicine reviews*, 49:101227, 2020.

[235] Janna Mantua, Nickolas Gravel, and Rebecca Spencer. Reliability of sleep measures from four personal health monitoring devices compared to research-based actigraphy and polysomnography. *Sensors*, 16(5):646, 2016.

[236] Jennifer L Martin and Alex D Hakim. Wrist actigraphy. *Chest*, 139(6):1514–1527, 2011.

[237] Christopher M Depner, Philip C Cheng, Jaime K Devine, Seema Khosla, Massimiliano de Zambotti, Rébecca Robillard, Andrew Vakulin, and Sean PA Drummond. Wearable technologies for developing sleep and circadian biomarkers: a summary of workshop discussions. *sleep*, 43(2):zsz254, 2020.

[238] World population prospects - population division.

[239] Omar B Ahmad, Cynthia Boschi-Pinto, Alan D Lopez, Christopher JL Murray, Rafael Lozano, Mie Inoue, et al. Age standardization of rates: a new who standard. *Geneva: World Health Organization*, 9(10), 2001.

[240] Gho | by category | mean body mass index trends among adults, age-standardized (kg/m^2) - estimates by who region.

[241] Soultana Macridis, Nora Johnston, Steven Johnson, and Jeff K Vallance. Consumer physical activity tracking device ownership and use among a population-based sample of adults. *PloS one*, 13(1):e0189298, 2018.

[242] John D Omura, Susan A Carlson, Prabasaj Paul, Kathleen B Watson, and Janet E Fulton. National physical activity surveillance: Users of wearable activity monitors as a potential data source. *Preventive medicine reports*, 5:124–126, 2017.

[243] Tim Althoff, Eric Horvitz, Ryen W White, and Jamie Zeitzer. Harnessing the web for population-scale physiological sensing: A case study of sleep and performance. In *Proceedings of the 26th international conference on World Wide Web*, pages 113–122, 2017.

[244] Serge Daan, DG Beersma, and Alexander A Borbély. Timing of human sleep: recovery process gated by a circadian pacemaker. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 246(2):R161–R183, 1984.

[245] David P Doane and Lori E Seward. Measuring skewness: a forgotten statistic? *Journal of statistics education*, 19(2), 2011.

[246] Hylton E Molzof, Sarah E Emert, Joshua Tutek, Mazheruddin M Mulla, Kenneth L Lichstein, Daniel J Taylor, and Brant W Riedel. Intraindividual sleep variability and its association with insomnia identity and poor sleep. *Sleep medicine*, 52:58–66, 2018.

[247] Erika W Hagen, Jodi H Barnet, Lauren Hale, and Paul E Peppard. Changes in sleep duration and sleep timing associated with retirement transitions. *Sleep*, 39(3):665–673, 2016.

[248] Organisation for Economic Co-Operation and Development. Ageing and employment policies—statistics on average effective age of retirement, 2017.

[249] Gladys Martinez, Kimberly Daniels, and Anjani Chandra. *Fertility of men and women aged 15-44 years in the United States: National Survey of Family Growth, 2006-2010*. Number 51. Department of Health and Human Services, Centers for Disease Control and ..., 2012.

[250] Barbara C Galland, Barry J Taylor, Dawn E Elder, and Peter Herbison. Normal sleep patterns in infants and children: a systematic review of observational studies. *Sleep medicine reviews*, 16(3):213–222, 2012.

[251] The World Bank. Air transport, passengers carried, 2018.

[252] Thomas Reilly, Jim Waterhouse, and Ben Edwards. Jet lag and air travel: implications for performance. *Clinics in sports medicine*, 24(2):367–380, 2005.

[253] Thomas Reilly, Jim Waterhouse, and Benjamin Edwards. Some chronobiological and physiological problems associated with long-distance journeys. *Travel Medicine and Infectious Disease*, 7(2):88–101, 2009.

[254] Ulf Aslak and Laura Alessandretti. Infostop: Scalable stop-location detection in multi-user mobility data. *arXiv preprint arXiv:2003.14370*, 2020.

[255] Hartmut Rosa. Social acceleration: ethical and political consequences of a desynchronized high–speed society. *Constellations*, 10(1):3–33, 2003.

[256] Philipp Lorenz-Spreen, Bjarke Mørch Mønsted, Philipp Hövel, and Sune Lehmann. Accelerating dynamics of collective attention. *Nature communications*, 10(1):1–9, 2019.

[257] Jean-Baptiste Michel, Yuan Kui Shen, Aviva Presser Aiden, Adrian Veres, Matthew K Gray, Joseph P Pickett, Dale Hoiberg, Dan Clancy, Peter Norvig, Jon Orwant, et al. Quantitative analysis of culture using millions of digitized books. *science*, 331(6014):176–182, 2011.

[258] James E Cutting, Kaitlin L Brunick, Jordan E DeLong, Catalina Iricinschi, and Ayse Candan. Quicker, faster, darker: Changes in hollywood film over 75 years. *i-Perception*, 2(6):569–576, 2011.

[259] Brett Hutchins. The acceleration of media sport culture: Twitter, telepresence and online messaging. *Information, communication & society*, 14(2):237–257, 2011.

[260] Hartmut Rosa. *Alienation and acceleration: Towards a critical theory of late-modern temporality*, volume 3. Aarhus Universitetsforlag, 2010.

[261] Sarah A Burgard and Jennifer A Ailshire. Gender and time for sleep among us adults. *American sociological review*, 78(1):51–69, 2013.

[262] Mathieu Nédélec, Shona Halson, Abd-Elbasset Abaidia, Said Ahmaidi, and Gregory Dupont. Stress, sleep and recovery in elite soccer: a critical review of the literature. *Sports Medicine*, 45(10):1387–1400, 2015.

[263] Peter Fowler, Rob Duffield, and Joanna Vaile. Effects of domestic air travel on technical and tactical performance and recovery in soccer. *International Journal of Sports Physiology and Performance*, 9(3):378–386, 2014.

[264] Sarah Forbes-Robertson, Edward Dudley, Pankaj Vadgama, Christian Cook, Scott Drawer, and Liam Kilduff. Circadian disruption and remedial interventions. *Sports Medicine*, 42(3):185–208, 2012.

[265] Thomas Reilly, Greg Atkinson, Ben Edwards, Jim Waterhouse, Torbjörn Åkerstedt, Damien Davenne, Björn Lemmer, and Anna Wirz-Justice. Coping with jet-lag: a position statement for the european college of sport science. *European Journal of Sport Science*, 7(1):1–7, 2007.

[266] Peter M Fowler, Wade Knez, Stephen Crowcroft, Amy E Mendham, Joanna Miller, Charli Sargent, Shona Halson, and Rob Duffield. Greater effect of east versus west travel on jet lag, sleep, and team sport performance. *Medicine & Science in Sports & Exercise*, 49(12):2548–2561, 2017.

[267] Maria-Raquel G Silva, António Pascoal, Hugo-Henrique Silva, and Teresa Paiva. Assessing sleep, travelling habits and jet lag in kite surfers according to competition level. *Biological Rhythm Research*, 47(5):677–689, 2016.

[268] Samira Bourgeois-Bougrine, Philippe Carbon, Charlotte Gounelle, Regis Mollard, and Alex Coblentz. Perceived fatigue for short-and long-haul flights: a survey of 739 airline pilots. *Aviation, space, and environmental medicine*, 74(10):1072–1077, 2003.

[269] Mikael Sallinen, Maria Sihvola, Sampsa Puttonen, Kimmo Ketola, Antti Tuori, Mikko Härmä, Göran Kecklund, and Torbjörn Åkerstedt. Sleep, alertness and alertness management among commercial airline pilots on short-haul and long-haul flights. *Accident Analysis & Prevention*, 98:320–329, 2017.

SLEEPJ, 2021, 1–16

doi: 10.1093/sleep/zsaa169
Advance Access Publication Date: 4 September 2020
Original Article

Original Article

Gender differences in nighttime sleep patterns and variability across the adult lifespan: a global-scale wearables study

Sigga Svala Jonasdottir[1], Kelton Minor[2,] and Sune Lehmann[1,2,*]

[1]Department of Applied Mathematics and Computer Science, Technical University of Denmark, Lyngby, Denmark and
[2]Copenhagen Center for Social Data Science, University of Copenhagen, Copenhagen, Denmark

*Corresponding author. Sune Lehmann, Technical University of Denmark, Anker Engelunds Vej 1, Building 101A, 2800 Kongens Lyngby, Denmark.
Email: sljo@dtu.dk.

Abstract

Study Objectives: Previous research on sleep patterns across the lifespan have largely been limited to self-report measures and constrained to certain geographic regions. Using a global sleep dataset of in situ observations from wearable activity trackers, we examine how sleep duration, timing, misalignment, and variability develop with age and vary by gender and BMI for nonshift workers.

Methods: We analyze 11.14 million nights from 69,650 adult nonshift workers aged 19–67 from 47 countries. We use mixed effects models to examine age-related trends in naturalistic sleep patterns and assess gender and BMI differences in these trends while controlling for user and country-level variation.

Results: Our results confirm that sleep duration decreases, the prevalence of nighttime awakenings increases, while sleep onset and offset advance to become earlier with age. Although men tend to sleep less than women across the lifespan, nighttime awakenings are more prevalent for women, with the greatest disparity found from early to middle adulthood, a life stage associated with child-rearing. Sleep onset and duration variability are nearly fixed across the lifespan with higher values on weekends than weekdays. Sleep offset variability declines relatively rapidly through early adulthood until age 35–39, then plateaus on weekdays, but continues to decrease on weekends. The weekend–weekday contrast in sleep patterns changes as people age with small to negligible differences between genders.

Conclusions: A massive dataset generated by pervasive consumer wearable devices confirms age-related changes in sleep and affirms that there are both persistent and life-stage dependent differences in sleep patterns between genders.

Statement of Significance

A global dataset from wearable devices enables a detailed understanding of age-related tendencies in sleep patterns, controlling for country-level and within-individual variation. During early adulthood, we find elevated levels of variability in sleep offset and duration along with high levels of weekend–weekday misalignment, suggesting that mismatches between internal timing and external demands are pervasive during this phase of human development. In older adulthood, reduced sleep duration and increased sleep disturbances may either contribute to, or correlate with, further age-related decline. Gender gaps in average sleep duration, timing and nighttime awakenings are apparent, despite considerable heterogeneity in circadian preferences. Information about parenting mobile application usage can be paired with big data from wearable devices to explore lifestage gender inequality in sleep quality. Further research on person-centered behavioral interventions that promote regular sleep–wake cycles are needed.

Key words: sleep; big data; aging; gender; sleep variability; sleep misalignment; sleep timing and duration

Submitted: 6 January, 2020; **Revised:** 4 August, 2020

Introduction

Sufficient sleep is fundamental to healthy human functioning. Brief, irregular and/or disturbed sleep are risk factors for infectious disease, cardiovascular disease, depression, and all-cause mortality [1–3]. Similar to other physiological functions, sleep patterns vary between people across the human population and change within individuals over the lifespan [4–6]. Meta-analyses and cross-sectional research provide convergent evidence that the ability to initiate and maintain sleep declines as people age, independent of factors such as medical co-morbidities and medication use [7, 8]. However, large-scale in situ data on changes in sleep patterns across the lifespan remain scarce and geographically constrained.

Age-associated changes in sleep include both decreases in total time asleep, deep (slow wave) sleep, and rapid eye movement sleep, as well as increases in sleep latency, time awake after sleep onset, stage 1 and stage 2 sleep [7, 9, 10]. Increases in nighttime awakenings with age suggest a decline in the buildup of homeostatic sleep pressure [11]. Evidence that healthy older adults exhibit less objective and subjective sleepiness after selective slow wave sleep deprivation indicates that sleep need may decline as adults age [12], while other research has shown that reduced sleep still negatively impacts cognitive performance, irrespective of age [13, 14]. A recent review concluded that while there is still no consensus, the current body of evidence largely supports the hypothesis that older adults have an impaired ability to generate sleep rather than a reduced sleep need [15]. Hence, additional insight is needed to characterize structural changes in sleep patterns over the lifespan and to better understand the underlying drivers.

Aside from changes in the duration and composition of sleep, adult aging is further characterized by changes in circadian regulatory processes, with phase advances and diminished amplitudes in daily core body temperature, melatonin, and cortisol rhythms [11]. These changes are associated with, and often proxied by, aging-related advances in sleep timing after adolescence, with individuals going to bed and waking up earlier with increasing age [16–20].

Previous research has established the importance of measuring sleep timing separately on both free and work days, since weekly social schedules constrain daily rhythms and can induce a misalignment with biological time [21].

In a series of large scale survey-based studies conducted with participants from four countries (Germany, Switzerland, the Netherlands, and Austria), both the difference between free and work day sleep duration and midsleep timing (social jet lag) were shown to decline with age [22, 23]. Importantly, the authors did not report how underlying differences in sleep onset and offset timing between work and free days may contribute to observed developments in misalignment. Since people exert more practical influence over the beginning and end of their sleep period compared with midsleep, an expectation of how onset and offset change on weekdays and weekends with age would be useful.

Moreover, while a recent study found that inter-individual variability in midsleep timing declines with age [24], far less is known about how intra-individual variability in sleep timing and duration within work and free day periods changes across the life course [25–27]. In situ sleep data collected over an extended period is needed to inform expectations about age-related developments in sleep timing, misalignment, and variability.

Prior research has established that the extent of adult age-related changes in sleep patterns is highly moderated by gender, with women reaching both puberty and their peak eveningness earlier than men in young adulthood and sleeping longer than men until age 50–60, a period of adulthood that coincides with menopause [23, 28–30]. Well-controlled laboratory studies have found that women exhibit phase advanced core body temperature and melatonin rhythms, as well as a shorter intrinsic circadian period compared with men [31, 32]. While some larger sleep surveys have found that females are more likely to be morning types compared with males from the end of adolescence to late adulthood [17, 33], others have found no apparent gender differences [34, 35] and one nationally representative study found the opposite trend [36]. Notably, these cross-sectional studies were conducted within different countries using slightly different methodologies, highlighting the need for an integrated global assessment of potential gender-related differences in sleep. Beyond long-term changes in sleep patterns, events occurring during certain stages of adult development can impact women and men differently. Pregnancy and the postpartum period are associated with dynamic physiological changes and behavioral demands known to disturb sleep quality for women, although limited in situ evidence exists comparing sleep disturbances for both women and men during young child-rearing [37, 38].

Despite revealing salient age-related changes in sleep patterns, there are a number of areas where previous research can be extended and improved. Early lifespan meta-analyses favored data from predominantly Western countries, averaged across different sleep assessment methodologies and depended extensively on short-term polysomnography recordings that may have disrupted participants' habitual sleep cycles [7, 9]. By comparison, cross-sectional sleep survey research has primarily relied on subjective measures prone to self-report, recall, and rounding biases [17, 24, 39, 40]. Moreover, large-scale sleep surveys typically ask for single estimates of work and free day sleep onset and offset times, and thus do not enable measurement of the intra-individual variability of sleep patterns during work and free day periods [26]. To address these limitations, recent research has drawn upon behavioral data from mobile phones and self-tracking apps to infer the dynamics of human activity and sleep in everyday contexts across large populations of device users [41–47]. Differing from mobile phones, the activity trackers employed in the current study were worn by users, enabling closely coupled measurements of human sleep patterns. Similar to wrist-actigraphs, wearable devices can automatically monitor sleep measurements in situ over extended periods of use, making it possible to study both average and time-varying sleep patterns in daily contexts [20, 48, 49]. Drawing on a dataset of objective sleep measurements and mobile application use statistics from a large sample of n = 69,650 wearable users over multiple years across 47 countries, we investigate the following questions:

1. After controlling for country-level and individual-level variation with mixed effects models, does global data from consumer wristband devices confirm gender differences in age-related changes in sleep duration, timing, and circadian misalignment?

2. How does intra-individual variability in sleep duration and timing develop across the life course separately on weekdays and weekends, adjusting for country-level variation?
3. Does child-rearing—as proxied by parenting mobile application use—predict life-stage gender differences in nighttime sleep disturbances?

The present study differs from prior self-report studies, most of which featured self-report data from single countries or regions, and—at the time of writing—represents the most geographically extensive analysis of age-associated changes in sleep using consistent, objectively recorded measures of sleep duration, timing, misalignment, and variability.

Methods

Data and demographics

The anonymized data set used in this study consists of sleep observations collected using smart wristbands from 2015 to 2018 (see the section Data collection below for full details). In total, we analyze 11.14 million nights of sleep observations arising from 69,650 adult nonshift workers, about a third of them women. In Table 1, we show the number of individuals and nights broken out by the demographic variables *age group*, *gender*, and *BMI category* and in Supplementary Table S1 lists all the countries in our sample and ratio of users residing there. We compare age statistics (median age) in our sample to information provided by the United Nation Population Division (UN) [50] for the five countries with the most users in the dataset in Supplementary Table S2. The median values in our sample and the overall population correspond well: users from Japan are slightly younger (by 1 year) while those from Taiwan and the United Kingdom match their respective reference populations. By comparison, users from Germany are younger (by 7 years) as well as those from Russia (by 5 years). We also compare age standardized BMI statistics of the study sample to population estimates provided by the World Health Organization (WHO) in Supplementary Table S3 [51, 52]. We find both men and women from all countries fall within or near the 95% confidence intervals (CIs) of the WHO reference values. Women from the UK fall 0.5 points above the 95% CI and women from Japan average 0.5 points below the 95% CI reference range.

Sleep duration, timing, and variability outcomes

We use nine sleep metrics to assess how sleep patterns change across the lifespan. *Sleep duration* specifies the total recorded time a person spent asleep during a given night. To quantify sleep timing, we use *sleep onset* (the registered point in time when a person fell asleep) and *sleep offset* (the recorded time when a person woke up). We measure the misalignment between an individual's internal biological clock and external social clock by applying a variant of the formula used to compute *social jetlag* [22]. Specifically, instead of calculating social jetlag for midsleep (see Supplementary Figure S8 for a comparison) we estimate weekend–weekday differences in sleep onset and offset. The weekend–weekday misalignment of sleep timing can lead to the loss of sleep duration on weekdays and partial compensation on weekends, which is quantified by estimating the weekend–weekday sleep duration difference. These metrics are calculated for each week of data collection, resulting in weekly repeated measurements for each user, which are then aggregated to produce user-level averages. We also study the *variability* of sleep onset, offset, and duration in order to estimate the regularity of people's sleep timing and duration. We quantify intra-individual variability as the standard deviation of a person's corresponding measurements for each sleep outcome, and compute this separately for weekends and weekdays [53].

The individual-level covariates for this study are gender (female/male) and BMI categories (underweight/normal weight/overweight/obese) which were labelled according to the World Health Organization classification [54, 55].

Additionally, we also include temporal variables for day category (weekday/weekend) to account for likely differences in the social structure over the course of the week. Since we do not

Table 1. Overview of the data set with a focus on demographics: age, gender, and BMI

	# of adult users in sample			# of night sleep observations		
	All	Male	Female	All	Male	Female
Total	69,650	47,656	21,993	11,144,539	7,673,495	3,471,044
Age groups						
19–24	5,466	3,745	1,721	579,315	383,761	195,554
25–29	8,976	5,813	3.163	1,105,037	698,471	406,566
30–34	11,224	7,414	3,810	1,559,445	1,022,033	537,412
35–39	9,796	6,584	3,212	1,520,749	1,024,874	495,875
40–44	9,315	6,435	2,880	1,591,014	1,092,829	498,185
45–49	9,934	6,994	2,940	1,844,717	1,308,604	536,113
50–54	7,164	5,059	2,105	1,395,210	999,842	395,368
55–59	4,445	3,180	1,265	879,907	646,560	233,347
60–67	3,330	2,433	897	669,145	496,521	172,624
BMI categories						
Underweight	2,272	1,197	1,075	350,954	173,220	177,734
Normal weight	34,063	22,101	11,962	5,773,876	3,843,837	1,930,039
Overweight	22,936	17,371	5,565	3,558,454	2,687,277	871,177
Obese	10,379	6,988	3,391	1,461,255	969,161	492,094

The table provides statistics for both the number of adults in the sample, as well as the number of nights analyzed. Note that the data set contains more men than women and more people within the normal weight range BMI category.

directly observe schedules, we assume the likelihood of workdays is highest on weekdays and work-free days is highest on weekends, similar to others [20, 24].

Data modeling

We analyzed the data set using R Version 3.5.1 [56]. Given the longitudinal and hierarchical structure of the data with repeated measurements within users, and users nested within their country of residence, observations are likely highly correlated on both levels (country and user). To account for this dependence within the data set, we adopt a mixed effects modeling framework [57]. Mixed effects models allow us to control for user and country-level variation while examining age-related trends in sleep patterns and assessing the influence of demographic factors. Concretely, the model can be specified in matrix form as

$$y = X\beta + Zu + \quad \varepsilon, \quad with \quad u \sim N_q(0, G) \quad and \quad \varepsilon \sim N_n(0, R),$$

with β representing the fixed effects parameters, u representing the random effects, X representing the $n * p$ design matrix for the fixed-effects parameters, and Z the $n * q$ design matrix describing the random effects. The models for weekday–weekend differences and sleep variability are defined without user random effects since these measures were computed for each user as single values. We use the lmerTest R package, the lmer function to fit the data set and apply Satterthwaite's degrees of freedom method to estimate the p-values for the significance of fixed factors [58, 59].

We center the age variable around its mean to help improve interpretability, prevent multicollinearity, and lower the scale of the variables to accommodate the inclusion of age squared in the model.

Data collection

Data were collected from 2015 to 2018 via Sony SmartBand (SmartBand Talk [SWR30] and SmartBand 2 [SWR12]), designed to track physical activity and sleep behavior. The waterproof wristbands use proprietary, internally validated algorithms based on movement registered by an internal accelerometer to estimate sleeping and waking states in 1-min epochs. When connecting the wristband at first, users received visual instruction on how and where (wrist) to place the device and were advised to wear it on their dominant side. All wearable data included in this study were wirelessly transmitted via Bluetooth to an accompanying mobile phone application, which also independently registered user mobile application usage statistics. Similar to many other wearable devices and wrist actigraphs, the devices used in the present study detect sleep timing and total sleep time but do not detect time in bed, preventing the further study of age-related changes in sleep latency and sleep efficiency. Moreover, although the armbands have been validated internally within SONY, we note that the wristbands have not been publicly validated using the gold standard of polysomnography as recommended inn the Sleep Research Society Workshop on wearable devices for the measurement of sleep [60]. The wristbands employed in this study have been shown to produce wake and sleep states that converge with objective measures of user mobile phone use patterns [46]. However, this global dataset offers unique methodological advantages; scale, longitudinal coverage, and ecologically valid observations. By using it, we follow a growing trend of utilizing commercial devices in sleep research to study sleep behavior in naturalistic settings at large scales [39, 61, 62]. Further, we have performed an extensive comparison of the findings here with multiple independent global sleep datasets. We find that this worldwide dataset externally converges with country-level sleep measures from separate large-scale datasets, demonstrates consistency over the period of observation and replicates age-related sleep trends from previously published self-report studies, including changes in sleep duration and timing. These full comparisons are presented in the Supplementary Information Sections *Comparison of country-level statistics to other publications* and *Consistency over time* and Results. Further, the wristbands employed in this study have been shown to produce wake and sleep states that converge with objective measures of user mobile phone use patterns [46].

Study participants consist of anonymized users who consented to share their data for research purposes. Age group, BMI category, gender, and country of residence were preprocessed from self-reported demographic information. All data analyses were carried out in accordance with the EU's General Data Protection Regulation 2016/679 (GDPR) and the regulations set out by the Danish Data Protection Agency. The GDPR describes regulations for data protection and privacy in the European Union and the European Economic Area. It also addresses the transfer of personal data outside the EU and EEA areas.

Data processing and inclusion criteria

To reduce the risk of including sleep observations from those suffering from insomnia, artificially shortened sleep observations due to users ceasing wristband use in the middle of the resting period, observations from nightshift workers or any other possible problems, outliers from the sleep data were removed by applying inclusion filters to sleep duration, onset, and offset. We adopt standard filters for sleep duration ($3 < duration < 13$), matching those applied by Roenneberg et al. [23]. These filters are more inclusive (by 2 h) than those used by Walch et al. [45] and Althoff [62] ($4 < duration < 12$). Furthermore, we apply the following conservative sleep timing filters. First, we remove all sleep observations with onset or offset times greater than one and a half standard deviations away from the sample average computed separately for weekdays and weekends and obtain the following time filters:

- 20:24≤ onset weekends ≤04:52
- 20:28 ≤ onset weekdays ≤03:59
- 03:59 ≤ offset weekends ≤12:52
- 03:21 ≤ offset weekdays ≤11:25

This results in the removal of 12% of sleep observations yielding a final dataset consisting of 11.14 million nights from 69,650 users. The full data preprocessing procedure is described in the section *Data Filtering* in the Supplementary Information.

To help ensure that sleep estimates are representative of typical sleeping behavior, we further require all participants to have a minimum threshold of sleep observations. Specifically, each user must have sleep observations extending over a minimum period of 4 weeks, with at least 1 weekday and weekend night per week, amounting to a minimum 8 nights per user (median 87 nights per user). We also limit our analysis to adults 19–67 years of age due to limited across-country data for older age groups. Each user is assigned a country of residence, defined as the country in which the majority of their sleep entries occur.

Sociocultural variation

This article focuses on how sleep duration, timing, misalignment, and variability develop with age and how other demographic factors such as gender and BMI may affect these trends. Hence, it is important to note that users in the sample reside across a wide range of countries around the world. Breaking the data set out by country of residence yields cohorts from 47 distinct countries with at least 150 users in each country. Supplementary Table S1 lists out all of the countries and percentage of users residing in each country, as well as the ratio of male users within each country. Table 2 shows the development of sleep onset and duration with age for men split up by the top 5 countries with the most users. It is evident that there is substantial heterogeneity in the amount and timing of sleep obtained between countries. The summary statistics reveal, consistent with the literature, that there are indeed large disparities in sleep patterns across cultures [20, 45, 63, 64]. Since the focus of the present research is to assess and identify age- and gender-related changes, we control for these baseline country-level differences through our mixed effects modeling framework described in Data modeling.

Results

Sleep timing and duration over the lifespan

In order to summarize the development of sleep onset, offset, and duration across the lifespan, we calculate each user's average value and then aggregate across our study sample by age, gender, and day type (weekday or weekend).

The resulting curves for sleep onset, offset, and duration are shown in Figures 1–3, D.

Development of onset

The main panel on Figure 1 shows that, overall, sleep onset becomes earlier as people age and that people tend to go to bed later on weekends (indicated by lighter colors); the difference between weekday and weekend is roughly constant for both men and women across all age-groups. There are large differences in mean onset time between men and women (more than 30 min for the 19–24 young adult age group), which progressively become smaller in magnitude across the lifespan, eventually falling out of the range of statistical significance for the 60–67 older adult age group. This eventual confluence of sleep onset is driven by a steeper age-related advance in sleep onset time for men than women. While the decline in sleep onset time is consistent for men, the rate of decrease in onset time for women nearly plateaus after the age 35–39 range. Even though the 95% CIs for the mean are narrow, the actual distribution of sleep onset is quite broad, as shown in Figure 1, A–C, which shows the distribution of onset time for the 19–24 group, the 40–44 age group, and the 60–67 group. In order to directly visualize the progression of sleep onset timing between genders, in Figure 1, E (weekends) and Figure 1, F (weekdays), we display the difference of male/female onset from the average curve (genders weighted equally). The gender gap in onset time appears to persist until around age 40, when the two curves begin to converge.

Figure 1, D plots the aggregated raw data from our sample; the displayed trends in sleep onset are confirmed by our modeling which adjusts for demographic covariates, and controls for individual and country baseline behavior. Sleep onset has a quadratic relationship with age ($p < 2 * 10^{-16}$, Supplementary Table S17). The model estimates a 29 ± 0.20 min difference between weekends and weekdays for women and 28 ± 0.20min difference for men (age group 40–44), with a

Table 2. Development of sleep onset and duration by age split up by the top five countries with the most users in the data set

	Japan, n = 17,231 (24.7%)	Germany, n = 7,140 (10.3%)	Russia, n = 5,095 (7.3%)	Taiwan, n = 5,028 (7.2%)	UK, n = 3,900 (5.6%)
Age groups	Average sleep onset (hh:mm)				
19–24	00:53	23:55	00:39	01:12	00:24
25–29	00:44	23:52	00:29	00:59	00:07
30–34	00:41	23:40	00:21	00:51	23:52
35–39	00:27	23:36	00:08	00:42	23:46
40–44	00:21	23:37	00:04	00:32	23:41
45–49	00:15	23:30	00:03	00:30	23:45
50–54	00:06	23:31	23:58	00:16	23:36
55–59	23:54	23:27	23:55	23:52	23:45
60–67	23:42	23:26	23:50	23:50	23:42
Age groups	Average sleep duration (h)				
19–24	6.6	7.3	7.0	6.7	7.3
25–29	6.4	7.1	7.0	6.7	7.3
30–34	6.4	7.1	7.0	6.6	7.2
35–39	6.3	7.0	7.0	6.4	7.1
40–44	6.3	6.9	7.0	6.5	7.1
45–49	6.2	7.0	7.0	6.4	7.0
50–54	6.2	7.0	7.0	6.6	7.0
55–59	6.3	6.9	7.0	6.6	7.0
60–67	6.4	7.2	7.1	6.5	7.1

Note there are strong differences between countries with a clear split between European and Asian countries.

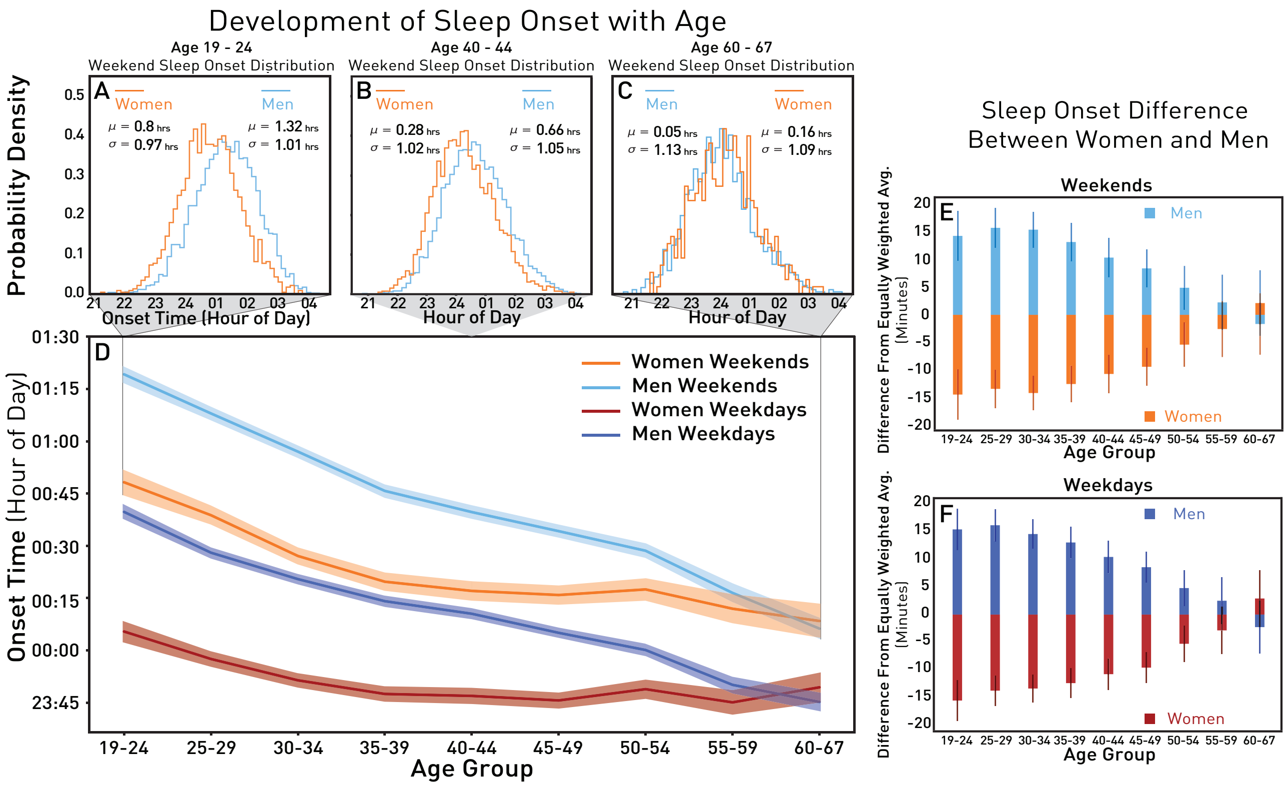


Figure 1. Distributions for sleep onset on weekends split up by gender for different age groups: (A) age 19–24, (B) age 40–44, and (C) age 60–67. The development of average sleep onset by age group split up by gender and day type (weekend/weekday). The red/orange colors correspond to women, light/dark blue colors correspond to men, darker colors represent weekdays and lighter colors signify weekends. The colored envelopes display 95% CIs around each age group mean (D). The equally weighted, between gender sleep onset difference by age group with 95% CI on weekends (E) and weekdays (F).

later average onset on weekends. When we consider the rate of decrease of sleep onset for men, we find the model results suggest an even steeper rate of decrease than the raw data (see Supplementary Figure S3 and Supplementary Table S17). Consequently, the difference between men and women at age 40–44 on weekdays is estimated to be 24 ± 1.5 min based on the raw data but 15 ± 1.5 min (95% CI) by the model. Furthermore, the model estimates the onset curves for men and women to intersect slightly earlier (within the 50–54 age range) than the raw data. From age 55 to 67, the model indicates that men are expected to exhibit earlier onset than women. The mixed effects model indicates that there is a larger range of country-level random effect values for onset (1.76 h) than offset (1.35 h). This finding is in accordance with the results from a study conducted in 2014 using surveys and smartphone data: country of residence appears to exert a stronger influence on adult sleep onset than offset [45].

Development of offset

Turning to the development of sleep offset, Figure 2, D shows that the mean value of sleep offset mostly decreases with age, and people tend to wake up earlier as they get older. On weekdays, the curve is nearly flat for women between ages 45 and 59, but there is an increase for the age interval 60–67. Men consistently decrease in wake-up time with age except the slight increase from 60 to 67 on weekdays. The contrast between weekends and weekdays is nearly fixed across the lifespan with an hour difference resulting in later wake-up time on weekends. The curves on Figure 2, D show roughly the opposite behavior of what we observed for sleep onset (Figure 1, D), with the 95% CI of the mean values for men and women overlapping until the middle of adulthood and thereafter diverging with men rising earlier than women. Thus, from age 19 to 39 women and men exhibit an average tendency to go to bed at different times yet wake up at similar times. The sleep offset curves for men and women diverge earlier on weekends (40–44) where the separation occurs one age group later (45–49) on weekdays. This can be seen even more clearly on Figure 2, E (weekends) and Figure 2, F (weekdays) which shows the difference of sleep offset by gender and age group from the equally weighted average of the curves for men and women.

Similar to the case of onset, the plotted mean offset values have small error bars (as indicated by the 95% confidence bands), while the actual distributions of sleep offset are quite broad. This is depicted in Figure 2, A–C, which shows the distribution of offset time for the 19–24 group, the 40–44 group, and the 60–67 group, respectively. We observe close agreement between the plots in Figure 2, D and the model results (Supplementary Table S19). Age, gender, and type of day are the most influential factors on wake-up time, which has a quadratic relationship with age ($p < 10^{-16}$, see Supplementary Table S19). For people aged 40–44, the model shows men to have the same sleep offset time as women, whereas on weekends they are expected to wake up 2.7 ± 0.20 min *earlier* (Supplementary Table S19). This is displayed in Figure 2, D, which shows that the sleep offset curves for men and women diverge earlier in the lifespan on weekends than weekdays.

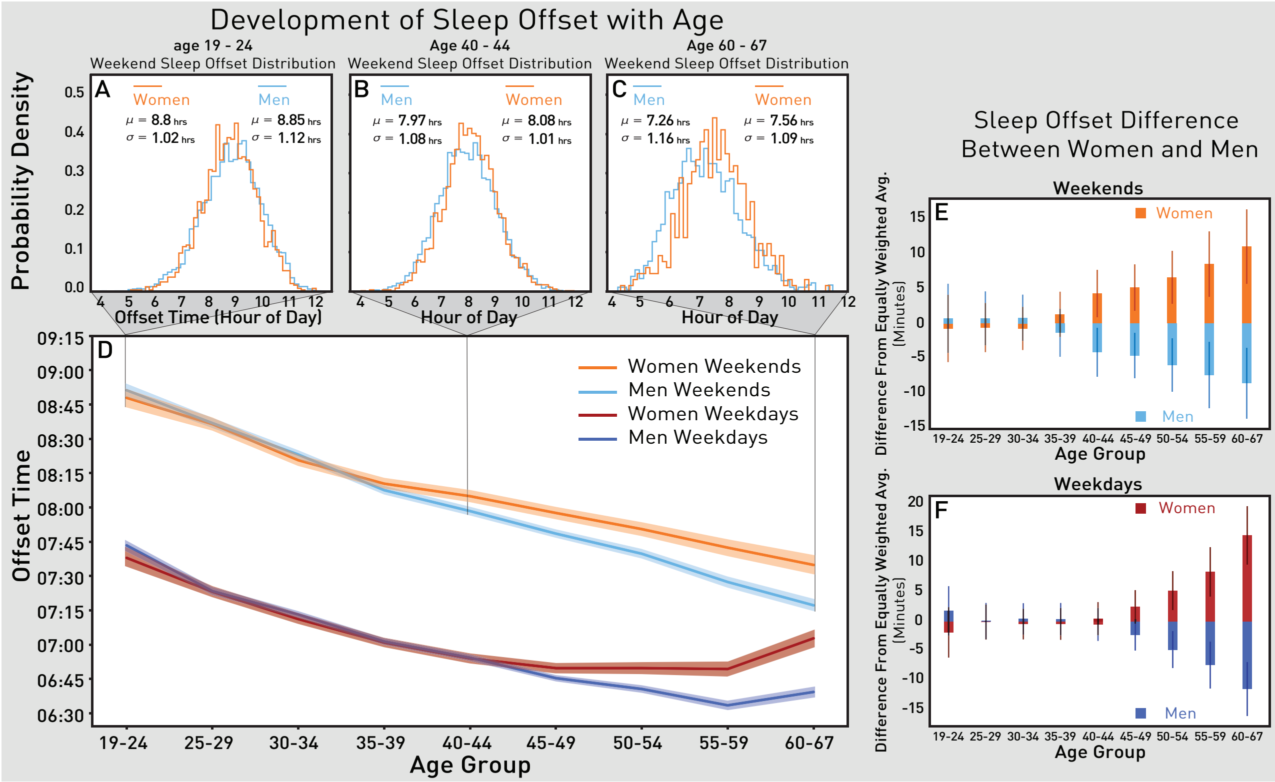


Figure 2. Distributions for sleep offset on weekends split up by gender for different age groups: (A) age 19–24, (B) age 40–44, and (C) age 60–67. The development of average sleep offset by age group split up by gender and day type (weekend/weekday). The red/orange colors correspond to women, light/dark blue colors correspond to men, darker colors represent weekdays and lighter colors signify weekends. The colored envelopes display 95% CIs around each age group mean (D). The equally weighted, between gender sleep offset difference by age group with 95% CI on weekends (E) and weekdays (F).

Development of duration

Figure 3, D shows that sleep duration tends to decrease across the lifespan. This development is nearly linear for weekends and less so on weekdays, with a small increase in duration for the oldest age group on weekdays. The difference between weekends and weekdays remains similar throughout the lifespan with a slightly smaller gap for men and women in the oldest group. This is highlighted in Figure 3, E (weekends) and Figure 3, F (weekdays) which show the difference of sleep duration by gender and age group from the equally weighted average of the curves for men and women. Although the average behavior shown in Figure 3, D exhibits statistically significant differences between men and women across different age groups, each aggregated group mean is derived from a broad range of underlying behavior as Figure 3, A–C shows the distributions of sleep duration for the 19–24 group, the 40–44 group, and the 60–67 group, respectively.

The mixed effects model for sleep duration, which controls for individual and country of residence variations, generally confirms the trends observed in the aggregated raw data plots visible in Figure 3 (for comparison of the raw data and model fit see Supplementary Figure S4). The weekend–weekday differences in duration are apparent in the model results but the magnitude of gender differences turn out smaller, due to different rates of change in sleep duration with age. Consequently, the curves for men and women come close to overlapping from age 55 to 67, see Supplementary Figure S4. Adjusting for BMI, the aggregated raw data estimates women at age 40–44 to sleep 23 $\pm$1.7 min longer than men, whereas the model estimates a difference of 11 $\pm$1.0 min (95% CI), see Supplementary Table S21 for estimates of fixed effects.

Development of nighttime awakenings with age

Having considered the progression of sleep onset, offset, and duration, we now assess how the prevalence of nighttime awakenings develops across adulthood. To quantify nighttime awakenings, we use wake after sleep onset (WASO) which refers to periods of wakefulness occurring after defined sleep onset and reflects sleep fragmentation [65]. For each registered night, WASO is the total time an individual is recorded awake (after defined sleep onset, but also occurring before defined sleep offset). Since sleep was recorded in 1-min epochs, only WASO measurements greater than 60 s were registered by the wristbands. We observe a large fraction of users with zero instances of WASO (85% of the users have a median WASO value of zero). This is in part because accelerometer-based fitness bands may underestimate sleep disruptions if individuals are awake but lying still in bed [66]. For that reason, our measure of nighttime awakenings may be conservative and correspond to relatively large sleep disruptions detectable by the embedded accelerometer. The percentage of users with nonzero median WASO is plotted by age group and gender in Figure 4. The percentage of individuals with nonzero median WASO increases with age; for the 19–24 age group, 4.5% of men and 9.7% of women have nonzero median WASO compared with 33.4% (men) and 35.6% (women) for the 60–67 age group.

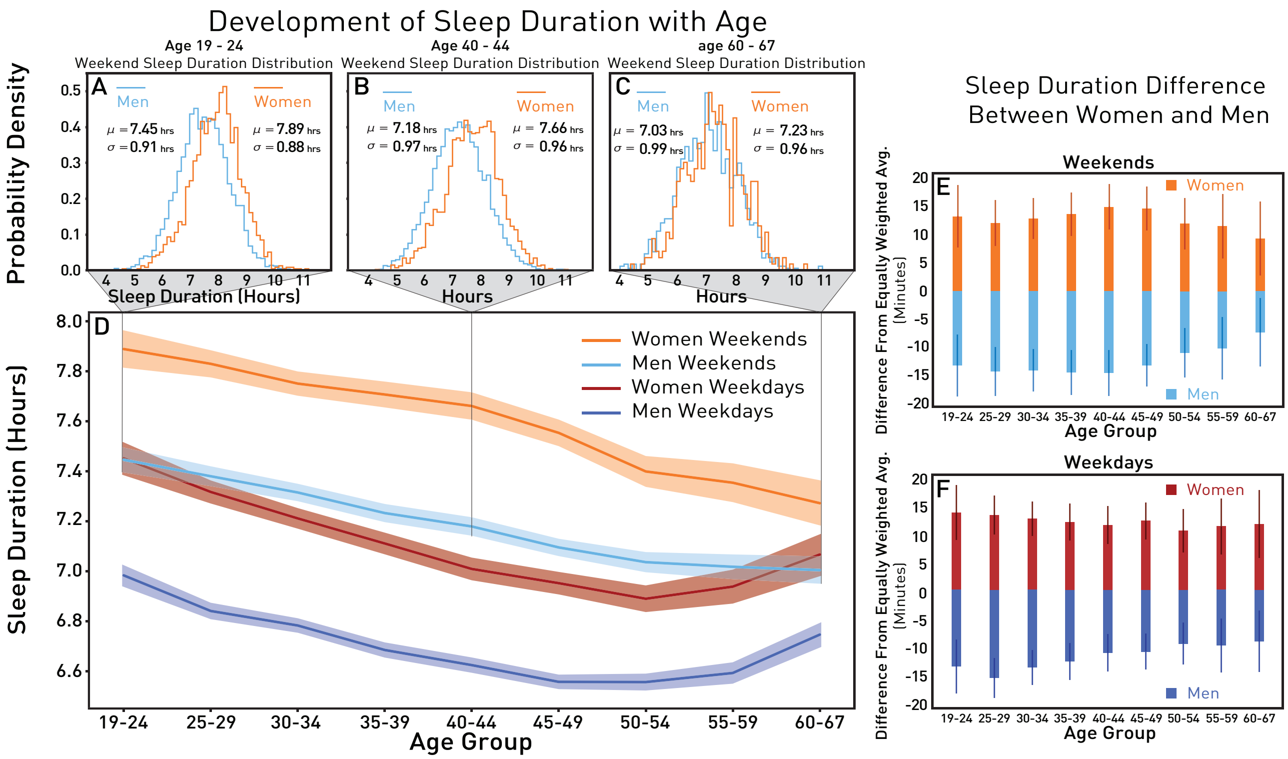


Figure 3. Distributions for sleep duration on weekends split up by gender for different age groups: (A) age 19–24, (B) age 40–44, and (C) age 60–67. The development of average sleep duration by age group with 95% CI split up by gender and day type (weekend/weekday). The red/orange colors correspond to women, light/dark blue colors correspond to men, darker colors represent weekdays and lighter colors signify weekends. The colored envelopes display 95% CIs around each age group mean (D). The equally weighted, between gender sleep duration difference by age group with 95% CI on weekends (E) and weekdays (F).

Figure 4 shows that a higher proportion of women have nonzero WASO medians than men from age 19 to 39. Specifically, 17.5% of women have nonzero WASO median values compared with 13.5% for men aged 19–39. The distribution of within individual median WASO differs significantly for men and women aged 19–39, estimated with two sample Kolmogorov–Smirnov (KS) statistics where $p = 3.06 * 10^{-21}$. Interestingly, this same stage of life marks a biological window where childbirth, infant rearing and child caretaking are more likely [67]. As a post hoc analysis, we investigate the hypothesis that increased prevalence of nighttime awakenings during early adulthood may be linked to tending to infants and young children which exhibit irregular sleep patterns for the first 0–2 years of life [68]. For that reason, we analyze the age group 19–39, where the difference between the two curves in Figure 4 diverges between genders. As a proxy for information regarding parental status and infant-rearing, we reference aggregated app-context information. Specifically, we can anonymously identify users as probable parents if they have apps installed on their phones intended for parents with young children ("parent apps"). We describe how we identify apps as "parent apps" in the Supplementary Information: Identifying "parent apps." We find that women in the age range 19–39 with a parent app installed on their devices have a significantly different distribution of median (denoted *M*) WASO than age-matched women without the application on their phone (estimated with two sample KS statistics, $p = 9.66 * 10^{-21}$), where M_{WASO} for women with parent app $= 184$ s and M_{WASO} for women without parent app $= 65$ s (see distribution Supplementary Figure S5). By comparison, the distribution of median WASO for young adult men with parent

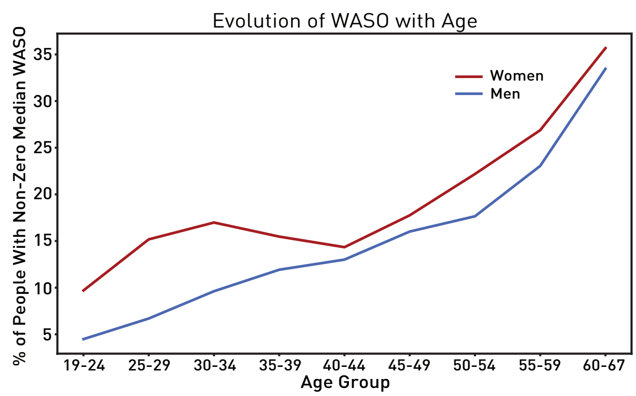


Figure 4. The percentage of people with nonzero median WASO, by age group and gender. The red color corresponds to women and blue color to men.

apps does not differ significantly from those without them (estimated with two sample KS statistics, $p = 0.228$) where M_{WASO} for men with parent app $= 52$ s and M_{WASO} for men without parent app $= 37$ s (see distribution Supplementary Figure S5).

Next, we examine the subset of sleep observations for users with M_{WASO} 0, but this choice of subset eliminates the skew arising from a large fraction of users with zero WASO measurements (Supplementary Figure S5).

After applying the same comparison, we find that women aged 19–39 with parent apps installed on their phones have a significantly different distribution of mean values (denoted μ) for their WASO than similarly aged women without parent apps (estimated with two-sample KS statistics, $p = 1.38 * 10^{-14}$) where μ_{WASO} for women with parent $app = 1105$ s and μ_{WASO} for women with parent $app = 874$ s (see distribution on Supplementary Figure S6). In contrast, when we carry out the same comparison for men, we find that their distributions do not differ significantly between the group with and without parent apps (estimated with two-sample KS statistics, $p = 0.207$) and μ_{WASO} for men with parent $app = 905$ s and μ_{WASO} for men without parent $app = 867$ s (see distribution Supplementary Figure S6).

Development of circadian misalignment with age

Many people (about 75% of the US and European labor force) maintain a conventional 5 day work week from 9 to 5 which constrains their weekly sleep behavior [69, 70]. This recurrent temporal pattern can lead to substantial sleep deprivation during weekdays and sleep compensation during weekends, in addition to a weekend–weekday contrast in sleep timing [23]. Figure 5, D illustrates the development of weekend–weekday sleep timing differences over the lifespan (green/pink colors for onset and blue/red colors for offset). From approximately age 19

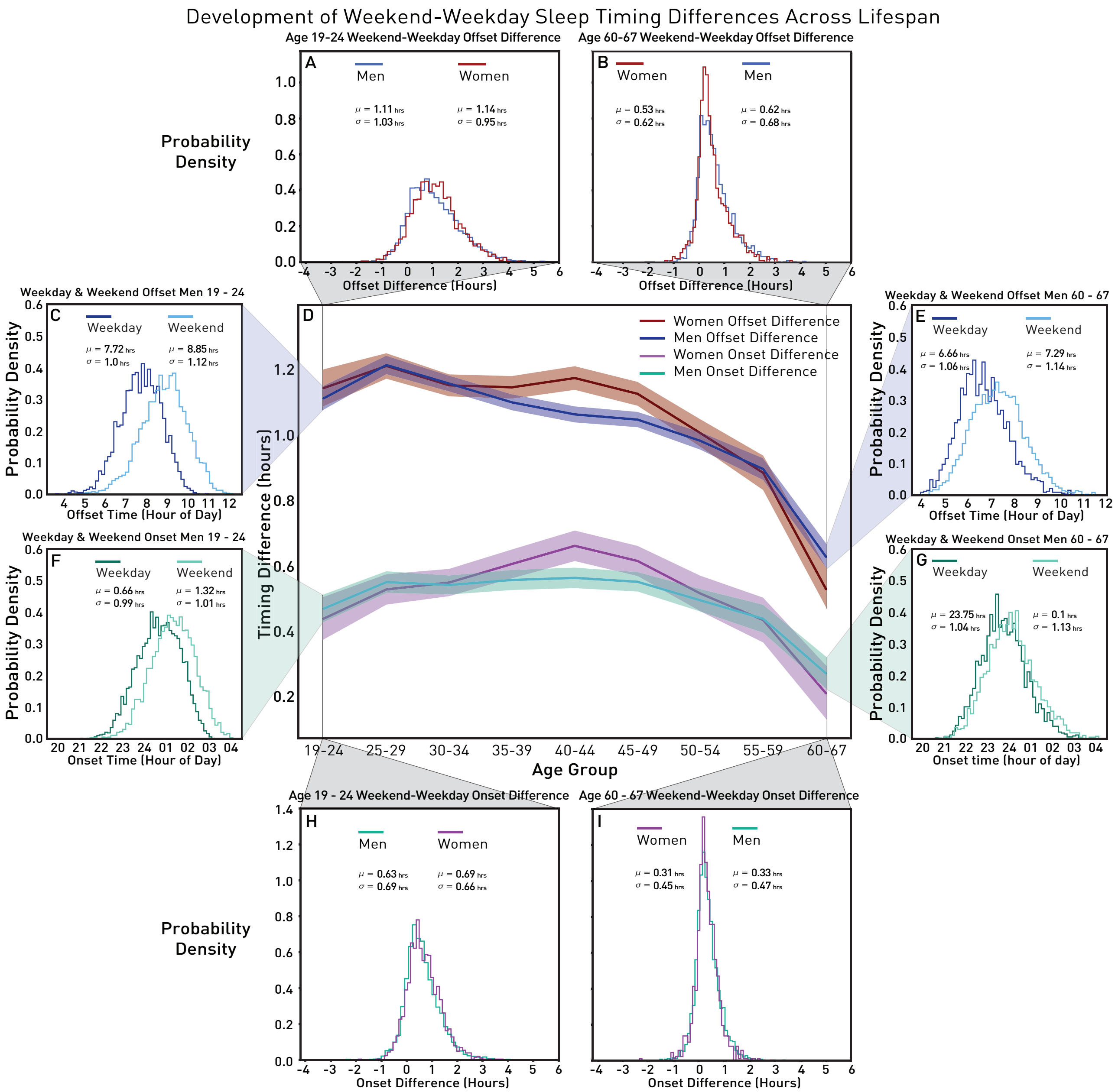


Figure 5. The development of weekend–weekday differences for sleep onset and offset by age group split by gender. The red/pink colors correspond to women, blue/green blue colors correspond to men, darker colors represent weekend–weekday offset difference and lighter colors signify weekend–weekday onset difference. The colored envelopes display 95% CIs around each age group mean (D). The distribution of weekend–weekday differences in sleep onset for different age groups: (A) age 19–24 and (B) age 60–67. The distribution for the sleep offset weekend–weekday differences for different age groups: (H) age 19–24 and (I) age 60–67. Distribution for sleep onset time on weekdays and weekends for different age groups: (C) age 19–24 and (E) age 60–67. Distribution for sleep offset time on weekdays and weekends for different age groups: (F) age 19–24 and (G) age 60–67.

to 55, average sleep offset tends to be 55–70 min later on weekends while onset tends to be 25–35 min later on weekends over the same period. Thus, adults in our sample tend to sleep half an hour more on weekends than weekdays. This result is confirmed in Figure 6, B which shows the development of weekend-weekday sleep duration difference with age. We identify sleep duration to be 25–40 min longer on weekends from age 19 to 55. After age 55, weekend–weekday misalignment in onset timing declines to 20 ± 2.6 min for older adult men, alongside a marginally larger decrease in offset misalignment to 38 ± 2.3 min (95% CI). The distributions for sleep offset and onset split by weekends and weekdays for age group 19–24 (offset Figure 5, C and onset Figure 5, F) and age group 60–67 (offset Figure 5, E and onset Figure 5, G) illustrate this contrast. For example, offset is on average 67 ± 2.2 min later on weekends for men age group 19–24 but 38 ± 2.3 min for men in the 60–67 year old group (95% CI). Interestingly, average sleep offset misalignment remains greater than sleep onset misalignment into older adulthood, despite an overall convergence toward more similar weekend and weekday schedules and a reduction in weekend–weekday sleep duration difference.

Figures 5, D and 6, B indicate that misalignment in both sleep timing and duration progress similarly for men and women across the majority of the lifespan. A possible exception is visible during the 35–49 age range, during which both sleep offset and onset misalignment are slightly greater for women. Increased offset misalignment in this period for women appears to be driven by later weekend offset times compared to men, while weekday offset times are similar for both genders (see section Sleep timing and duration over the lifespan). This general similarity between genders is confirmed when observing the overlapping distribution for weekday–weekend differences of sleep onset and offset times for men and women, respectively, age group 19–24 (offset on Figure 5, A and onset on Figure 5, H) and age group 60–67 (offset Figure 5, B and onset Figure 5, I).

As before, we consider the potential biases in the data set when drawing conclusions from the figures and compare the aggregated empirical data to our mixed effects model. Our primary inferences from Figures 5 and 6 are verified by our modeling results presented in Supplementary Tables S23, S25, and S27. After controlling for country and adjusting for BMI in the mixed effects model, the slight difference between middle-aged men and women (age group 40–44, see Figure 5, D), is no longer evident or negligible due to small effect size (the model estimates men to have a 1.9 ± 1.0 min higher weekend–weekday sleep offset difference and 3.0 ± 0.8 min higher weekend–weekday sleep onset difference than women [95% CI]).

Sleep variability over the lifespan

Figure 7, A and B shows the development of adult onset and offset variability with age (green/purple colors correspond to onset and blue/red to offset, while the darker shades represent weekdays and lighter shades indicate weekends). Interestingly, we find that onset variability, measured as the intra-individual standard deviation of onset time, is nearly fixed across the lifespan at 1.1 h on weekdays and 1.3 h on weekends. By comparison, offset variability decreases relatively rapidly for age group 19–24, both for men (weekdays 1.2 ± 0.015 h and weekends 1.5 ± 0.016 h) and women (weekdays 1.3 ± 0.021 h and weekends 1.4 ± 0.021 h) up until age 35–39, remaining around 0.9 h on weekdays while continuing to decrease on weekends at a gradual rate. Variability for all measurements (onset, offset, and duration) is always higher on weekends than weekdays. We find that young adults have more variable sleep offset times than onset times both on weekends and weekdays. Figure 7, C and D shows that the difference between offset variability and onset variability is positive and higher across early adulthood (19–29) for men and women on both weekends and weekdays. The weekend difference between offset and onset variability is larger for men across the age 19–34 range, while the weekday difference is larger for women in the 19–24 and 25–29 age groups.

Figure 7, F illustrates the development of sleep duration variability over the lifespan, which decreases gradually with age such that the youngest group of men have only 14 ± 1.2 min higher sleep variability than the oldest group on weekends. From Figure 7, A, B, and F, we observe small significant differences between men and women; higher onset variability both on weekends and weekdays after early adulthood and consistently higher sleep offset variability on weekends for all age groups. When comparing these results to our mixed effects models which control for the influence of country and demographic covariates, we find that all of the general conclusions inferred from the descriptive plots in Figure 7 are verified (Supplementary Tables S29, S31, and S33). Taking the age 40–44 group as an example, the model estimates a 3-min higher onset variability for men than women on weekdays, 2-min greater

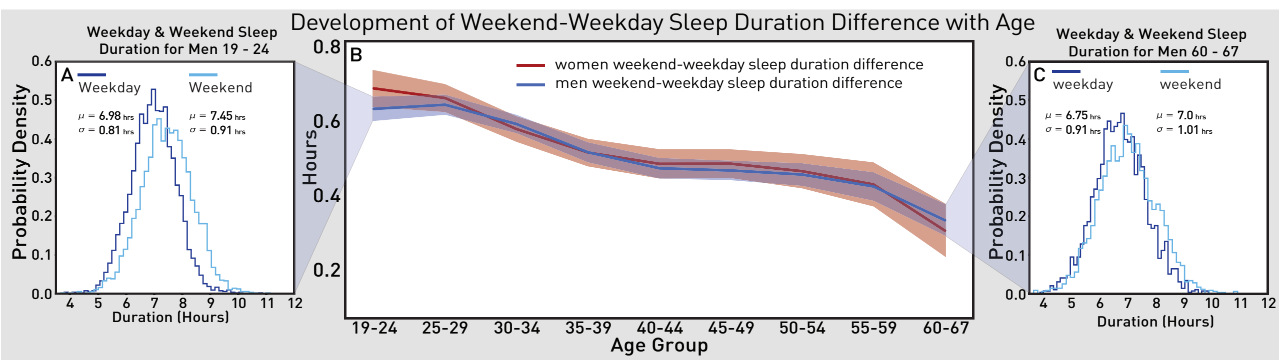


Figure 6. The development of weekend–weekday sleep duration difference by age group split up by gender. The red color corresponds to women and blue color to men. The colored envelopes display 95% CIs around each age group mean (B). Distribution of weekend–weekday duration differences for different age groups: (A) age 19–24 and (C) age 60–67.

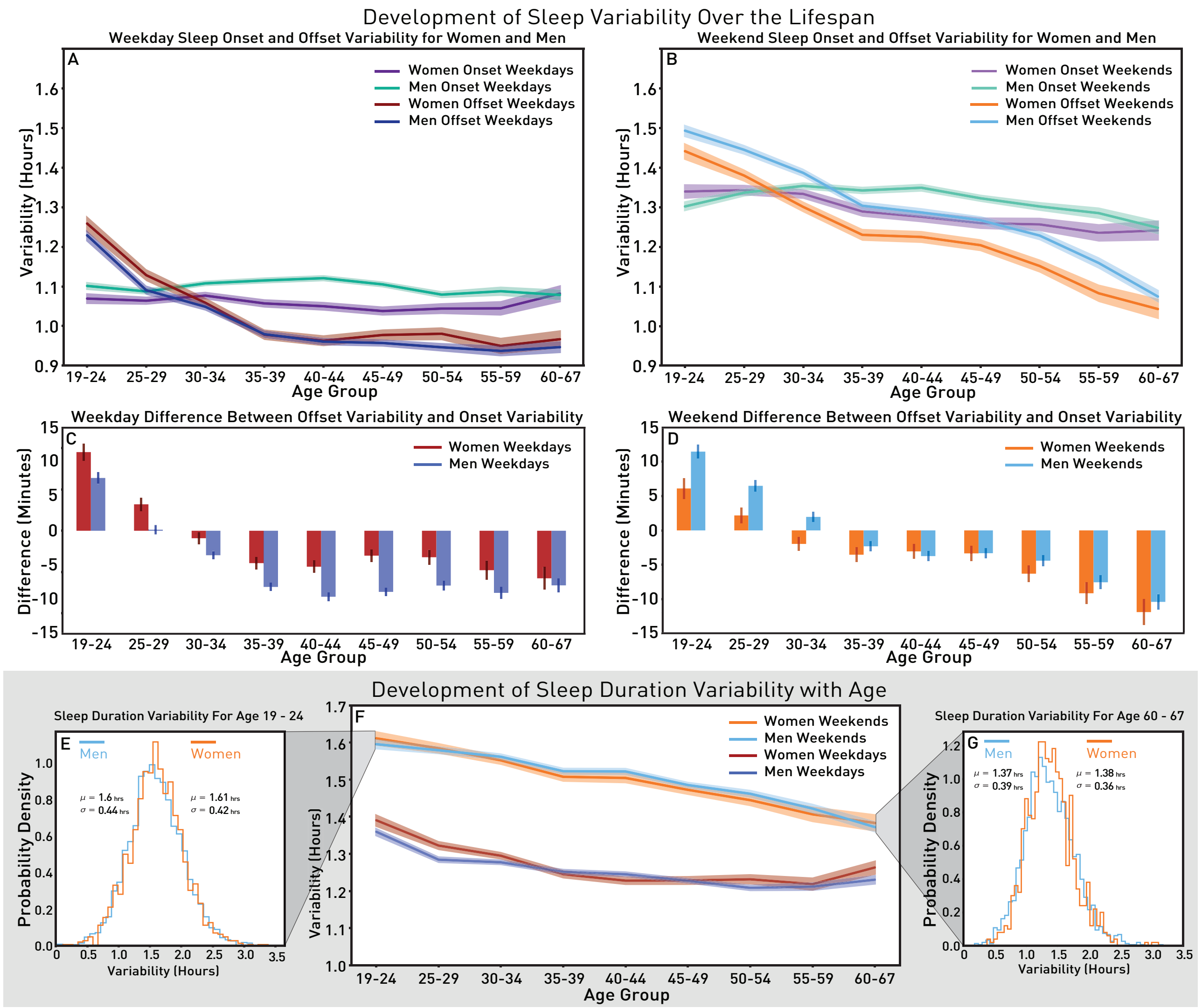


Figure 7. The development of sleep onset and offset variability with 95% CI by age group and split by gender and type of day: (A) weekdays and (B) weekends. Red/orange curves is offset variability for women, light/dark blue is offset variability for men, light/dark purple is onset variability for women, and light/dark green is onset variability for men. Darker colors represent weekdays and lighter colors weekends. The colored envelopes display 95% CIs around each age group mean. The difference between sleep onset and offset variability with 95% CI by age group on weekdays (C) and weekends (D). The development of sleep duration variability by age group split up by gender and day type (weekday/weekend). The red/orange colors correspond to women, light/dark blue colors correspond to men, darker colors represent weekdays and lighter colors signify weekends. The colored envelopes display 95% CIs around each age group mean (F). Distribution of sleep duration variability by gender for different age groups: (E) age 19–24 and (G) age 60–67.

onset variability on weekends and 5-min higher offset variability on weekends, which are trends that can also be identified on Figure 7, A and B.

The effect of BMI

In Table 3, we list average sleep onset, offset, and duration for men and women within the four BMI groups estimated with the mixed effects model (95% CI) for age group 40–44 on weekdays. Overall, differences in sleep timing between BMI groups are statistically insignificant and/or small in effect. As one exception to this trend, we find that men within the normal BMI range sleep on average 17 ± 2.8 min more than those in the obese BMI category, and men in the underweight category sleep on average 24 ± 4.7 min more than those in the obese category. We carry out further discussion concerning these results in section called "BMI Discussion" in the Supplementary Information.

Discussion

Drawing on a massive global sleep dataset comprised of 11.14 million sleep observations from 69,650 adults spanning 47 countries, we confirm the presence of age-related changes in sleep duration, timing, misalignment, and variability. After controlling for baseline country-level variation using mixed effects models, we find that younger adulthood is marked by both delayed sleep onset and offset, and higher intra-individual sleep duration variability, offset variability, weekend–weekday misalignment, and weekend–weekday sleep duration difference compared with older adulthood. Conversely, sleep duration is shorter and nighttime awakenings are more prevalent during older adulthood. Only sleep onset variability exhibits little to no difference across the youngest and oldest age groups in our sample. Certain changes in sleep behavior progress consistently across most age groups observed, while others appear to be highly life-stage and/or gender dependent. In contrast to studies based on

Table 3. Mixed effects model estimates of average sleep onset, offset, and duration for different BMI and gender groups on weekdays age 40–44 (95% CI)

	Onset hh:mm ± m		Offset hh:mm ± m		Duration hours ± hours	
	Women	Men	Women	Men	Women	Men
Underweight	23:55 ± 8	00:10 ± 8	07:05 ± 7	07:11 ± 8	7.03 ± 0.0938	6.95 ± 0.114
Normal weight	23:45 ± 7	00:01 ±7	06:58 ± 6	06:58 ± 6	7.03 ± 0.0938	6.84 ± 0.0952
Overweight	23:45 ± 7	00:01 ± 7	06:55 ± 6	06:55 ± 6	6.98 ± 0.0972	6.72 ± 0.102
Obese	23:52 ± 8	00:10 ± 8	06:54 ± 6	06:54 ± 6	6.86 ± 0.0988	6.56 ± 0.106

single or short-term observations, our unique dataset of millions of multi-night sleep recordings enables a consistent, detailed understanding of age-related tendencies in sleep patterns including intra-individual variability within inferred working and leisure periods as well as misalignment between them. We confirm several recognized age-related changes in human sleep and provide novel evidence of gender differences across key phases of adult development. Further, in the Supplementary Information, we provide (1) a comparison of our data with multiple independent large scale and global sleep datasets, (2) we explicitly compare our global estimates of social jetlag to those from Roenneberg et al. [23], and perform a quantitative exploration of underlying regional differences that appear to drive this disparity of our results compared to Roenneberg et al., and lastly, (3) we perform an exploratory analysis of the possible effect of retirement age, which varies by country and demonstrates that different regional policies appear to affect people's sleep patterns.

Misalignment

Circadian misalignment has been found to be associated with negative health outcomes such as obesity, metabolic risk factors for diabetes, and atherosclerotic cardiovascular disease, as well as adverse behaviors such as drinking and smoking which can negatively impact healthy human development [23, 71]. Notably, we find considerably lower levels of social jetlag in our sample across all observed age groups compared with the values reported by Roenneberg et al. [23] (Supplementary Figure S9). In the Roenneberg et al. [23] study, the sample consisted of questionnaire respondents from predominantly four European countries (Germany, Switzerland, the Netherlands, and Austria). Thus, the discrepancy may reflect a mismatch between global and regional circadian preferences, recall biases linked to the questionnaire and/or other unobserved differences. Constraining our sample to only include the same primary European countries as Roenneberg et al. yields markedly higher values of social jetlag across the lifespan than in our full global sample as well as altered age-related gender differences during early and late adulthood, with men incurring marginally more social jetlag than women—consistent with the age-related dependencies identified by Roenneberg et al. By contrast, in our global sample middle-aged women have marginally more social jetlag than men, with negligible gender differences in other age groups. Comparing social jetlag levels between regional strata of our sample from Asia and Europe suggests that social jetlag for young adults may be over twice as large in the same European region sampled by Roenneberg et al. [23], and ~1.5 times larger for middle-aged and older adults (Supplementary Figure S10). This provides suggestive evidence that the gap between the attenuated magnitude of social jetlag in our full sample relative to Roenneberg et al.'s may not merely be due to different means of data collection and associated measurement error (objective multi-night recording vs. self-report questionnaires). Rather, underlying regional differences appear to play an important role. We contend that accounting for underlying country-level variation is important to prevent biased global estimates of salient sleep outcomes and age-related developments.

In line with previous research, we find weekend–weekday differences in sleep timing and duration to be more pronounced among younger adults, with these elevated differences (Supplementary Figure S8) driven primarily by earlier sleep offset on weekdays and later offset on weekends than weekdays. Weekend–weekday misalignment in sleep timing and duration slightly decrease with age and decline more rapidly around the age range of 55–59, leading to near convergence of sleep onset timing on weekends and weekdays for older adults aged 60–67. Reduced misalignment in older adulthood may signal the social onset of exiting the labor force for retirement. However, some misalignment in both sleep offset, midsleep and duration persists across the age groups observed, indicating that pervasive work schedules likely continue to exert an influence on people's sleep–wake cycles through most of the adult lifespan. Our results indicate that sleep research involving adults should account for weekend–weekday heterogeneity in sleep patterns, even in older populations where nonstandard weekday schedules might otherwise be presumed.

Variability

A growing body of research indicates that irregular sleep is linked to maladaptive responses adverse to human health [25, 53, 72–80] Outside of research on weekend–weekday misalignment, limited evidence exists about age-related changes in sleep variability in sleep patterns within individuals, particularly during weekdays and within weekends [81]. Taken together, two recent cross-sectional studies found that between-individual onset, offset, duration, and chronotype variability decrease with age [24, 45]. Similarly, a sleep diary-based study found that intraindividual variability in sleep duration decreases with age [27]. By comparison, our data set shows that intraindividual variability in sleep onset is close to fixed over the lifespan—implying that young, middle-aged, and older adults may have persistently variable sleep onset times, whereas variability in wake-up times decreases with age, likely driving the observed decline in sleep duration variability.

We find that young adults tend to have more variable offset times than onset times, a trend which inverts after age 35 due to a decrease in sleep offset variability. One possible explanation is that a concurrent rise in weekday alarm clock use to meet fixed workplace, childcare, and/or other social commitments might drive this reduction in offset variability. Across our population, variability measurements are consistently higher on weekends, confirming that sleep patterns are more regulated on weekdays, in line with the alarm clock hypothesis. However, the gradual decline in both weekend sleep offset variability and duration variability across most of the lifespan suggests that both endogenous and exogenous factors may be involved.

Nighttime awakenings

Previous research and reviews found that women are at a greater risk than men to develop insomnia, and both insomnia and other sleep disorders are more prevalent in women during pregnancy and the postpartum period [37, 82–84]. Ours is the first study to use the contextual information encoded in app usage as a proxy for parental status to explore lifestage gender inequality in sleep quality. Importantly, the gender difference we observe in nighttime awakenings is more pronounced from young to middle adulthood than from middle to late adulthood. Supporting the hypothesis that increased sleep disturbances for women during early adulthood might be driven by childbirth and raising young children, we find a significant difference in the median WASO between women with parent apps installed on their phone and age-matched women without such applications. When we applied the same comparison to the two corresponding groups of men, we found no significant difference, a finding which suggests that the gap in prevalence of nighttime awakenings between women and men aged 19–39 may be driven by the presence of infants or young children—as well as gender-associated caretaking norms—which disproportionately interrupt the sleep of female parents. This finding is in agreement with a panel study on changes in sleep satisfaction and sleep duration after childbirth, which also found a less pronounced decrease in sleep satisfaction for men than women [85]. Others have interpreted disturbed sleep after childbirth as a contributor and/or symptom of postpartum depression [86, 87].

In line with previous observational studies that suggest age-related increases in WASO, we find that the prevalence of people regularly experiencing sleep disturbances increases with age [7, 9]. A greater proportion of women than men regularly experience some time awake after sleep onset across all age groups observed in our study, indicating that more women may have difficulty maintaining sleep even though women on average sleep longer than men. Taken together, these findings contribute to the nascent literature on the unequal burden of child rearing on women's sleep quality [88]. The use of parents apps to identify individuals with young children illustrates the promise of using contextual information related to app usage as a novel way to understand the connection between sleep and overall behavior.

Sleep timing and duration

Epidemiological studies have demonstrated that men are, on average, later chronotypes than women until 40–50 years of age, after which their circadian phase advances to overlap or become earlier [17, 24]. Our study both confirms (Supplementary Figure S7) and expands on this finding by documenting the underlying dynamics between sleep onset and offset across these age groups that shape the full sleep period and its relative position. We find that men tend to have a later sleep onset than women up until 50–54 years of age, while up until the age range of 35–39 there is no significant difference in offset time between men and women. Thereafter, from middle to late adulthood, women tend to rise later. Taken together, this inversion may be indicative of gender-gaps in both domestic and labor demands during this period from mid-late adulthood [89]. It is possible that the general overlap in wake-up times for women and men during young to middle adulthood may be due to temporarily convergent external demands characteristic of this phase of development, such as attending university, work, tending to infants and/or raising young children, etc. By choosing to focus our analysis on both the beginning and end of the sleep period, rather than just its midpoint (Supplementary Figure S7) as commonly used in epidemiological sleep studies, we capture these differences and changes which have not been consistently described before at a global scale.

The finding that men sleep less than women on average across age groups [23, 63], confirmed by our study, is believed to have both a biological and social basis [89, 90]. For instance, we find that the sleep surplus for women relative to men is largest during young to middle adulthood when sleep interruptions are considerably more common for women than for men, likely due to the differential burden of caregiving. Thus, a combination of imbalanced nocturnal demands and socially imposed offset timing due to labor schedules may drive the observed gender differences in onset. Indeed, from middle to late adulthood average onset times converge and average offset times diverge. Furthermore, in line with previous research [6, 7, 9], we find that average sleep duration declines with age, with increasing portions of the average sleep distributions for both men and women falling below 7 h until weekday sleep duration slightly rebounds after age 60, a phase associated with attenuated working demands due to retirement [90]. Interestingly, later weekday wake up timing in late adulthood was apparent in Germany and the United Kingdom, but was not evident for Japan within the age range of our sample. Thus regional heterogeneity in transitioning out of the labor force may be reflected by differences in the manifestation of partial sleep timing recovery (see Supplementary Information: Analysis of effective retirement age with three-way interaction of age, gender, and country). However, such recovery in sleep offset appears to be consistently more subdued for older men than women. A recent global cross-sectional study found that acute cognitive deficits in reasoning and verbal ability can arise from sleeping less than 7–8 h regardless of age [13]. Importantly, average weekday sleep duration for men in our sample was consistently under 7 h across all age groups observed.

Limitations

Several considerations should be weighed when interpreting the results of this study. First, the wearable fitness bands used rely on in-built accelerometers and proprietary algorithms developed and internally validated by a global mobile technology company. Accelerometry-based consumer sleep trackers are known to slightly overestimate sleep duration and underestimate sleep disruptions [66], suggesting that actual sleep duration may be slightly lower than recorded in this study and that our estimates for the prevalence of people with frequent nighttime awakenings may be conservative. Second, age and BMI were self-reported. Despite possible recall bias, we find good general agreement between the World Health Organization's country-level estimates of median age and age standardized BMI and the corresponding estimates from our data set. Nevertheless, our sampled population of wearable users may not be representative of the wider population due to potential unobserved factors also associated with wearable device ownership, such as post-secondary education attainment [91]. Third, given the cross-sectional design of the current study, we cannot statistically identify whether the observed trends across age groups correspond to within-individual changes over the lifespan or rather reflect generational differences in normative sleep patterns. An exemplary longitudinal study which analyzed the change in diurnal timing preferences of 567 males in Finland across a 23 year period found that sleep timing shifted to become earlier with age, supporting the former intuition [92]. However, we cannot distinguish whether the age-related sleep patterns we observe are primarily driven by physiological or social developments associated with different stages of adulthood. Fourth, similar to others, we use weekends as a proxy for free days where individuals were not working to help distinguish between endogenously and exogenously driven changes [20, 24]. This assumption does not hold for the subset of our sample who might be unemployed or otherwise follow irregular (e.g. service industry) work schedules. Thus, our estimates of misalignment may be slightly conservative. Despite these limitations, our primary results converge with recognized age-related trends in both sleep duration and timing [9, 28]. Furthermore, these trends appear to be generally consistent across multiple geographic regions and sociocultural contexts.

Implications

Massive data sets generated by pervasive consumer wearable devices can provide globally consistent measurements and thus can contribute unique and confirmatory insights about the development of human sleep patterns. Interestingly, the wide and overlapping distributions of sleep times between genders across the life course suggests that even though there are characteristic differences in mean values, overgeneralization of gender differences should be avoided. Underlying heterogeneity in sleep duration and timing across the life course proves the rule rather than the exception. Early weekday work schedules and norms likely constrain the varied circadian preferences of individuals, contributing to misalignment. Furthermore, given the pervasive asymmetry between weekend and weekday sleep patterns as well as variability in day-to-day sleep timing, research on behavioral interventions that promote regular sleep wake cycles is needed. Rather than impose standard morning start times, organizations might explore and evaluate person-centered work schedules and jobs that match the diverse circadian preferences of individuals, evident in this study and others.

Supplementary Material

Supplementary material is available at *SLEEP* online.

Funding

Sune Lehmann acknowledges funding from the Independent Research Fund Denmark (Microdynamics of Influence in Social Systems) and the Villum Foundation (Nation-scale Social Networks).

Conflict of interest statement. None declared.

References

1. Irwin MR. Why sleep is important for health: a psychoneuroimmunology perspective. *Annu Rev Psychol*. 2015;**66**:143–172.
2. Domínguez F, *et al*. Association of sleep duration and quality with subclinical atherosclerosis. *J Am Coll Cardiol*. 2019;**73**(2):134–144.
3. Cappuccio FP, *et al*. Sleep duration and all-cause mortality: a systematic review and meta-analysis of prospective studies. *Sleep*. 2010;**33**(5):585–592.
4. Vitiello MV. Sleep in normal aging. *Sleep Med Clin*. 2006;**1**(2):171–176.
5. Espiritu JRD. Aging-related sleep changes. *Clin Geriatr Med*. 2008;**24**(1):1–14.
6. Li J, *et al*. Sleep in normal aging. *Sleep Med Clin*. 2018;**13**(1):1–11.
7. Floyd JA, *et al*. Age-related changes in initiation and maintenance of sleep: a meta-analysis. *Res Nurs Health*. 2000;**23**(2):106–117.
8. Foley D, *et al*. Sleep disturbances and chronic disease in older adults: results of the 2003 National Sleep Foundation Sleep in America Survey. *J Psychosom Res*. 2004;**56**(5):497–502.
9. Ohayon MM, *et al*. Meta-analysis of quantitative sleep parameters from childhood to old age in healthy individuals: developing normative sleep values across the human lifespan. *Sleep*. 2004;**27**(7):1255–1273.
10. Floyd JA, *et al*. Changes in REM-sleep percentage over the adult lifespan. *Sleep*. 2007;**30**(7):829–836.
11. Pace-Schott EF, *et al*. Age-related changes in the cognitive function of sleep. *Prog Brain Res*. 2011;**191**:75–89.
12. Dijk DJ, *et al*. Age-related reduction in daytime sleep propensity and nocturnal slow wave sleep. *Sleep*. 2010;**33**(2):211–223.
13. Wild CJ, *et al*. Dissociable effects of self-reported daily sleep duration on high-level cognitive abilities. *Sleep*. 2018;**41**(12). doi: 10.1093/sleep/zsy182
14. Gadie A, *et al*.; Cam-CAN. How are age-related differences in sleep quality associated with health outcomes? An epidemiological investigation in a UK cohort of 2406 adults. *BMJ Open*. 2017;**7**(7):e014920.
15. Mander BA, *et al*. Sleep and human aging. *Neuron*. 2017;**94**(1):19–36.
16. Adan A, *et al*. Circadian typology: a comprehensive review. *Chronobiol Int*. 2012;**29**(9):1153–1175.
17. Roenneberg T, *et al*. Epidemiology of the human circadian clock. *Sleep Med Rev*. 2007;**11**(6):429–438.
18. Kantermann T, *et al*. Comparing the morningness-eveningness questionnaire and munich chronotype questionnaire to the dim light melatonin onset. *J Biol Rhythms*. 2015;**30**(5):449–453.
19. Luca G, *et al*. Age and gender variations of sleep in subjects without sleep disorders. *Ann Med*. 2015;**47**(6):482–491.
20. Kuula L, *et al*. Using big data to explore worldwide trends in objective sleep in the transition to adulthood. *Sleep Med*. 2019;**62**:69–76.

21. Roenneberg T, *et al*. Life between clocks: daily temporal patterns of human chronotypes. *J Biol Rhythms*. 2003;**18**(1):80–90.
22. Wittmann M, *et al*. Social jetlag: misalignment of biological and social time. *Chronobiol Int*. 2006;**23**(1–2):497–509.
23. Roenneberg T, *et al*. Social jetlag and obesity. *Curr Biol*. 2012;**22**(10):939–943.
24. Fischer D, *et al*. Chronotypes in the US—Influence of age and sex. *PLoS One*. 2017;**12**(6):e0178782.
25. Bei B, *et al*. Beyond the mean: a systematic review on the correlates of daily intraindividual variability of sleep/wake patterns. *Sleep Med Rev*. 2016;**28**:108–124.
26. Knutson KL, *et al*. Intra-individual daily and yearly variability in actigraphically recorded sleep measures: the CARDIA study. *Sleep*. 2007;**30**(6):793–796.
27. Dillon HR, *et al*. Variability in self-reported normal sleep across the adult age span. *J Gerontol B Psychol Sci Soc Sci*. 2015;**70**(1):46–56.
28. Roenneberg T, *et al*. A marker for the end of adolescence. *Curr Biol*. 2004;**14**(24):R1038–R1039.
29. Tonetti L, *et al*. Sex difference in sleep-time preference and sleep need: a cross-sectional survey among Italian pre-adolescents, adolescents, and adults. *Chronobiol Int*. 2008;**25**(5):745–759.
30. Bailey M, *et al*. Sex differences in circadian timing systems: implications for disease. *Front Neuroendocrinol*. 2014;**35**(1):111–139.
31. Cain SW, *et al*. Sex differences in phase angle of entrainment and melatonin amplitude in humans. *J Biol Rhythms*. 2010;**25**(4):288–296.
32. Duffy JF, *et al*. Sex difference in the near-24-hour intrinsic period of the human circadian timing system. *Proc Natl Acad Sci U S A*. 2011;**108**(Suppl 3):15602–15608.
33. Adan A, *et al*. Gender differences in morningness-eveningness preference. *Chronobiol Int*. 2002;**19**(4):709–720.
34. Paine SJ, *et al*. The epidemiology of morningness/eveningness: influence of age, gender, ethnicity, and socioeconomic factors in adults (30–49 years). *J Biol Rhythms*. 2006;**21**(1):68–76.
35. Zimmermann LK. Chronotype and the transition to college life. *Chronobiol Int*. 2011;**28**(10):904–910.
36. Merikanto I, *et al*. Relation of chronotype to sleep complaints in the general Finnish population. *Chronobiol Int*. 2012;**29**(3):311–317.
37. Mallampalli MP, *et al*. Exploring sex and gender differences in sleep health: a Society for Women's Health Research Report. *J Womens Health (Larchmt)*. 2014;**23**(7):553–562.
38. Pengo MF, *et al*. Sleep in women across the life span. *Chest*. 2018;**154**(1):196–206.
39. Girschik J, *et al*. Validation of self-reported sleep against actigraphy. *J Epidemiol*. 2012;**22**(5):462–468.
40. Lauderdale DS, *et al*. Self-reported and measured sleep duration: how similar are they? *Epidemiology*. 2008;**19**(6):838–845.
41. Stopczynski A, *et al*. Measuring large-scale social networks with high resolution. *PLoS One*. 2014;**9**(4):e95978.
42. Abdullah S, *et al*. Towards circadian computing: early to bed and early to rise makes some of us unhealthy and sleep deprived. In: proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing; 2014:673–684.
43. Aledavood T, *et al*. Digital daily cycles of individuals. *Front Phys*. 2015;**3**:73.
44. Christensen MA, *et al*. Direct measurements of smartphone screen-time: relationships with demographics and sleep. *PLoS One*. 2016;**11**(11):e0165331.
45. Walch OJ, *et al*. A global quantification of "normal" sleep schedules using smartphone data. *Sci Adv*. 2016;**2**(5):e1501705.
46. Cuttone A, *et al*. SensibleSleep: a Bayesian model for learning sleep patterns from smartphone events. *PLoS One*. 2017;**12**(1):e0169901.
47. Monsivais D, *et al*. Seasonal and geographical impact on human resting periods. *Sci Rep*. 2017;**7**(1):10717.
48. Fagherazzi G, *et al*. An international study on the determinants of poor sleep amongst 15,000 users of connected devices. *J Med Internet Res*. 2017;**19**(10):e363.
49. Roenneberg T, *et al*. Human activity and rest in situ. *Methods Enzymol*. 2015;**552**:257–283.
50. United Nations, Department of Economic and Social Affairs, Population Division. *World Population Prospects 2019, Online Edition*. Median Age of Population. 2019.
51. Ahmad OB, *et al*. Age standardization of rates: a new WHO standard. *Geneva World Heal Organ*. 2001;**9**(10):10.
52. WHO. *Global Health Observatory Data Repository: Mean Body Max Index Trends Among Adults, Age-Standardized. Estimates by Country*. Geneva, Switzerland: World Health Organization; 2019.
53. Molzof HE, *et al*. Intraindividual sleep variability and its association with insomnia identity and poor sleep. *Sleep Med*. 2018;**52**:58–66.
54. Consultation WHO. *Obesity: Preventing and Managing the Global Epidemic*. Geneva, Switzerland: WHO Technical Report Series 894; 2000.
55. WHO Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *Lancet (London, England)*. 2004;**363**(9403):157.
56. R Core Team. *R: A Language and Environment for Statistical Computing*. 2018. https://www.r-project.org/.
57. Faraway JJ. *Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models*. Boca Raton, FL: Chapman and Hall/CRC; 2016.
58. Kuznetsova A, *et al*. lmerTest package: tests in linear mixed effects models. *J Stat Softw*. 2017;**82**(13):2–3.
59. Luke SG. Evaluating significance in linear mixed-effects models in R. *Behav Res Methods*. 2017;**49**(4):1494–1502.
60. Depner CM, *et al*. Wearable technologies for developing sleep and circadian biomarkers: a summary of workshop discussions. *Sleep*. 2020;**43**(2). doi:10.1093/sleep/zsz254
61. Ancoli-Israel S, *et al*. The role of actigraphy in the study of sleep and circadian rhythms. *Sleep*. 2003;**26**(3):342–392.
62. Althoff T, *et al*. Harnessing the web for population-scale physiological sensing: a case study of sleep and performance. In: Proceedings of the 26th International Conference on World Wide Web; 2017:113–122.
63. Ong JL, *et al*. Large-scale data from wearables reveal regional disparities in sleep patterns that persist across age and sex. *Sci Rep*. 2019;**9**(1):3415.
64. Lo JC, *et al*. Young adults' sleep duration on work days: differences between East and West. *Front Neurol*. 2014;**5**:81.
65. Shrivastava D, *et al*. How to interpret the results of a sleep study. *J Community Hosp Intern Med Perspect*. 2014;**4**(5):24983.
66. Kolla BP, *et al*. Consumer sleep tracking devices: a review of mechanisms, validity and utility. *Expert Rev Med Devices*. 2016;**13**(5):497–506.

67. Martinez GM, *et al*. Fertility of men and women aged 15–44 in the United States: national survey of family growth, 2011–2015. *Natl Health Stat Report*. 2018;**51**(113):1–17.
68. Galland BC, *et al*. Normal sleep patterns in infants and children: a systematic review of observational studies. *Sleep Med Rev*. 2012;**16**(3):213–222.
69. Eurofund. *Fifth European Working Conditions Survey. of the European Union PO, ed*. 2012. https://www.eurofound.europa.eu/sites/default/files/ef_publication/field_ef_document/ef1182en.pdf. Accessed July 18, 2020.
70. U.S. Bureau of Labor Statistics. Job flexibilities and work schedules summary—2017–2018. 2019. http://apps.who.int/gho/data/view.main.CTRY12461?lang=en. Accessed July 1, 2020.
71. Wong PM, *et al*. Social jetlag, chronotype, and cardiometabolic Risk. *J Clin Endocrinol Metab*. 2015;**100**(12):4612–4620.
72. Gooley JJ. How much day-to-day variability in sleep timing is unhealthy? *Sleep*. 2016;**39**(2):269–270.
73. Lemola S, *et al*. Variability of sleep duration is related to subjective sleep quality and subjective well-being: an actigraphy study. *PLoS One*. 2013;**8**(8):e71292.
74. Telzer EH, *et al*. Sleep variability in adolescence is associated with altered brain development. *Dev Cogn Neurosci*. 2015;**14**:16–22.
75. Becker SP, *et al*. Intraindividual variability of sleep/wake patterns in relation to child and adolescent functioning: a systematic review. *Sleep Med Rev*. 2017;**34**:94–121.
76. Patel SR, *et al*.; Osteoporotic Fractures in Men (MrOS); Study of Osteoporotic Fractures (SOF) Research Groups. The association between sleep patterns and obesity in older adults. *Int J Obes (Lond)*. 2014;**38**(9):1159–1164.
77. Okun ML, *et al*. Sleep variability, health-related practices, and inflammatory markers in a community dwelling sample of older adults. *Psychosom Med*. 2011;**73**(2):142–150.
78. Veeramachaneni K, *et al*. Intraindividual variability in sleep and perceived stress in young adults. *Sleep Health*. 2019;**5**(6):572–579.
79. Slavish DC, *et al*. Intraindividual variability in sleep and comorbid medical and mental health conditions. *Sleep*. 2019;**42**(6). doi:10.1093/sleep/zsz052
80. Murray JM, *et al*.; Delayed Sleep on Melatonin (DelSoM) Study Group. Sleep regularity is associated with sleep-wake and circadian timing, and mediates daytime function in delayed sleep-wake phase disorder. *Sleep Med*. 2019;**58**:93–101.
81. Taylor BJ, *et al*. Bedtime variability and metabolic health in midlife women: the SWAN sleep study. *Sleep*. 2016;**39**(2):457–465.
82. Moline ML, *et al*. Sleep in women across the life cycle from adulthood through menopause. *Sleep Med Rev*. 2003;**7**(2):155–177.
83. Zhang B, *et al*. Sex differences in insomnia: a meta-analysis. *Sleep*. 2006;**29**(1):85–93.
84. Kendall-Tackett KA. *Depression in New Mothers: Causes, Consequences and Treatment Alternatives*. Oxfordshire, UK: Routledge; 2016.
85. Richter D, *et al*. Long-term effects of pregnancy and childbirth on sleep satisfaction and duration of first-time and experienced mothers and fathers. *Sleep*. 2019;**42**:5–7.
86. Newport DJ, *et al*. The treatment of postpartum depression: minimizing infant exposures. *J Clin Psychiatry*. 2002;**63**(Suppl 7):31–44.
87. Goyal D, *et al*. Fragmented maternal sleep is more strongly correlated with depressive symptoms than infant temperament at three months postpartum. *Arch Womens Ment Health*. 2009;**12**(4):229–237.
88. Venn S, *et al*. The fourth shift: exploring the gendered nature of sleep disruption among couples with children. *Br J Sociol*. 2008;**59**(1):79–97.
89. Burgard SA, *et al*. Gender and time for sleep among U.S. adults. *Am Sociol Rev*. 2013;**78**(1):51–69.
90. Hagen EW, *et al*. Changes in sleep duration and sleep timing associated with retirement transitions. *Sleep*. 2016;**39**(3):665–673.
91. Macridis S, *et al*. Consumer physical activity tracking device ownership and use among a population-based sample of adults. *PLoS One*. 2018;**13**(1):e0189298.
92. Broms U, *et al*. Long-term consistency of diurnal-type preferences among men. *Chronobiol Int*. 2014;**31**(2):182–188.

Travel serves to balance skewed sleep patterns

Sigga Svala Jónasdóttir, James P. Bagrow and Sune Lehmann

February 27, 2021

Abstract

The interplay of travel and sleep is well known, with travel generally expected to have a deleterious effect on sleep. However, a detailed understanding of the changes in sleep associated with travel at an epidemiological scale has been limited by a lack of large-scale data. Here we show that travel has a balancing effect on sleep, where underslept individuals attain more sleep while travelling compared to at-home behaviour, while individuals who tend to sleep more than 7.5 hours on average, sleep less when sleeping away from home. The analysis of our global data-set with 3.2 million nights (220.000 travel-nights) for 19.000 users reveals a systematic change in sleep duration and timing due to travel relative to typical at-home behaviour. The change in sleep quantity for travel-nights depends linearly on typical nighttime sleep at home and decreases as median sleep duration increases. On average, wake-up time advances to later hours on weekdays when travelling, but moves to earlier hours on weekends. Our study emphasises the potential of identifying novel sleep behaviours in large behavioural data-sets from consumer wearable devices, and may inspire future studies to further examine how environment and behaviour affects human sleep.

Introduction

Attaining sufficient sleep is critical to many aspects of human health [1–3]. Short and irregular sleep duration contributes to molecular, immune, and neurological changes that play a role in disease development, increasing, for example, the risk of obesity and cardiovascular diseases, and substantially affecting mood, motor and cognitive performance [1, 4–10]. Despite the importance of sleep to health, average sleep duration has continued to decrease among economically developed countries: for example, 30% of the US population slept on average less than 6 hours in 2013, compared to 3% in 1963 [11–13].

Concurrently, travel has increased dramatically over the past two decades, with the number of air-travelers nearly tripling [14]. There are good reasons to think that traveling impacts sleep negatively. Travel and new resting environments are known to influence sleep quantity and quality. Indeed, the *First night effect* (FNE) was first documented in 1964 where sleep-initiation difficulty and prolonged sleep-onset latency was found to occur on the first night of sleep taking place in sleep laboratory [15, 16]. Later, in 2016, Tamaki *et al.* showed that FNE is a consequence of a single brain-hemisphere displaying elevated alertness in new and unfamiliar environments. The hemisphere with reduced sleep depth showed more enhanced response to external stimuli during resting period [17].

Travel fatigue and *Jet-lag* are conditions which can cause sleep complications when travelling [18–24]. Travel fatigue is associated with any long journey, regardless of the mode of transport characterised by tiredness, disorientation and headaches which usually last only for a day or so, but when flying across several time zones there is the added effect of jet-lag with longer-lasting ramification [18–21]. Jet-lag is due to desynchronisation of the body's internal clock and the new time zone an individual enters after long-distance travel [22–24].

Jet-lag is not only limited to travel. *Social jetlag* is a term and measure used in sleep epidemiology to quantity the difference between weekend-weekday behaviour, and if measured high, likely happening due to constraints of early-morning work schedule on weekdays, which are relieved on weekends [25, 26].

Most of the existing research to understand the effect of travel on sleep, aims to understand the physiological and behavioural changes among professional athletes, and has been carried out as small scale studies (typically 10-30 study subjects) or to understand subjective fatigue and alertness among aircrew staff [18, 27–34]. These studies have found no significant difference in sleep quantity and quality before and after short-haul air travel (without crossing of time zones) [28, 35–38]. However, if journeys cross time-zones, the outcome is different. Jet-lag was found to cause sleep issues in new time zones, including reduced sleep duration, more frequent and longer nighttime awakenings, delayed sleep onset after eastward travel, and advanced sleep offset after westward travel [22, 35, 39]. While multiple effects have been discussed, the quantitative changes in sleep due to travel have not been researched in an epidemiological context. Here, we address this gap in the literature through a large and global data-set of sleep activity data recorded with wearable devices. The dataset consists of $\sim$ 19 000 users residing in 95 countries with more than 3.2 million nights and thereof $\sim$ 220 000 away from home

Our work sheds new light on the effect of travel and new resting environments on sleep behaviour. Specifically, we find that sleep during travel tends to depend on sleep patterns at home, specifically that it serves a balancing function: People with shorter than average home-sleep duration tend to have longer nighttime sleep during travel, while those who have longer than average home-sleep duration, tend to sleep less during travel.

Methods

Data collection

The dataset was collected from 2015 to 2019 via wristbands designed to track physical activity and sleep behavior. The wristbands use proprietary, internally validated algorithms based on movement registered by an internal accelerometer to estimate sleeping and waking states in 1-minute intervals. The 1-minute sleep states are used to infer sleep onset, offset and duration for each night where nighttime awakening or sleep fragmentation is also accounted for and quantified as wake after sleep onset (WASO). When first connecting the wristband to their smartphone, users receive visual instruction on how and where (wrist) to place the device and they are advised to wear it on their dominant side. Measurements produced by the wristbands exhibit a high degree of face validity and converge with estimates of age-related changes from the literature [40]. Measurements have also been validated by comparing country-level estimates of sleep onset, offset and duration to numbers from other publications [40]. By using these wristbands, we follow a growing trend of utilizing commercial devices in sleep research to study sleep behavior in naturalistic settings at large scales [40–43].

Users are anonymous and self-report their age, gender, height and weight. The location data originates from GPS traces; these are not collected at a fixed sampling rate but estimates are updated when there is a change is the motion-state of the device (if the accelerometer registers a change). All data analysis was carried out in accordance with the EU's General Data Protection Regulation 2016/679 (GDPR) and the regulations set out by the Danish Data Protection Agency. The GDPR describes regulations for data protection and privacy in the European Union and the European Economic Area; it also addresses the transfer of personal data outside the EU and EEA areas.

Data pre-processing

To reduce the risk of including sleep observations from those suffering from insomnia, artificially shortened night observations due to users ceasing wristband use in the middle of the resting period, observations from night-shift workers or any other possible data errors, outliers were removed. The details of the process is described step-by-step in the SI (*Data Pre-processing*).

We transformed the raw location data to *stop-locations* using the `infostop` algorithm [44], converting traces to stops, each with an ID, start, and end time. We discard sleep observations without associated stop-locations. We define a person's *sleep location* as the stop location with start-time closest to the sleep onset. To ensure consistency we only accept locations where sleep begins and user does not leave the location until after the sleep has ended. We expect people to sleep at home for the majority of the time, and therefore use sleep location to infer home location. The location where most nights take place is defined to be a user's *home location*. We remove users from the dataset if their percentage of nights-at-home is lower than 70%. We use this threshold to ensure that we select individuals with a fixed home location and retain approximately 80 % of the users by applying this selection criteria. Henceforth, we refer to nights that take place at least 20 km away from home as *travel-nights*.

We use the median sleep duration to quantify typical sleep duration (for nights recorded at home). In order for the median to be representative of an individual's typical behaviour we require all participants to have a minimum of 10 nights recorded at home; in this we treat weekends and weekdays separately. In the SI we provide evidence that 10 nights is a reasonable threshold (*Filtering & Inclusion Criteria*).

As we wish to understand the quantitative effect of travel on sleep duration and timing, we also require users to have a minimum number of nights recorded away from home. We set this minimum to two travel days (by day-type; weekdays/weekends). Again, we justify this choice using robustness checks and down-sampling as shown in the SI (see *Results*). Note we separate analyses into day-type (weekend vs. weekday) and users may be included in the analysis for a single such day-type or both.

After the pre-processing, the final data-set used for analyses consists of 2.4 million weekday nights (6.0% away from home) from about 19 300 users and 0.8 million weekend nights (9.3% away from home) from 13 300 users. An in-depth exploration of how users are distributed by demographics and data coverage is presented in the SI (*Data Coverage & Demographics*).

Data modeling

In order to support our main findings we employ a mixed effects model – a panel data analysis with a hierarchical linear model where the relationship between the change in sleep duration away from home (relative to regular behaviour) and typical sleep duration at home is explored [45].

The mixed effects model enables us to retain the hierarchical structure of the data – repeated measurements within users and instead of estimating Δ as a single measurement per user ($\Delta = \mu - M_{home}$ where μ is the average sleep duration for travel-nights), we estimate it for every recorded night for each user, defined as

$$\Delta_{i,j} = duration_{i,j} - M_j,$$

where $i = 1, ..., N$ and $j = 1, ..., K$ and where N is the total number of nights for user j and K

is the total number of users. The mixed effects model is then specified in matrix form as

$$\mathbf{y} = X\beta + Zu + \epsilon, \qquad \text{with } u \sim N_q(0, G) \text{ and } \epsilon \sim N_n(0, R), \tag{1}$$

with β representing the fixed effects parameters, u representing the random effects, X representing the $n \times p$ design matrix for the fixed-effects parameters, and Z the $n \times q$ design matrix describing the random effects. The dependent variable is $\Delta_{i,j}$ and fixed effects parameters are the demographic variables (gender, generation, BMI category, region of residence) and home. The independent variable is median sleep duration (continuous) and we control for individual baseline behaviour since user is a random effect (intercept).

We analyzed the data set using R Version 3.5.1 and the `lmerTest` package, the `lmer` function to fit the data set and apply Satterthwaite's degrees of freedom method to estimate the p-values for the significance of fixed effects [46–48]. The model is reduced by removing insignificant fixed effects (one at a time) with the `drop1` function which utilizes F-test (one-sided) for its estimates. We center medium sleep duration around its sample mean to help improve interpretability and prevent multi-collinearity.

Results

Measuring change in sleep duration due to travel On Figure 1A we present an example of data collected for a single user. We use the the median sleep duration, M_{home} to quantify the typical sleep duration at home. In order to evaluate the behaviour when travelling we estimate average sleep duration for travel-nights (denoted μ_{travel}). We define $\Delta_s = \mu_s - M_{home}$ as the change in sleep duration relative to typical behaviour, where the state $s \in \{home, travel\}$. The variable Δ_s is estimated for each user in our sample. We explain the rationale for comparing mean to median for both home and travel nights below.

Sleep-behavior during travel depends linearly on sleep at home. We first explore whether the change in sleep duration away from home depends on typical sleep duration at home by plotting the distribution of $\Delta_s = \mu_s - M_{home}$ where $s \in \{home, travel\}$ for individuals with different median sleep duration. The results are shown in Figure 1B, where users are grouped into *sleep groups* by rounding their median to the nearest half-hour bin. The distributions are broad, but we see a clear trend that the average Δ_{travel} moves from positive to negative values as the median sleep duration increases. This implies that individuals who sleep little at home (duration $\leq$ 5.0 hours) tend sleep longer when they are away from home. On the opposite end of the spectrum, those who sleep longer at home (duration $\geq$ 9 hours), sleep less when they are away from home. To quantify this trend, we calculate the average Δ_{travel} for each sleep group (ranging from 4.5 - 9.5 hours) which reveals an approximately

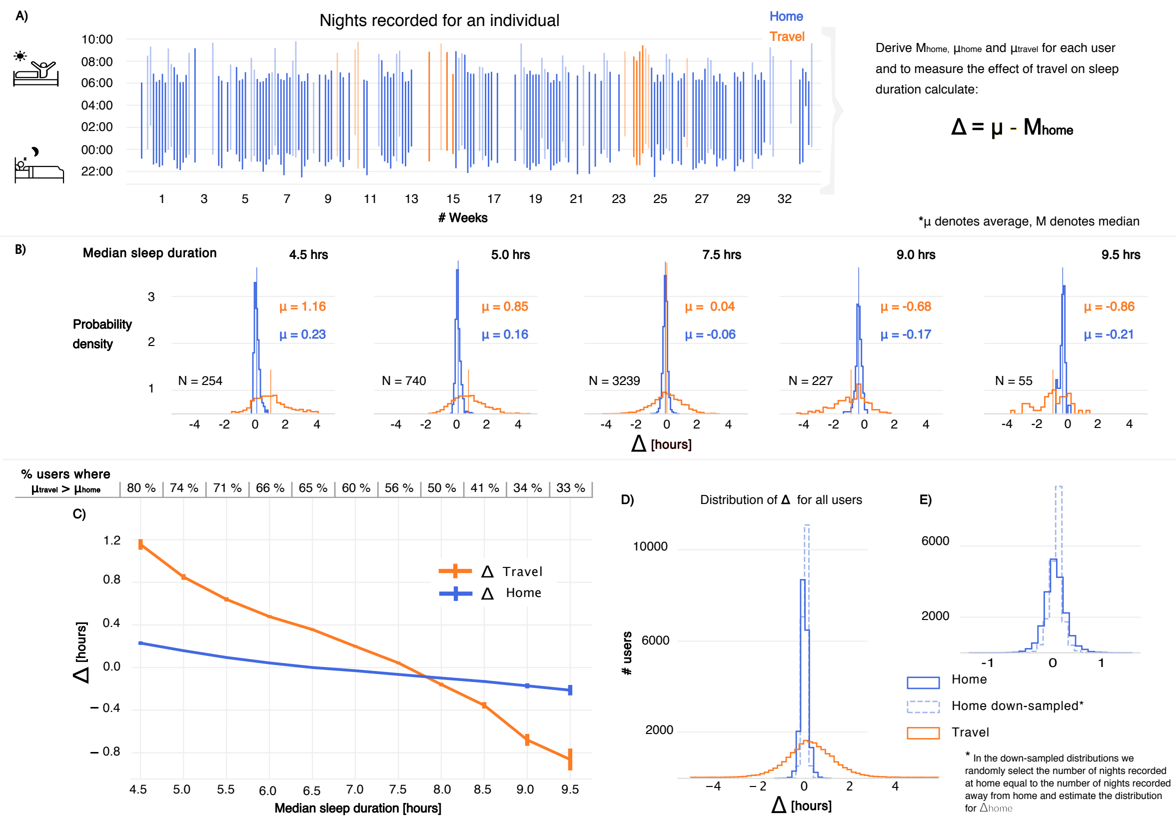


Figure 1: Sleep activity patterns and the relative change in sleep duration for travel-nights
A: For every individual we measure sleep onset, offset and duration for each recorded night. From these records we derive three measurements; median sleep duration at home (M_{home}) and average sleep duration at home and away from home (μ_{home} & μ_{travel}). To measure change in sleep duration due to travel, relative to typical behaviour at home, we derive a new measure $\Delta_{travel} = \mu_{travel} - M_{home}$. **B**: The distributions of Δ_{travel} (orange) and Δ_{home} (blue) for groups with different median sleep duration where users are grouped together by rounding their median to the nearest half-hour bin (referred to as *sleep groups*). **C**: The average Δ_{travel} for all sleep groups (median duration ranging from 4.5 - 9.5 hours) with the standard error of the mean (SEM). **D**: The distribution for Δ_{home}, $\Delta_{home\ DS}$ and Δ_{travel} for all users. **E**: A larger visual representation (more narrow range of the x-axis) for the distribution for Δ_{home} and $\Delta_{home\ DS}$ from panel D.

linear dependence of Δ_{travel} on typical sleep duration at home; error bars in Figure 1C show the standard error of the mean (SEM).

Baseline effect for home nights. In Figures 1B & C, we also plot $\Delta_{home} = \mu_{home} - M_{home}$ (blue color). This is to illustrate a baseline effect, which relates to the observed systematic change in sleep duration away from home. This baseline effect is a decreasing linear trend of Δ_{home} (blue line in Figure 1C), which shows that there is a systematic difference between mean and median as a function of median sleep duration for nights at home.

Our hypothesis is that the slope of Δ_{home} arises because of *sleep-wake homeostasis*, a physiological process which regulates sleep pressure. For example, a person who tends to sleep less than physiologically needed, will build up sleep pressure from the last adequate sleep episode which can be eliminated by a long nighttime sleep (a 'catch-up' night) [49,50]. These 'catch-up' nights can result in a skewed distribution of sleep duration, with a disproportionately larger right tail; a positive skew (exemplified on Figure S6). Similarly, we expect a negative skew (a heavy left tail of the distribution) for individuals who tend to have longer nighttime sleep than they can sustain.

This behavior is confirmed on Figure S6, which shows that 95 % of users sleeping 4.5 hours at home have a longer average than median sleep duration and on the contrary, 93 % of those sleeping 9.5 hours have shorter average than median sleep duration. This explains why Δ_{home} is positive for median sleep duration of less than 7 hours in Figure 1C, and negative for median sleep duration longer than 7 hours. The weak linear trend of Δ_{home} and median sleep duration on Figure 1C (which we believe is due to the process of sleep-wake homeostasis) explains our comparison of Δ_{travel} with Δ_{home} – to obtain the absolute effect of travel on sleep.

Results are robust despite imbalanced sample size of travel and home nights. To directly compare Δ_{travel} and Δ_{home} we plot both distributions together in Figure 1D. Visually, the two distributions are very different with a much broader distribution for travel-nights. To rule out that our results are due to this imbalance in sample sizes (e.g. that the broad range of Δ_{travel} is to due to lower sampling rate for travel nights), we perform an individual-level down-sampling of nights at home to balance our data sample. The distribution of down-sampled home-values, $\Delta_{home\ DS}$, (light blue colored distribution) is shown on Figure 1D & E. The down-sampled distribution is, in fact, slightly narrower than Δ_{home} and remains quite different from the broad range of behaviour observed for the distribution of Δ_{travel} (for a more detailed description see *Down-sampling nights at home* in SI).

The illustrations in Figure 1 only shows behaviour on weekdays since we follow the convention of sleep research and analyze weekdays and weekends separately. In the next section, we include data from weekends to understand the effect of travel on weekend nights.

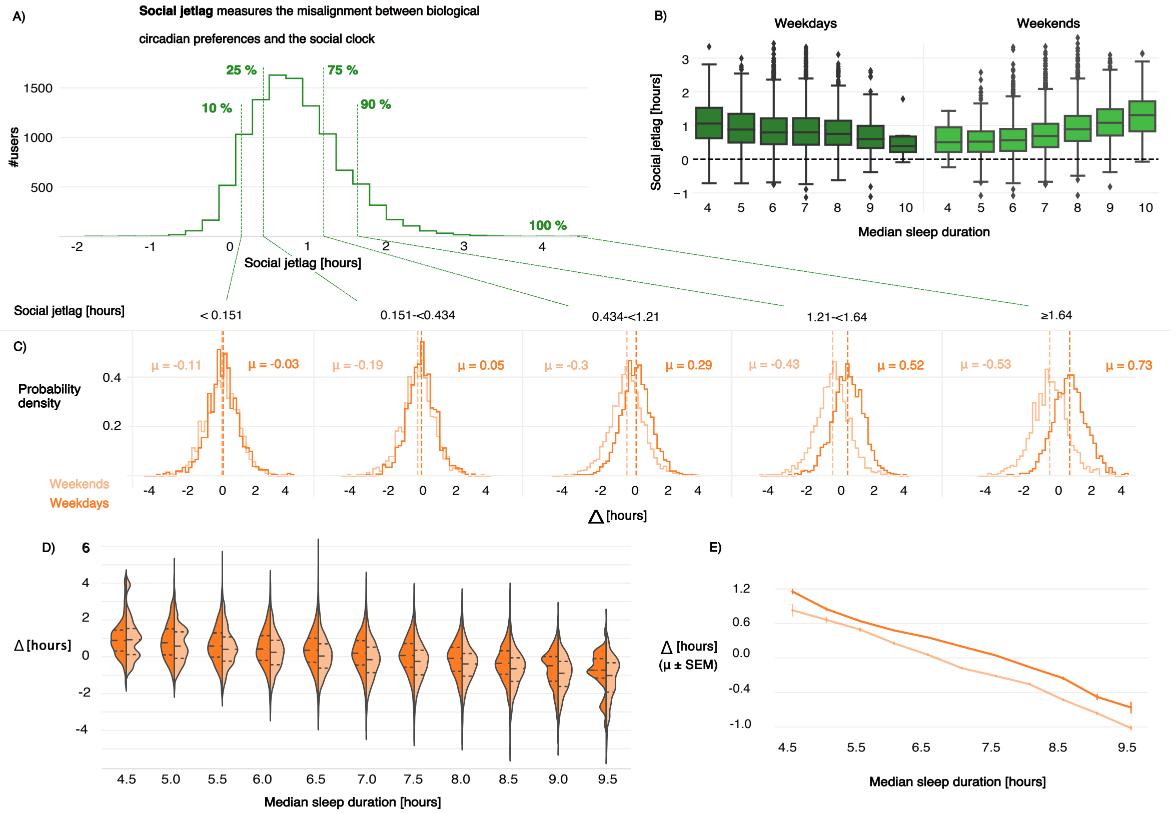


Figure 2: Disproportionate effect of travel on individuals with high social jetlag and the connection between weekends and weekdays

A: Defines social jetlag and visualises how it distributes across the user sample. **B**: Box-plots of social jetlag for users with different range of sleep duration. The horizontal lines represents the median, box limits correspond to upper and lower quartiles, whiskers define the 1.5x interquartile range and points are outliers. **C**: Distribution for Δ_{travel} on weekends and weekdays for groups of users with different range of social jetlag (defined by percentiles) **D**: The distributions of Δ_{travel} for groups with different range of sleep duration by day type – weekends (lighter orange color), weekdays (darker orange color) and the dotted lines mark the quartiles of the distributions. **E**: The average Δ_{travel} plotted with the standard error of the mean (SEM) by sleep groups (half-hour bins for median sleep duration) on weekdays (dark orange color) and weekends (light orange color)

Amount of social jetlag impacts the change in sleep during travel. Social jetlag (SJ) was conceptualised by Wittman et al. (2006) and quantifies the difference between 'biological time preferences' and 'the social clock'. Stated more plainly, social jetlag measures the difference between weekday and weekend sleep behaviour [25]. In Figure 2A we show how social jetlag is distributed across our sample. Most users (80 %) have some amount of social jetlag ranging from 9-98 minutes. Figure 2B shows that social jetlag depends on sleep duration and individuals with high social jetlag typically sleep little on weekdays (4-5 hours) and a lot on weekends (9-10 hours). This large quantitative difference is usually attributed to constraints from an early work schedule on weekdays (social clock), causing substantial sleep deprivation on weekdays and sleep compensations during weekends (biological time preferences) [25,26]. In Figure 2C we plot the distribution of Δ_{travel} for weekdays (dark orange color) and weekends (light orange color) for groups of users with different range of social jetlag (defined by percentiles). We observe a larger effect of travel on sleep duration for individuals with high values of social jetlag (SJ), and users in the top 10th percentile (SJ > 98 minutes) gain on average 44 minutes of sleep when nights take place away from home on weekdays but lose 32 minutes of sleep on weekends.

Effects of travel on weekend nights. Next we examine how sleep duration changes for travel-nights on weekends and compare it to the patterns observed previously for weekday-nights (see Figure 1). Figure 2D shows the distributions for Δ_{travel} on weekdays (dark orange color) and weekends (light orange color) organized by sleep groups, where the dotted black lines represent the distribution quartiles. Figure 2E illustrates the averages for Δ_{travel} by sleep groups with the SEM. The relationship between Δ_{travel} and typical sleep duration on weekends is fundamentally the same as for weekdays; the change in sleep duration during travel decreases as the sleep duration at home increases. However, the relative change is slightly larger in the positive direction (line pushed further up on y-axis) on weekdays compared to weekends when observing the distribution averages and quartiles on Figure 2D and E. These differences can be explained by the fact that people are usually more constrained by time and alarm clocks on weekdays, consequently sleeping less than they might need and therefore more susceptible to gain sleep - the opposite is expected for weekends; more room to lose sleep [25,26].

Sleep onset shows a similar behavior to duration for travel nights. Above we have observed the systematic change in sleep duration for travel-nights, but sleep duration is derived from two variables; bed time (sleep onset) and wake-up time (sleep offset). We now investigate whether the effect of travel extends to sleep onset and offset. In order to explore this question, we use the same methodology as above. Thus, we calculate $\Delta_{onset\ travel} = \mu_{onset\ travel} - M_{onset\ home}$ and $\Delta_{offset\ travel}$. These quantities are then aggregated into averages by user groups, defined by percentiles (10th, 25th, 50th, 75th, & 90th) of the distribution of median sleep duration (see Figure 3A1 for weekdays and A2 for weekends). The average $\Delta_{onset\ travel}$ (blue color) and $\Delta_{offset\ travel}$ (yellow color) are shown with the SEM for weekdays on Figure 3B1 and weekends on Figure 3B2. We find that the change in bed time depends on the duration of home-sleep; those who sleep less than 6.2 hours on weekdays (bottom 25th percentile) go to bed earlier on weekday travel-nights. For those sleeping 7.5 hours or less (bottom 50th percentile), the travel bed-time on weekends is advanced to earlier hours. The dependence of $\Delta_{onset\ travel}$ on typical sleep duration is approximately linear and bed-time advances from earlier to later hours (relative to typical behaviour at home) as typical home-sleep duration increases.

Sleep offset shows opposite behavior on weekdays and weekends for travel nights . Wake-up time during travel tends to be later for all users on weekdays but earlier on weekends (see yellow curves on Figure 3B1 and B2). The users in the bottom 10th percentile on weekdays and top 10th percentile on weekends change their behaviour the most relative to typical hours at home, waking up 33 ± 2 minutes later on weekdays and 46 ± 2 minutes earlier on weekends when nights take place away from home. The top 10th percentile on weekdays and bottom 10th percentile on weekends change their behaviour the least (shift of 8 ± 2 minutes in wake-up time). The middle group of users (10-90th percentile in the distribution of median sleep duration) exhibit more homogeneity on weekdays where the change in wake-up hours on weekdays is 22-28 minutes later, whereas the range is broader on weekends and a slight linear dependence with typical sleep duration at home (wake-up time occurring 19-35 minutes earlier than at home). The difference between the change in sleep timing (onset and offset) due to travel on weekends and weekdays can be explained by the fact that sleep patterns have a tendency to be shifted to earlier hours than is natural to individuals on weekdays due to morning work schedule [25,26]. This constraint seems to extend over to the relative change in sleep timing away from home since bed and wake-up times are almost only shifted to later hours for most groups on weekdays and earlier hours on weekends.

Confirming robustness of results via mixed effects models. Our data set contains males and females, a wide range of ages, and originates from users across the world, and sleep behavior has been shown to depend on these demographic indicators [26, 40, 51–56]. In the analysis above, we ignore this heterogeneity and explore sleep behavior during travel averaged across our entire population. In order to understand the effects of the underlying heterogeneity on our results, we now explore the relationship between the change in sleep duration for travel nights (Δ_{travel}) and typical sleep duration at home (M_{home}) using mixed effects model.

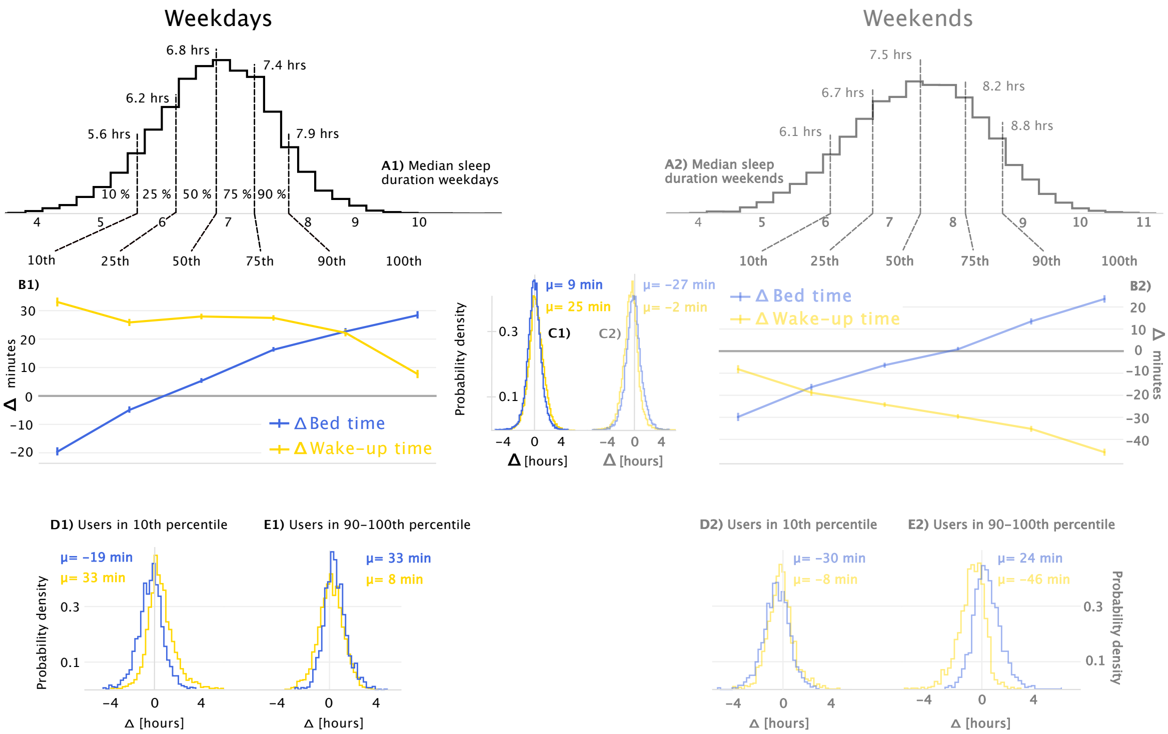


Figure 3: Change in sleep onset and offset for travel-nights
A.1 & A.2: The distributions of sleep duration on weekdays (A.1) and weekends (B.1) with the 10th, 25th, 50th, 75th and 90th percentiles marked with dotted lines. **B1 & B2**: The change in sleep onset and offset relative to typical home-sleep ($\Delta_{onset\ travel}$ and $\Delta_{offset\ travel}$) aggregated into averages (with SEM) by user groups defined by percentiles of sleep duration illustrated on (A1 & A2). **C1 & C2**: The distributions of $\Delta_{onset\ travel}$ and $\Delta_{offset\ travel}$ for all users in the sample. **D1 & D2**: The distributions of $\Delta_{onset\ travel}$ and $\Delta_{offset\ travel}$ for users with the lowest sleep duration on weekdays (D1) and weekends (D2) (bottom 10th percentile) **E1 & E2**: The distributions of $\Delta_{onset\ travel}$ and $\Delta_{offset\ travel}$ for users with the highest sleep duration on weekdays (E1) and weekends (E2) (90-100th percentile)

Specifically, we analyse the effect of the following covariates: generations (Millenials, Gen X & Baby Boomers), gender (Male/Female), East/West (Asia/North America & Europe) and BMI category (Normal Weight/Overweight/Obese). All of these are defined formally in the SI (see *Data Coverage & Demographics*). We implement a model with a three-way interaction term between home (True/False), every demographic variable and median duration centered around the mean (the model is defined formally in *Mixed effects model* in the SI).

Our mixed effects model confirms the large difference between the rate of decrease for Δ at home and away from home; 24 ± 0.5 minute larger decline when travelling than at home (for an hour increment in typical sleep duration) on weekdays and 20 ± 1 minutes on weekends (see Tables S4 and S5 in SI). The region of residence is the most influential covariate in terms of level of significance and effect size both on weekends and weekdays. The difference between East and West is small in terms of the rate of decrease of Δ_{travel} (1 minute difference between the slopes on weekdays and none on weekends), but much greater considering the intercept – which is 36 ± 0.5 minutes higher for users in the West on weekdays and 28 ± 1 minutes higher on weekends. Gender is measured significant as a single term on weekday nights with 7.0 ± 0.5 minute difference between intercepts (higher for women). To provide an overview of these results, we list the model estimates of Δ_s where $s \in \{home, travel\}$ for different median sleep duration (4.5, 7.5 & 9.5 hours) and by most important covariates in Table 1 (also illustrated visually on Figures S7 and S8 in the SI).

	WEEKDAYS					
	TRAVEL			**HOME**		
Sleep duration [hours]	**4.5**	**7.5**	**9.5**	**4.5**	**7.5**	**9.5**
West Men	1.66 ± 0.017	0.218 ± 0.016	-0.74 ± 0.037	0.168 ± 0.0076	-0.0522 ± 0.0023	-0.199 ± 0.089
East Men	1.08 ± 0.013	-0.407 ± 0.028	-1.40 ± 0.056	0.182 ± 0.012	-0.0825 ± 0.0072	-0.259 ± 0.020
West Women	1.82 ± 0.029	0.316 ± 0.030	-0.684 ± 0.069	0.168 ± 0.0076	-0.0522 ± 0.0023	-0.199 ± 0.089
East Women	1.23 ± 0.025	-0.309 ± 0.043	-1.34 ± 0.088	0.182 ± 0.012	-0.0825 ± 0.0072	-0.259 ± 0.020
	WEEKENDS					
West Men	1.00 ± 0.078	-0.528 ± 0.020	-1.55 ± 0.086	0.225 ± 0.0094	-0.045 ± 0.0049	-0.225 ± 0.014
East Men	1.22 ± 0.052	-0.0586 ± 0.0055	-0.910 ± 0.043	0.225 ± 0.0094	-0.045 ± 0.0049	-0.225 ± 0.014

Table 1: Estimates of Δ_s where $s \in home, travel$ from mixed effects model for different sleep groups and demographics (the most important in terms of significance and effect size from model results)

Distance only has a small effect on our main findings. One possible hypothesis is that our results may depend on how far someone travels. To investigate this question, we include distance (three categories) as a covariate in the mixed effects model (< 1000 km, $1000 - 2500$ km & > 2500 km). Initially we explored this question using a model which both included nights at home and away from home. It turned out, however, that in this setting distance simply behaved as an extra proxy for the distinction between home/travel – signalling that this is a small effect. Next, we employed a mixed effects model which only includes travel-nights. The model is defined with two-way interaction term between each covariate and `distance`, as well as a interaction term between `distance` and median sleep duration (the model is defined explicitly in *Mixed effects model: Exploring the effect of distance* in the SI).

The interaction term between distance and median sleep duration is not significant, thus the slope is estimated to be the same for all distance categories. However, distance has a significant effect as a single term. Nights further than 2500 km away from home have 10 minute lower intercept on weekdays and 15 minutes lower on weekends compared to the other two categories, < 1000 km and $1000 - 2500$ km (see estimates of fixed effects in Tables S6 for weekdays and S7 on weekends in the SI). Distance has also a significant interaction term with other covariates and the distinction between East and West is even more enhanced for trips more than 1000 km away from home on weekdays and in the $1000 - 2500$ km category on weekends. On weekdays, for example, the difference between East and West for trips $< 1000km$ away from home is estimated to be the same as for the full model without a covariate for distance (36 minutes), but for trips $1000 - 2500$ km away from home there is added 11 minutes to the baseline (total 47 minutes) and for nights in > 2500 km distance a additional 7 minute difference (total 43 minutes).

Results are robust when varying the amount travel-nights. An important parameter in our analysis is how many nights of travel-sleep a user must have to be included in our data-set. We explore whether our results depend on the minimum number of travel days, we examine the estimates of fixed effects while the inclusion criteria changes, ranging from minimum 2 to 12 travels days per user. For this purpose, we use a simplified version of the model defined in *Robustness* in SI. This analysis shows that our estimates of fixed effects persist but in some instances become slightly smaller in magnitude. In some cases, the estimated effects fall just outside the range of standard error of the mean for the full data-set, see Tables S8 and S9 in the SI. However, the differences with respect to the full data set are small, and overall we confirm our findings. For example, the differences between the slope for home-nights and travel-nights is $[-0.394, -0.386]$ (estimates with SEM) for the full data-set but $[-0.376, -0.364]$ with minimum twelve travel days (for weekdays). This difference is larger on weekends, $[-0.387, -373]$ for two travel days (full data-set) but $[-0.34 : -32]$ for twelve. However, one most consider that there is less data coverage on weekends (5 weekdays versus 2 weekend days a week) and more variability as well, which could be exacerbating the difference [40]. Overall the same results are found when number of travel days required per user is increased, with some indications of a slight change in magnitude.

Discussion

Drawing on a data-set of 3.2 million nights and thereof 220 000 recorded away from home for approximately 19 000 users, we observe a systematic change in sleep duration and timing (onset & offset) for travel nights, relative to typical at-home behaviour. The change in sleep duration due to travel depends linearly on typical sleep quantity at home and decreases as median sleep duration increases - a pattern identified both for weekdays and weekends. Our main finding is that sleep during travel tends to have a balancing effect. Under-slept individuals tend to sleep more than at home when travelling, while individuals whose overall nighttime is characterized by long duration, tend to sleep less when nights take place away from home. The change in sleep onset and offset for travel-nights supports the observed changes in sleep duration. Wake-up time is on average advanced to later hours on weekdays compared to typical nights at home, but to earlier hours on weekends. The change in bed time for travel nights is linearly dependent on typical sleep duration at home, and is advanced to later hours as median sleep duration increases.

The dependence of the change in sleep duration for travel nights on typical sleep duration at home is found both for the case of weekdays and weekends, where individuals are slightly more inclined to gain sleep on weekdays than weekends. This latter finding is likely associated with the constraints of the social clock and is further supported by the fact that misaligned individuals (individuals who have high social jetlag) are disproportionately effected by travel [25, 26]. Our results show that on average wake-up time is shifted to later hours during travel nights on weekdays but to earlier hours on weekends, while the change in bed-time for travel nights is linearly dependent on median sleep duration at home. This highlights the fact that wake-up time is a more controllable factor when it comes to sleep, since individuals can set an alarm to wake-up at specific hour but cannot necessarily fall asleep at a predefined point in time. Sleep onset depends on intrinsic biological rhythm but also influenced by external factors. A previous study indicated the same results, where individuals seemed to catch longer nighttime sleep on weekends by shifting their bed-time marginally more than wake-up time [40].

We observe different effects of travel on sleep by demographic variables where the most significant and influential factor is region of residence – a variable which identifies whether an individual lives in the East (Asia) or West (North America & Europe). Those residing in the East are more inclined to lose sleep when travelling, whereas those in the West tend to gain sleep. We cannot provide a specific explanation for this difference, but speculate that the result may be related to the baseline for at-home behaviour. Individuals in the East cohort sleep on average less than those in the West – 6.4 versus 7.1 hrs weekdays and 6.9 versus 7.8 hrs on weekends – a pattern also identified in other studies [53–56].

While unprecedented in terms of number of users, our work does have some limitations.

First, there are relatively few nights recorded away from home ($\sim$ 7 % out of 3.2 million nights). We require users to have at least two travel nights to be included in the sample, but that sampling rate might not reflect the full range of behaviour for an individual. To mitigate this limitation, we do analyse the effect of travel with panel data analysis using hierarchical linear model which uses all data-points simultaneously to examine the effect of covariates while controlling for individual baseline behaviour and characteristics. We also perform a down-sampling for nights recorded at home (to be equal as number of travel days) which demonstrates that the large distinction between the distribution of Δ at home and away from home persist with the same sampling rate. Most importantly, when changing the inclusion criteria from 2 to 12 travel days while comparing estimates of fixed effects our results remain unchanged. *Second*, approximately 80 % of travel nights are recorded within the distance category of less than 1000 km away from home, and those nights are on average 240 and 280 km away from home (weekdays/weekends). Thus, our sample is biased towards relatively short-distance travel, however, and thus unlikely to be effected by jetlag. *Third*, age and BMI were self-reported which could recall bias. In a previous study, however, we found good agreement between the World Health Organization's country-level estimates of median age and age standardized BMI for our data-set [40]. Our sample of users may also not be representative of the wider population due to potential unobserved factors also associated with wearable device ownership [57]. *Fourth*, we note that the wristbands have not been publicly validated using the gold standard of polysomnography as recommended in the Sleep Research Society Workshop on wearable devices for the measurement of sleep [58]. However, we find 1) our data-set converges with country-level sleep measures from separate large-scale data-sets, 2) demonstrates consistency over the period of observation and 3) replicates age-related sleep trends from previously published self-report studies, including changes in sleep duration and timing [40]. The devices have also been internally validated by the manufacturer.

Due to the nature of the data sampling, we cannot know whether individuals are travelling to a new destination or not, but that could be influential considering First Night Effect, and therefore we suggest it to be considered in future studies [16, 17]. Similarly, since the data is observational, the purpose of the trip – business versus pleasure – is unknown which could have an effect.

The effect of travel on sleep behaviour has not been studied for a cohort of this size before and most of the research has aimed to understand the effect of travel on sleep to optimize athletic performance or to apprehend fatigue among aircrews. [18, 27–34, 59]. Interestingly, one of these studies identified the same pattern as we do – travel was negatively correlated with sleep duration on weekdays among kite surfers (N=94) [59]. Generally, travel is believed to have deleterious effects on sleep, but our study reveals that travel provides a more complex impact on the sleep of travellers, providing respite to underslept individuals, while the deleterious effects are reserved for those who tend to be well-rested [18–24].

References

[1] Michael R Irwin. Why sleep is important for health: a psychoneuroimmunology perspective. *Annual review of psychology*, 66:143–172, 2015.

[2] Francesco P Cappuccio, Lanfranco D'Elia, Pasquale Strazzullo, and Michelle A Miller. Sleep duration and all-cause mortality: a systematic review and meta-analysis of prospective studies. *Sleep*, 33(5):585–592, 2010.

[3] Michael A Grandner, Nicholas J Jackson, Bilgay Izci-Balserak, Rebecca A Gallagher, Renee Murray-Bachmann, Natasha J Williams, Nirav P Patel, and Girardin Jean-Louis. Social and behavioral determinants of perceived insufficient sleep. *Frontiers in neurology*, 6:112, 2015.

[4] Faith S Luyster, Patrick J Strollo, Phyllis C Zee, and James K Walsh. Sleep: a health imperative. *Sleep*, 35(6):727–734, 2012.

[5] Francesco P Cappuccio, Frances M Taggart, Ngianga-Bakwin Kandala, Andrew Currie, Ed Peile, Saverio Stranges, and Michelle A Miller. Meta-analysis of short sleep duration and obesity in children and adults. *Sleep*, 31(5):619–626, 2008.

[6] Christopher Ryan King, Kristen L Knutson, Paul J Rathouz, Steve Sidney, Kiang Liu, and Diane S Lauderdale. Short sleep duration and incident coronary artery calcification. *Jama*, 300(24):2859–2866, 2008.

[7] Francesco P Cappuccio, Daniel Cooper, Lanfranco D'elia, Pasquale Strazzullo, and Michelle A Miller. Sleep duration predicts cardiovascular outcomes: a systematic review and meta-analysis of prospective studies. *European heart journal*, 32(12):1484–1492, 2011.

[8] June J Pilcher and Allen I Huffcutt. Effects of sleep deprivation on performance: a meta-analysis. *Sleep*, 19(4):318–326, 1996.

[9] Andrea N Goldstein and Matthew P Walker. The role of sleep in emotional brain function. *Annual review of clinical psychology*, 10:679–708, 2014.

[10] Rebecca A Bernert, Joanne S Kim, Naomi G Iwata, and Michael L Perlis. Sleep disturbances as an evidence-based suicide risk factor. *Current psychiatry reports*, 17(3):15, 2015.

[11] Earl S Ford, Timothy J Cunningham, and Janet B Croft. Trends in self-reported sleep duration among us adults from 1985 to 2012. *Sleep*, 38(5):829–832, 2015.

[12] Matthew P Walker. A societal sleep prescription. *Neuron*, 103(4):559–562, 2019.

[13] Charles A Czeisler. Perspective: casting light on sleep deficiency. *Nature*, 497(7450):S13–S13, 2013.

[14] The World Bank. Air transport, passengers carried, 2018.

[15] HW Agnew Jr, Wilse B Webb, and Robert L Williams. The first night effect: an eeg studyof sleep. *Psychophysiology*, 2(3):263–266, 1966.

[16] Masako Tamaki, Hiroshi Nittono, Mitsuo Hayashi, and Tadao Hori. Examination of the first-night effect during the sleep-onset period. *Sleep*, 28(2):195–202, 2005.

[17] Masako Tamaki, Ji Won Bang, Takeo Watanabe, and Yuka Sasaki. Night watch in one brain hemisphere during sleep associated with the first-night effect in humans. *Current biology*, 26(9):1190–1194, 2016.

[18] J Waterhouse, T Reilly, and B Edwards. The stress of travel. *Journal of sports sciences*, 22(10):946–966, 2004.

[19] Thomas Reilly, Jim Waterhouse, and Ben Edwards. Jet lag and air travel: implications for performance. *Clinics in sports medicine*, 24(2):367–380, 2005.

[20] Thomas Reilly, Jim Waterhouse, and Benjamin Edwards. Some chronobiological and physiological problems associated with long-distance journeys. *Travel Medicine and Infectious Disease*, 7(2):88–101, 2009.

[21] Michele Lastella, Gregory D Roach, and Charli Sargent. Travel fatigue and sleep/wake behaviors of professional soccer players during international competition. *Sleep health*, 5(2):141–147, 2019.

[22] Jim Waterhouse, Thomas Reilly, Greg Atkinson, and Ben Edwards. Jet lag: trends and coping strategies. *The lancet*, 369(9567):1117–1129, 2007.

[23] Venkataramanujan Srinivasan, D Warren Spence, Seithikurippu R Pandi-Perumal, Ilya Trakht, and Daniel P Cardinali. Jet lag: therapeutic use of melatonin and possible application of melatonin analogs. *Travel Medicine and Infectious Disease*, 6(1-2):17–28, 2008.

[24] Timothy I Morgenthaler, Teofilo Lee-Chiong, Cathy Alessi, Leah Friedman, R Nisha Aurora, Brian Boehlecke, Terry Brown, Andrew L Chesson Jr, Vishesh Kapur, Rama Maganti, et al. Practice parameters for the clinical evaluation and treatment of circadian rhythm sleep disorders. *Sleep*, 30(11):1445–1459, 2007.

[25] Marc Wittmann, Jenny Dinich, Martha Merrow, and Till Roenneberg. Social jetlag: misalignment of biological and social time. *Chronobiology international*, 23(1-2):497–509, 2006.

[26] Till Roenneberg, Karla V Allebrandt, Martha Merrow, and Céline Vetter. Social jetlag and obesity. *Current Biology*, 22(10):939–943, 2012.

[27] Mathieu Nédélec, Shona Halson, Abd-Elbasset Abaidia, Said Ahmaidi, and Gregory Dupont. Stress, sleep and recovery in elite soccer: a critical review of the literature. *Sports Medicine*, 45(10):1387–1400, 2015.

[28] Peter Fowler, Rob Duffield, and Joanna Vaile. Effects of domestic air travel on technical and tactical performance and recovery in soccer. *International Journal of Sports Physiology and Performance*, 9(3):378–386, 2014.

[29] Sarah Forbes-Robertson, Edward Dudley, Pankaj Vadgama, Christian Cook, Scott Drawer, and Liam Kilduff. Circadian disruption and remedial interventions. *Sports Medicine*, 42(3):185–208, 2012.

[30] Thomas Reilly, Greg Atkinson, Ben Edwards, Jim Waterhouse, Torbjörn Åkerstedt, Damien Davenne, Björn Lemmer, and Anna Wirz-Justice. Coping with jet-lag: a position statement for the european college of sport science. *European Journal of Sport Science*, 7(1):1–7, 2007.

[31] Peter M Fowler, Wade Knez, Stephen Crowcroft, Amy E Mendham, Joanna Miller, Charli Sargent, Shona Halson, and Rob Duffield. Greater effect of east versus west travel on jet lag, sleep, and team sport performance. *Medicine & Science in Sports & Exercise*, 49(12):2548–2561, 2017.

[32] Samira Bourgeois-Bougrine, Philippe Carbon, Charlotte Gounelle, Regis Mollard, and Alex Coblentz. Perceived fatigue for short-and long-haul flights: a survey of 739 airline pilots. *Aviation, space, and environmental medicine*, 74(10):1072–1077, 2003.

[33] Mikael Sallinen, Maria Sihvola, Sampsa Puttonen, Kimmo Ketola, Antti Tuori, Mikko Härmä, Göran Kecklund, and Torbjörn Åkerstedt. Sleep, alertness and alertness management among commercial airline pilots on short-haul and long-haul flights. *Accident Analysis & Prevention*, 98:320–329, 2017.

[34] Cátia Reis, Catarina Mestre, Helena Canhão, David Gradwell, and Teresa Paiva. Sleep complaints and fatigue of airline pilots. *Sleep Science*, 9(2):73–77, 2016.

[35] Peter Fowler, Rob Duffield, and Joanna Vaile. Effects of simulated domestic and international air travel on sleep, performance, and recovery for team sports. *Scandinavian journal of medicine & science in sports*, 25(3):441–451, 2015.

[36] Louise K Richmond, Brian Dawson, Glenn Stewart, Stuart Cormack, David R Hillman, and Peter R Eastwood. The effect of interstate travel on the sleep patterns and performance of elite australian rules footballers. *Journal of Science and Medicine in Sport*, 10(4):252–258, 2007.

[37] Teneale Alyce McGuckin, Wade Heath Sinclair, Rebecca Maree Sealey, and Paul Bowman. The effects of air travel on performance measures of elite australian rugby league players. *European journal of sport science*, 14(sup1):S116–S122, 2014.

[38] Josephine Arendt. Managing jet lag: Some of the problems and possible new solutions. *Sleep medicine reviews*, 13(4):249–256, 2009.

[39] T Reilly, G Atkinson, and J Waterhouse. Travel fatigue and jet-lag. *Journal of sports sciences*, 15(3):365–369, 1997.

[40] Sigga Svala Jonasdottir, Kelton Minor, and Sune Lehmann. Gender differences in nighttime sleep patterns and variability across the adult lifespan: A global-scale wearables study. *Sleep*, 2020.

[41] Jennifer Girschik, Lin Fritschi, Jane Heyworth, and Flavie Waters. Validation of self-reported sleep against actigraphy. *Journal of epidemiology*, 22(5):462–468, 2012.

[42] Sonia Ancoli-Israel, Roger Cole, Cathy Alessi, Mark Chambers, William Moorcroft, and Charles P Pollak. The role of actigraphy in the study of sleep and circadian rhythms. *Sleep*, 26(3):342–392, 2003.

[43] Tim Althoff, Eric Horvitz, Ryen W White, and Jamie Zeitzer. Harnessing the web for population-scale physiological sensing: A case study of sleep and performance. In *Proceedings of the 26th international conference on World Wide Web*, pages 113–122. International World Wide Web Conferences Steering Committee, 2017.

[44] Ulf Aslak and Laura Alessandretti. Infostop: Scalable stop-location detection in multi-user mobility data. *arXiv preprint arXiv:2003.14370*, 2020.

[45] Julian J Faraway. *Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models*. Chapman and Hall/CRC, 2016.

[46] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2018.

[47] Alexandra Kuznetsova, Per B Brockhoff, and Rune Haubo Bojesen Christensen. lmertest package: tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 2017.

[48] Steven G Luke. Evaluating significance in linear mixed-effects models in r. *Behavior research methods*, 49(4):1494–1502, 2017.

[49] Alexander A Borbély. A two process model of sleep regulation. *Hum neurobiol*, 1(3):195–204, 1982.

[50] Serge Daan, DG Beersma, and Alexander A Borbély. Timing of human sleep: recovery process gated by a circadian pacemaker. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 246(2):R161–R183, 1984.

[51] Till Roenneberg, Tim Kuehnle, Myriam Juda, Thomas Kantermann, Karla Allebrandt, Marijke Gordijn, and Martha Merrow. Epidemiology of the human circadian clock. *Sleep medicine reviews*, 11(6):429–438, 2007.

[52] Gianina Luca, José Haba Rubio, Daniela Andries, Nadia Tobback, Peter Vollenweider, Gérard Waeber, Pedro Marques Vidal, Martin Preisig, Raphaël Heinzer, and Mehdi Tafti. Age and gender variations of sleep in subjects without sleep disorders. *Annals of medicine*, 47(6):482–491, 2015.

[53] L Kuula, M Gradisar, K Martinmäki, C Richardson, D Bonnar, K Bartel, C Lang, L Leinonen, and AK Pesonen. Using big data to explore worldwide trends in objective sleep in the transition to adulthood. *Sleep Medicine*, 62:69–76, 2019.

[54] Ju Lynn Ong, Jesisca Tandi, Amiya Patanaik, June C Lo, and Michael WL Chee. Large-scale data from wearables reveal regional disparities in sleep patterns that persist across age and sex. *Scientific reports*, 9(1):3415, 2019.

[55] Olivia J Walch, Amy Cochran, and Daniel B Forger. A global quantification of "normal" sleep schedules using smartphone data. *Science advances*, 2(5):e1501705, 2016.

[56] June C Lo, Ruth LF Leong, Kep-Kee Loh, Derk-Jan Dijk, and Michael WL Chee. Young adults' sleep duration on work days: differences between east and west. *Frontiers in neurology*, 5:81, 2014.

[57] Soultana Macridis, Nora Johnston, Steven Johnson, and Jeff K Vallance. Consumer physical activity tracking device ownership and use among a population-based sample of adults. *PloS one*, 13(1):e0189298, 2018.

[58] Christopher M Depner, Philip C Cheng, Jaime K Devine, Seema Khosla, Massimiliano de Zambotti, Rébecca Robillard, Andrew Vakulin, and Sean PA Drummond. Wearable technologies for developing sleep and circadian biomarkers: a summary of workshop discussions. *sleep*, 43(2):zsz254, 2020.

[59] Maria-Raquel G Silva, António Pascoal, Hugo-Henrique Silva, and Teresa Paiva. Assessing sleep, travelling habits and jet lag in kite surfers according to competition level. *Biological Rhythm Research*, 47(5):677–689, 2016.

[60] WHO Consultation. Obesi'i'y: Preventing and managing the global epidemic. 2000.

[61] Expert Consultation Who. Appropriate body-mass index for asian populations and its implications for policy and intervention strategies. *Lancet (London, England)*, 363(9403):157, 2004.

Supplementary Information

Data Availability Raw data are not publicly available to preserve users' privacy (according to the Privacy Policy of the wearable devices). Aggregated and anonymized data supporting the findings of this study are available from the corresponding authors upon request. Figure 1A contains raw data-points from the data-set.

Code Availability The code used to generate the results of this paper is available for download on github [link in published paper].

Data Pre-processing

The raw data consists 1-minute epochs of sleep activity which are aggregated into nights with sleep onset, offset, duration and wake-time after sleep onset (WASO). For each night a user can wake up multiple times but each awakening can only last for 60 minutes or less. In order to obtain nighttime sleep (exclude naps) and remove outliers we apply the following standard filters for sleep duration: $3 \leq$ duration ≤ 13 introduced by Roenneberg et.al [26]. Next we look at the distribution of sleep onset and offset separately on weekdays and weekends (Figure 1) and set filters to be 3 standard deviation away from the mean, or

- 20:24 $\leq$ onset weekends $\leq$ 04:52 and 03:59 $\leq$ offset weekends $\leq$12:52
- 20:28 $\leq$ onset weekdays $\leq$ 03:59 and 03:21 $\leq$ offset weekdays $\leq$ 11:25

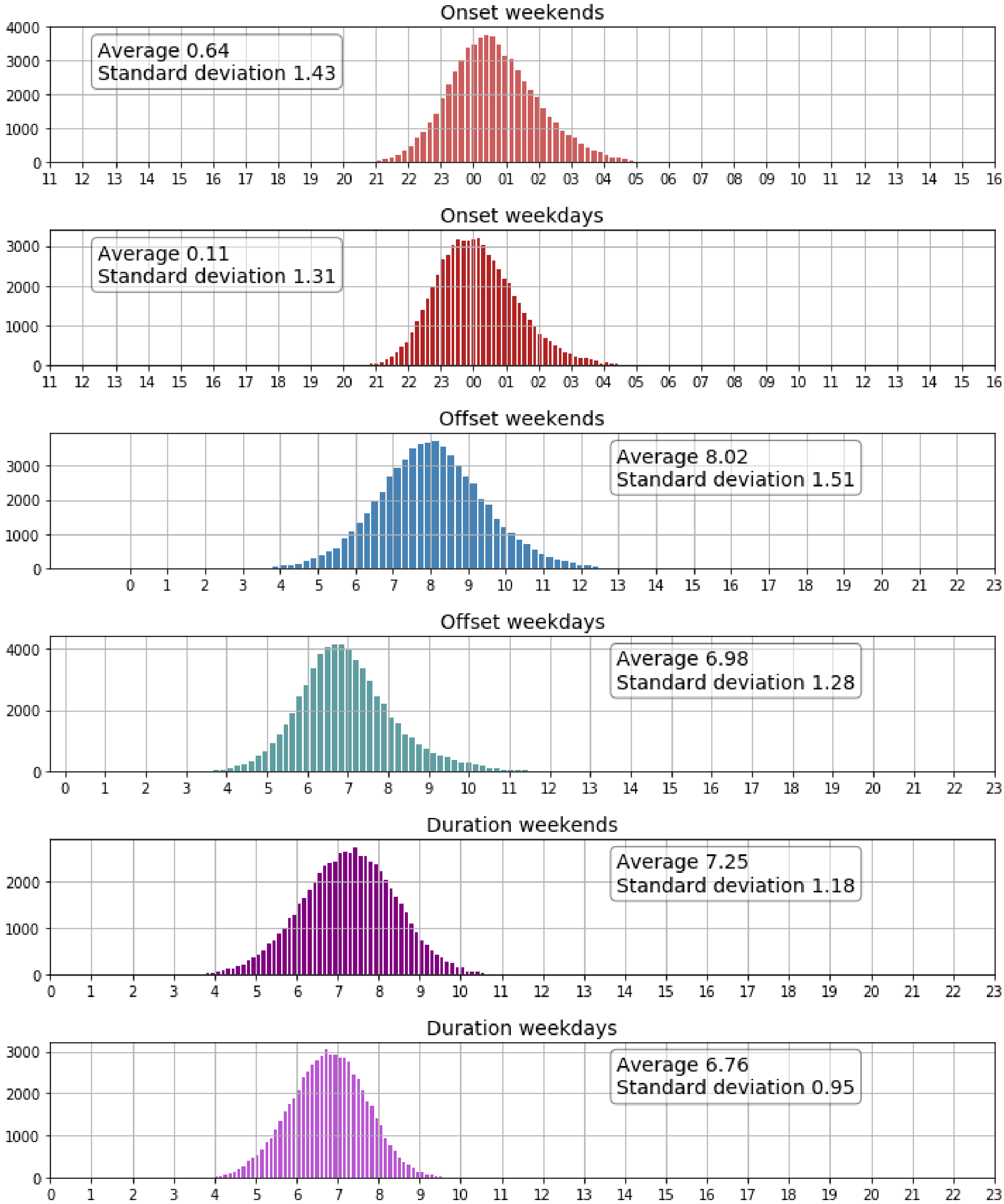


Figure 1: Distribution of sleep onset, offset duration plotted separately for weekdays and weekends

Filtering & Inclusion Criteria

To motivate our choice for the minimum number of nights required per user, we examine the development of the standard deviation for sleep duration by the number of days recorded, both at home and away from home (Figure 2). The standard deviation seems to stabilize around 10 recorded nights, both at home and away from home. That threshold is reasonable for at nights at home but would eliminate majority of our data (more than 90 %) if applied to travel-nights. Thus, we decided to require users of two recorded nights away from home by *day type*. One should pay attention to the fact that users can be included for the analysis just on either weekdays or weekends – not necessarily both day types.

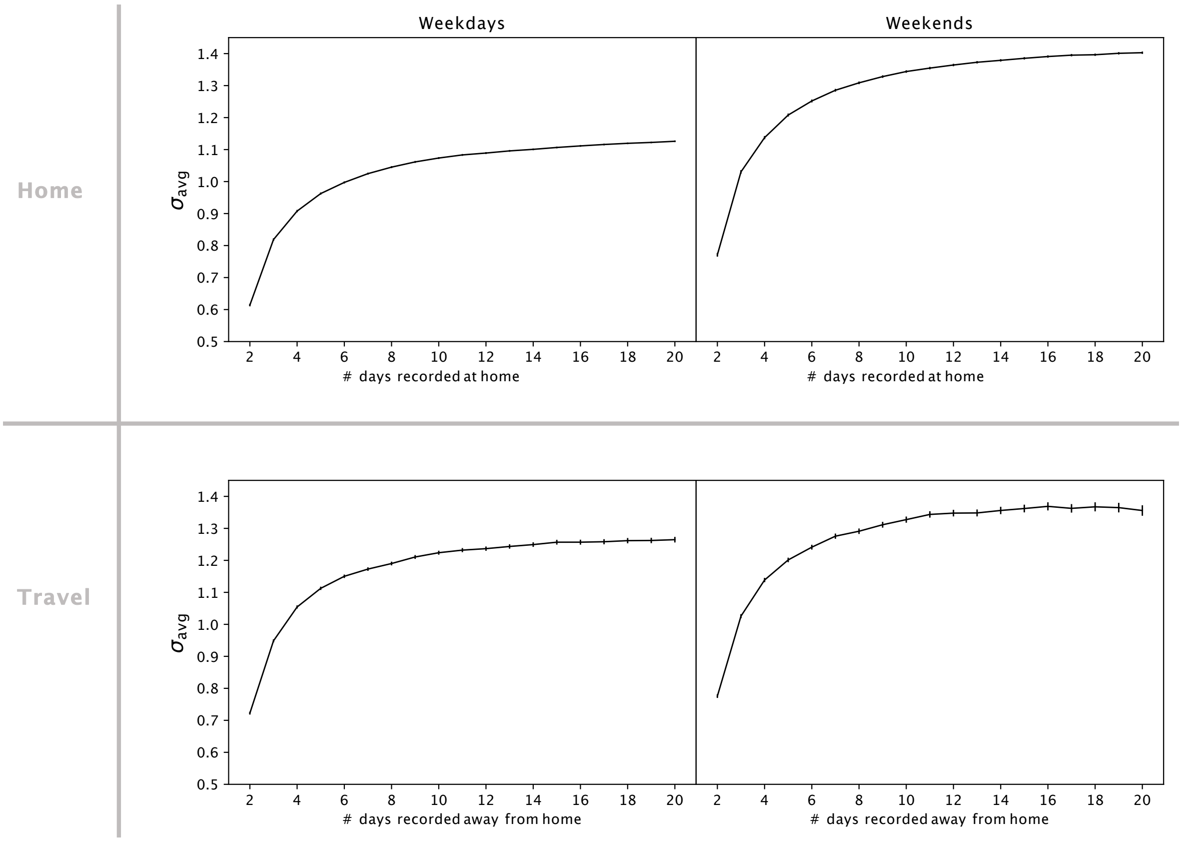


Figure 2: Development of the standard deviation of sleep duration aggregated by number of nights recorded at home and away from home

In most of our analysis we use the median sleep duration to quantify typical at home behaviour, consequently we also examine how the distributions for the standard error of the median (SEMe) develops as the inclusion criteria changes (Figure 3). Naturally, the distributions become tighter, the average and standard deviation decrease in magnitude as the number of days required per user is increased. We chose to require users of 10 recorded nights at home on by day type.

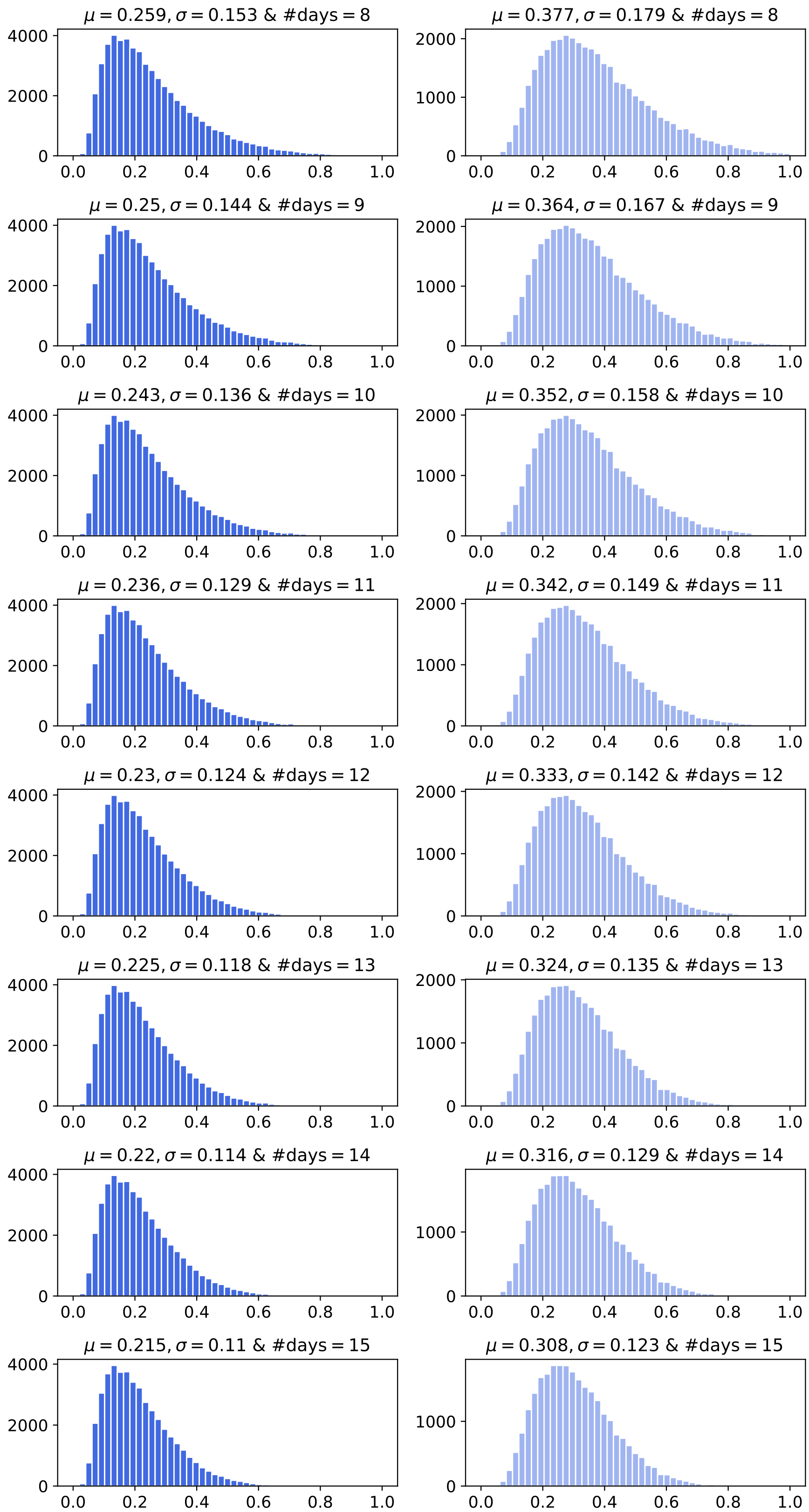


Figure 3: Distributions for the standard error of the median (SEMe) by day type while changing the number home-nights required per user (weekdays left column and weekends right column)

Data Coverage & Demographics

Sleep behaviour is dependent on internal and external processes and differentiates by demographic variables such as gender, age, cultural context and day type. To explore those effects we use individual-level covariates; gender (female/male), generations (Baby Boomers born 1946-64, Gen X born 1965-80, Millennial's born 1981-96 and Gen Z born 1997 or later) and BMI categories (underweight/normal weight/overweight/obese) which were labelled according to the World Health Organization classification [60, 61]. There are large disparities in sleep patterns across cultures, especially the contrast between Eastern (Asia) and Western (Europe and North America) regions. Studies have shown that sleep duration is shorter and bed times later among people residing in the East than those living in the West [53–56]. Thus, we use region of residence (also called `East/West`) as a covariate where East represents residents in Asia and West for those living N-America and Europe. All plots and models are implemented separately for weekdays and weekends because of likely differences in the social structure over the course of the week. Since we do not directly observe schedules, we assume the likelihood of work days is highest on weekdays and work-free days is highest on weekends [25]. Trips are categorized by distance; less than 1000 km, 1000-2500 km and more than 2500 km travelled.

Figure 4 visualises 1) the distribution of the number days users have recorded at home and away from home, 2) how users distribute by gender, BMI categories and generations, 3) lists out the ten largest geographic regions and 4) shows how far away from home travel-nights typically are. Approximately 1/3 of the sample is women and 2/3 men. Most users are either normal weight or overweight and from generation x or millennial's. Users distribute similarly by age and BMI for both genders. Most travel-nights are 1000 km or less away from home. There are dominant geographic regions in the sample, where more than 60 % of users live in 5 countries. We do explore and control for the effect of *all* demographic variables in our analysis.

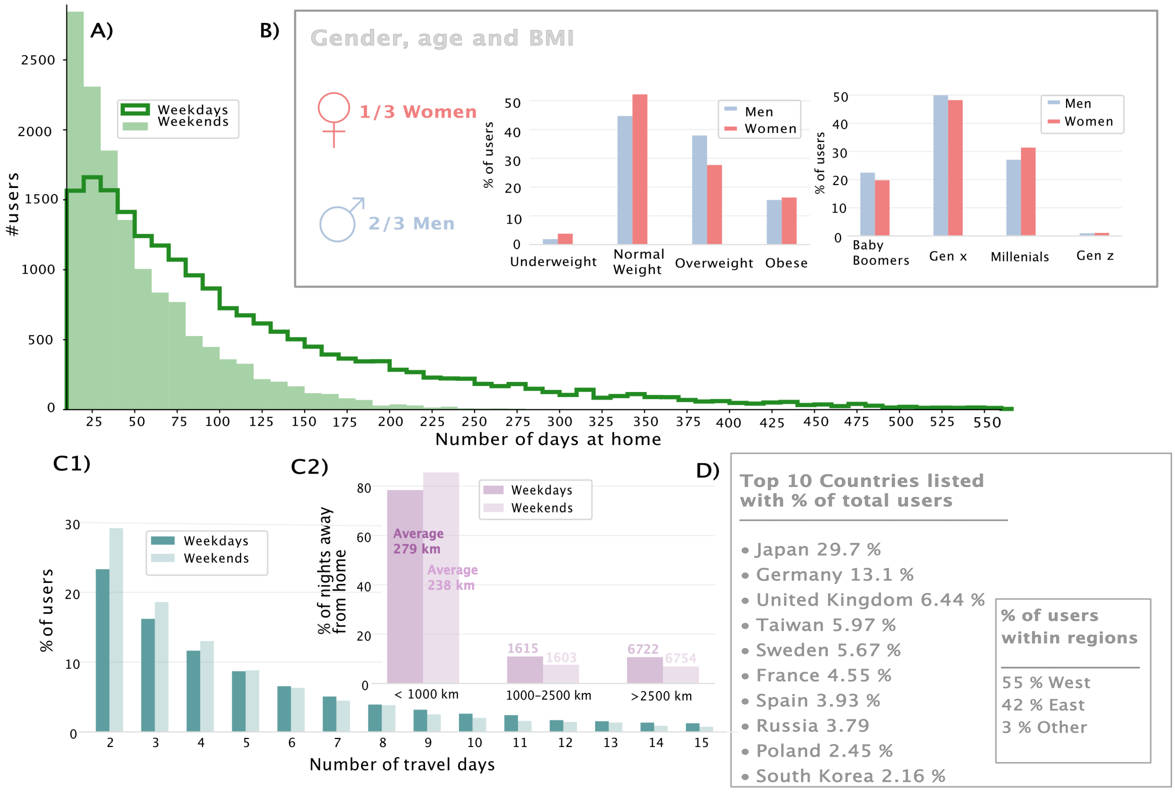


Figure 4: A) Displays the distribution of number of days users have recorded at home on weekends and weekdays. **B)** Illustrates the ratio of female and male users and how they distribute by BMI categories and generations. **C1)** Shows the percentage of users by number of travel days, separately on weekdays and weekends. **C2)** Displays how travel-nights distribute by distance categories and lists the average distance within each category, separately on weekends and weekdays. **D**) Lists the top ten countries with most residence and the percentage of users living there, as well as the percentage of users living within the three regions (East/West/Other)

Down-sampling nights at home

As mentioned in the manuscript - one of the limitations of this study is the disproportionate number of nights recorded away from home in comparison to nights at home (6 % of weekdays and 9.3 % of weekends are travel-nights). One might consider that the change in sleep behaviour away from home could be happening incidentally – meaning that if we randomly choose the same amount of nights at home as number of nights recorded away from home, then the sample distributions for Δ_{home} and Δ_{travel} would look more alike.

To contest to that presumption, we perform down-sampling such that we randomly select nights at home to be equal to the number of nights recorded away from home (for each user) and compare the sample distributions (both visually and by percentiles) for $\Delta_{home\ DS}$ and Δ_{home}. The process is described step-by-step;

- Repeat 50 times;
  - For each user we randomly choose N_{travel} nights recorded at home
  - For those randomly drawn nights, we estimate Δ_{home} and store it for each user
- Estimate $\Delta_{home\ DS}$ for each user from the 50 trials
- Estimate the quartiles for the sample distribution of $\Delta_{home\ DS}$

Results are listed in Tables 2 and 3 and distributions also visualised on Figure 5. The distribution for $\Delta_{home\ DS}$ is actually narrower than for the full sample. That can be rationalized by the fact that 70 % of users have 5 or less days recorded away from home but when we examined the development of the standard deviation by number of data-points (see section *Filtering & Inclusion Criteria* in SI) – the standard deviation increases and is not stabilized until there are about 10 recorded nights. The distribution for $\Delta_{home\ DS}$ moves further away from the distribution Δ_{travel} when down-sampled.

Iteration	**1**	**2**	**3**	**4**	**5**	**Full sample – Home**	**Full sample – Travel**
Minimum	-0.565	-0.588	-0.596	-0.619	-0.617	-1.39	-5.25
Lower quartile	-0.0532	-0.0515	-0.0533	-0.0524	-0.0534	-0.110	-0.417
Median	0	0	0	0	0	-0.0140	0.239
Upper quartile	0.0342	0.0340	0.0340	0.0363	0.0346	0.086	0.933
Maximum	0.726	0.754	0.711	0.735	0.766	1.16	5.98

Table 2: Sample quartiles of $\Delta_{home\ DS}$ [hours] home-nights are randomly selected and equal to the number of travel-nights on weekdays

Iteration	**1**	**2**	**3**	**4**	**5**	**Full sample – Home**	**Full sample – Travel**
Minimum	-0.692	-0.841	-0.680	-0.752	-0.746	-1.39	-5.25
Lower quartile	-0.0763	-0.0783	-0.0777	-0.0785	-0.0789	-0.110	-0.417
Median	0	0	0	0	0	-0.0140	0.239
Upper quartile	0.0114	0.0110	0.0117	0.0112	0.0134	0.086	0.933
Maximum	0.712	0.604	0.586	0.6454	0.6322	1.16	5.98

Table 3: Sample quartiles of $\Delta_{home\ DS}$ [hours] home-nights are randomly selected and equal to the number of travel-nights on weekends

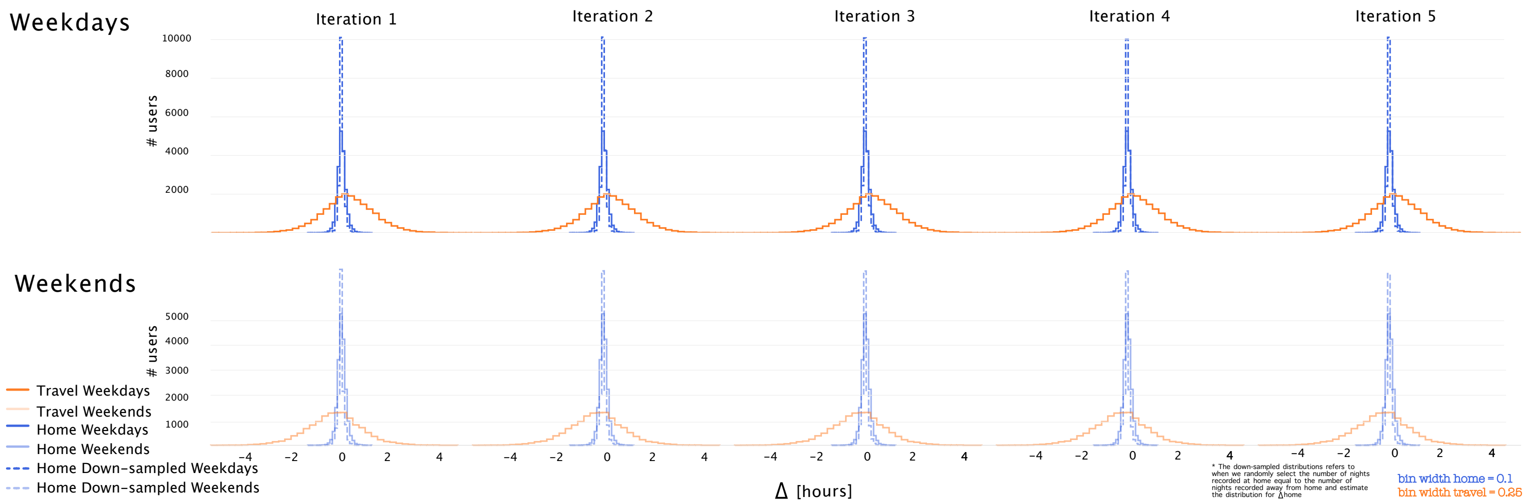


Figure 5: Distributions of $\Delta_{home\ DS}$ on weekdays and weekends (from the five iterations described above) with Δ_{home} and Δ_{travel}

The baseline effect at home

Sleep groups	4.5	5.0	7.0	9.0	9.5
% users where μ > M *	95.2 %	86 %	40 %	19 %	7 %
% users where μ ≤ M *	4.8 %	14 %	60 %	81 %	93 %
μ_{skew} *	1.02	0.67	-0.12	-0.73	-0.66

*μ denotes average, M denotes median

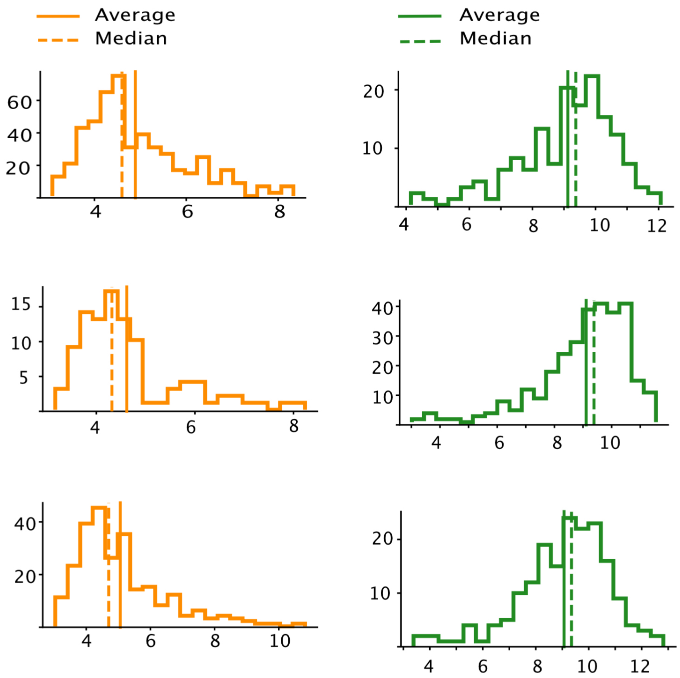


Figure 6: Here we illustrate how the asymmetry of an individual's distribution emerges due to homeostasis. In the table the majority of individuals who regularly have shorter nighttime sleep at home (4.5 or 5.0 hours) have a median larger than the mean and a positively skewed distribution – indicating heavier right tail. The opposite can be observed for individuals typically obtaining longer nighttime sleep – where majority of the users have an averages smaller than the median and a negative skew suggesting disproportional tendency for shorter nights. The distributions on the bottom of the figure are representative for six randomly selected users, 3 of which have low sleep duration (orange color) and 3 who have high sleep duration (green color).

Mixed effects model

To analyse the data-set with mixed effects model we implement a model with a three-way interaction term between home (True/False), every demographic variable and median duration (centered around the mean). Measurements are nested within user (random effect) and the model is defined in equation below.

$$\begin{aligned}
Y_i \quad = \quad & \mu + \alpha(\text{duration_center}_i) + \beta(\text{home}_i) + \delta(\text{bmi_cat}_i) + \epsilon(\text{east_west}_i) + \\
& \zeta(\text{gender}_i) + \eta(\text{generation}_i) + \theta(\text{home}_i \times \text{duration_center}_i) + \iota(\text{home}_i \times \text{bmi_cat}_i) + \\
& \kappa(\text{home}_i \times \text{east_west}_i) + \lambda(\text{home}_i \times \text{gender}_i) + \nu(\text{home}_i \times \text{generation}_i) + \\
& \xi(\text{duration_center}_i \times \text{bmi_cat}_i) + \pi(\text{duration_center}_i \times \text{east_west}_i) + \\
& \rho(\text{duration_center}_i \times \text{gender}_i) + \sigma(\text{duration_center}_i \times \text{generation}_i) + \\
& \tau(\text{duration_center}_i \times \text{home}_i \times \text{bmi_cat}_i) + \upsilon(\text{duration_center}_i \times \text{home}_i \times \text{east_west}_i) + \\
& \phi(\text{duration_center}_i \times \text{home}_i \times \text{gender}_i) + \chi(\text{duration_center}_i \times \text{home}_i \times \text{generation}_i) + \\
& + y(\text{user}_i) + \epsilon_i \text{ where } i = 1, \ldots 773\,132 \; or \; i = 1, \ldots 2\,386\,370 \\
& \text{ Furthermore } y(\text{user}_i) \sim N(0, \sigma_w^2), \text{ and } \epsilon_i \sim N(0, \sigma^2)
\end{aligned}$$

Estimates of fixed effects

Fixed effect	Estimate	Std. Error	P-value
(Intercept)	-3.092e-03	2.847e-03	0.28
dur_C	-7.357e-02	3.298e-03	<2e-16 ***
homefalse	5.589e-01	7.955e-03	<2e-16 ***
east_westeast	-2.004e-02	2.748e-03	3.25e-13 ***
genderFEMALE	2.367e-03	2.706e-03	0.381774
bmi_cat2	-2.639e-03	2.646e-03	0.318754
bmi_cat3	-8.979e-03	3.554e-03	0.011546 *
generationbaby boomers	-1.982e-02	2.915e-03	1.07e-11 ***
generationmillenials	1.231e-02	2.930e-03	2.66e-05 ***
dur_C:homefalse	-4.070e-01	7.447e-03	<2e-16 ***
homefalse:east_westeast	-5.943e-01	7.931e-03	<2e-16 ***
homefalse:genderFEMALE	1.121e-01	8.115e-03	<2e-16 ***
homefalse:bmi_cat2	-3.994e-02	7.602e-03	1.49e-07 ***
homefalse:bmi_cat3	-1.399e-01	1.022e-02	<2e-16 ***
homefalse:generationbaby boomers	-7.463e-02	8.321e-03	<2e-16 ***
homefalse:generationmillenials	-6.835e-02	8.478e-03	7.53e-16 ***
dur_C:east_westeast	-1.452e-02	2.982e-03	1.14e-06 ***
dur_C:genderFEMALE	5.005e-03	3.008e-03	0.096 .
dur_C:bmi_cat2	-1.738e-03	2.885e-03	0.55
dur_C:bmi_cat3	5.294e-03	3.798e-03	0.16
dur_C:generationbaby boomers	1.156e-03	3.101e-03	0.71
dur_C:generationmillenials	-1.167e-02	3.326e-03	0.00045 ***
dur_C:homefalse:genderFEMALE	-1.940e-02	8.768e-03	0.027 *
dur_C:homefalse:bmi_cat2	3.116e-02	8.327e-03	0.00018 ***
dur_C:homefalse:bmi_cat3	7.372e-02	1.100e-02	2.06e-11 ***
dur_C:homefalse:generationbaby boomers	4.440e-02	8.991e-03	7.87e-07 ***
dur_C:homefalse:generationmillenials	-7.136e-02	9.698e-03	1.87e-13 ***

Table 4: Estimates of fixed effects from mixed effects models for home-nights and travel-nights on weekdays

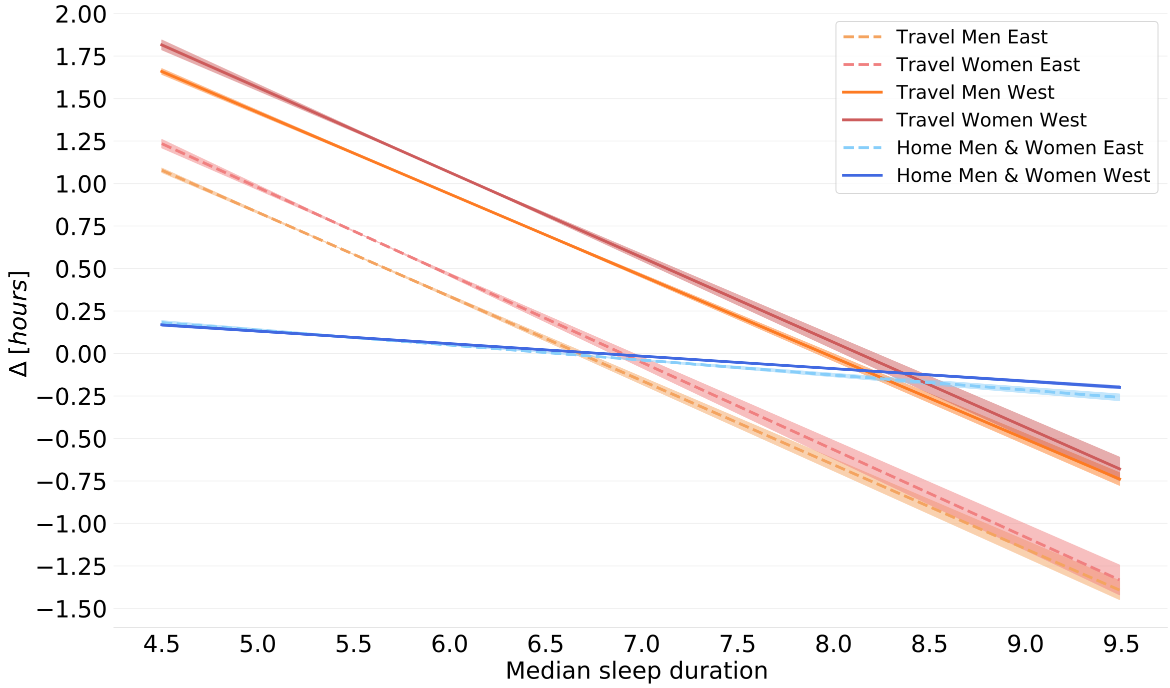


Figure 7: Illustration for the mixed effect model for the most important fixed effects in terms of significance and effect size on weekdays

Fixed effect	Estimate	Std. Error	P-value
(Intercept)	-4.158e-02	4.732e-03	<2e-16 ***
dur_C	-9.015e-02	4.732e-03	<2e-16 ***
homefalse	1.193e-02	1.416e-02	0.40*
east_westeast	1.623e-03	4.554e-03	0.72
genderFEMALE	-5.180e-03	4.343e-03	0.23
bmi_cat2	-6.223e-03	4.301e-03	0.15
bmi_cat3	-7.051e-03	5.850e-03	0.23
generationbaby boomers	-2.446e-02	4.684e-03	1.80e-07 ***
generationmillenials	1.375e-02	4.920e-03	0.0052 **
dur_C:homefalse	-3.355e-01	1.409e-02	<2e-16 ***
homefalse:east_westeast	-4.659e-01	1.426e-02	<2e-16 ***
homefalse:genderFEMALE	-2.479e-02	1.353e-02	0.067 .
homefalse:bmi_cat2	-7.138e-02	1.330e-02	7.93e-08 ***
homefalse:bmi_cat3	-1.376e-01	1.770e-02	7.58e-15 ***
homefalse:generationbaby boomers	-1.897e-03	1.476e-02	0.90
homefalse:generationmillenials	7.500e-04	1.474e-02	0.96
dur_C:east_westeast	-5.679e-03	4.489e-03	0.21
dur_C:genderFEMALE	2.864e-03	4.334e-03	0.51
dur_C:bmi_cat2	6.384e-04	4.207e-03	0.88
dur_C:bmi_cat3	5.256e-03	5.587e-03	0.35
dur_C:generationbaby boomers	4.635e-03	4.525e-03	0.31
dur_C:generationmillenials	-1.546e-02	4.992e-03	0.0020 **
dur_C:homefalse:east_westeast	-8.511e-02	1.388e-02	8.74e-10 ***
dur_C:homefalse:genderFEMALE	-5.676e-02	1.329e-02	1.94e-05 ***
dur_C:homefalse:bmi_cat2	2.853e-02	1.295e-02	0.028 *
dur_C:homefalse:bmi_cat3	5.596e-02	1.695e-02	0.00096 ***
dur_C:homefalse:generationbaby boomers	3.018e-02	1.426e-02	0.034 *
dur_C:homefalse:generationmillenials	-6.285e-02	1.477e-02	2.10e-05 ***

Table 5: Estimates of fixed effects from mixed effects models for home-nights and travel-nights on weekends

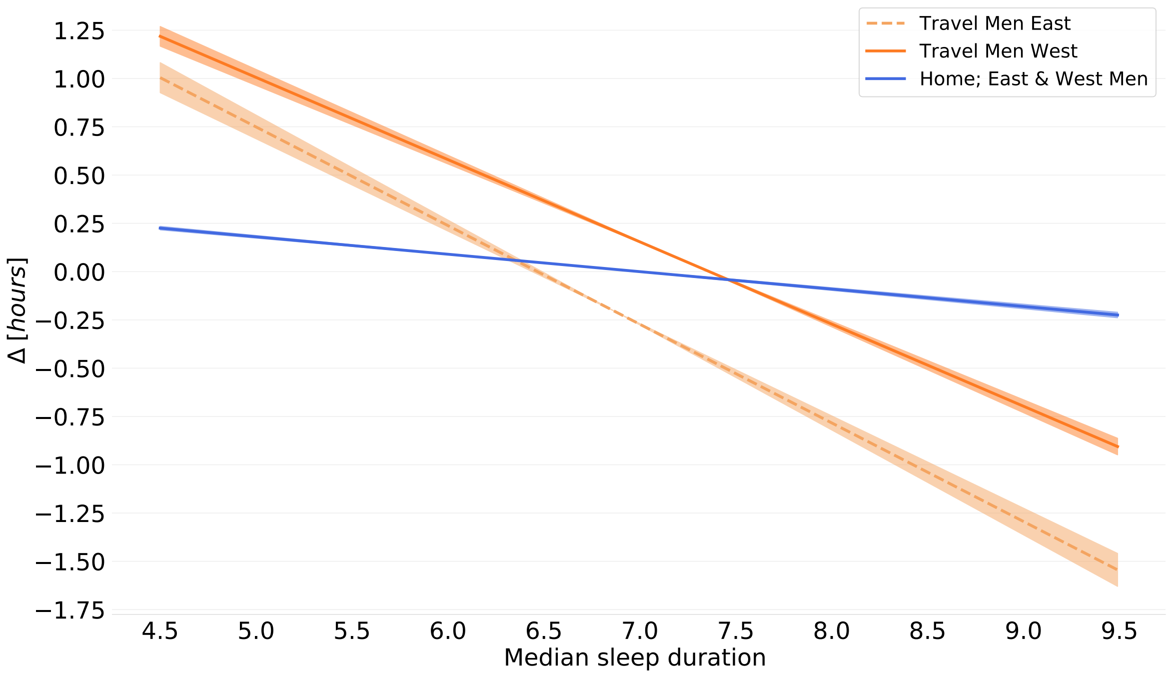


Figure 8: Illustration for the mixed effect model for the most important fixed effects in terms of significance and effect size on weekends

Mixed effects model: Exploring the effect of distance

To explore the effect of `distance` on the relationship between Δ_{travel} and typical sleep duration we implement a mixed effect model only for days away from, defined in equation below.

$$\begin{aligned} Y_i \quad = \quad & \mu + \alpha(\text{duration\_center}_i) + \beta(\text{distance\_category}_i) + \delta(\text{bmi\_cat}_i) + \epsilon(\text{east\_west}_i) + \\ & \zeta(\text{gender}_i) + \eta(\text{generation}_i) + \theta(\text{distance\_category}_i \times \text{duration\_center}_i) + \\ & \iota(\text{distance\_category}_i \times \text{bmi\_cat}_i) + \kappa(\text{distance\_category}_i \times \text{east\_west}_i) + \\ & \lambda(\text{distance\_category}_i \times \text{gender}_i) + \nu(\text{distance\_category}_i \times \text{generation}_i) + \\ & y(\text{user}_i) + \epsilon_i \text{ where } i = 1, \ldots 73\,265 \ or \ i = 1, \ldots 144\,582 \\ & \text{ Furthermore } y(\text{user}_i) \sim N(0, \sigma_w^2), \text{ and } \epsilon_i \sim N(0, \sigma^2) \end{aligned}$$

Estimates of fixed effects

Fixed effects	Estimate	Std. Error	P-value
(Intercept)	5.922e-01	1.675e-02	<2e-16 ***
dur_C	-4.867e-01	8.127e-03	<2e-16 ***
distance_category2500	5.271e-02	3.354e-02	0.12
distance_category2500++	-1.705e-01	3.555e-02	1.61e-06 ***
east_westeast	-5.926e-01	1.642e-02	<2e-16 ***
genderFEMALE	1.170e-01	1.619e-02	5.13e-13 ***
bmi_cat2	-7.324e-02	1.591e-02	4.16e-06 ***
bmi_cat3	-1.751e-01	2.099e-02	<2e-16 ***
generationbaby boomers	-1.050e-01	1.774e-02	3.36e-09 ***
generationmillenials	-9.293e-02	1.714e-02	5.98e-08 ***
distance_category2500:east_westeast	-1.823e-01	3.258e-02	2.19e-08 ***
distance_category2500++:east_westeast	-1.127e-01	3.361e-02	0.0008 ***
distance_category2500:genderFEMALE	-1.116e-01	3.269e-02	0.00064 ***
distance_category2500++:genderFEMALE	-5.883e-03	3.607e-02	0.87
distance_category2500:bmi_cat2	-1.968e-02	3.318e-02	0.55
distance_category2500++:bmi_cat2	9.754e-02	3.504e-02	0.0054 **
distance_category2500:bmi_cat3	3.896e-03	4.430e-02	0.93
distance_category2500++:bmi_cat3	1.698e-01	4.863e-02	0.00048 ***
distance_category2500:generationbaby boomers	-1.009e-02	3.627e-02	0.78
distance_category2500++:generationbaby boomers	2.282e-02	3.771e-02	0.55
distance_category2500:generationmillenials	6.709e-02	3.601e-02	0.062
distance_category2500++:generationmillenials	1.041e-01	3.955e-02	0.0085 **

Table 6: Estimates of fixed effects from mixed effects models for travel-nights on weekdays

Fixed effects	**Estimate**	**Std. Error**	**P-value**
(Intercept)	-2.887e-02	1.610e-02	0.073 .
dur_C	-4.880e-01	8.952e-03	<2e-16 ***
distance_category2500	-5.428e-02	4.038e-02	0.179
distance_category2500++	-2.481e-01	4.506e-02	3.67e-08 ***
east_westeast	-4.587e-01	2.008e-02	<2e-16 ***
bmi_cat2	-7.956e-02	1.922e-02	3.51e-05 ***
bmi_cat3	-1.602e-01	2.560e-02	4.09e-10 ***
distance_category2500:east_westeast	-2.180e-01	5.098e-02	1.90e-05 ***
distance_category2500++:east_westeast	-5.567e-02	5.489e-02	0.310
distance_category2500:bmi_cat2	6.107e-02	5.313e-02	0.250
distance_category2500++:bmi_cat2	5.953e-02	5.724e-02	0.298
distance_category2500:bmi_cat3	-2.831e-02	7.159e-02	0.692
distance_category2500++:bmi_cat3	2.508e-01	8.116e-02	0.002 **

Table 7: Estimates of fixed effects from mixed effects models for travel-nights on weekends

Robustness

$$
\begin{aligned}
Y_i \;=\; & \mu + \alpha(\text{duration\_center}_i) + \beta(\text{home}_i) + \delta(\text{bmi\_cat}_i) + \epsilon(\text{east\_west}_i) + \\
& \zeta(\text{gender}_i) + \eta(\text{generation}_i) + \theta(\text{home}_i \times \text{duration\_center}_i) + \\
& \iota(\text{home}_i \times \text{bmi\_cat}_i) + \kappa(\text{home}_i \times \text{east\_west}_i) + \\
& \lambda(\text{home}_i \times \text{gender}_i) + \nu(\text{home}_i \times \text{generation}_i) + \\
& y(\text{user}_i) + \epsilon_i \text{ where } i = 1, \ldots 773\,132 \; or \; i = 1, \ldots 2\,386\,370 \\
& \text{ Furthermore } y(\text{user}_i) \sim N(0, \sigma_w^2), \text{ and } \epsilon_i \sim N(0, \sigma^2)
\end{aligned}
$$

Fixed effect	**travel days $\geq$ 2**	**travel days $\geq$ 4**	**travel days $\geq$ 6**	**travel days $\geq$ 8**	**travel days $\geq$ 10**	**travel days $\geq$ 12**
	Estim. $\pm$ SEM	**Estim. $\pm$ SEM**	**Estim. $\pm$ SEM**	**Estim. $\pm$ SEM**	**Estim. $\pm$ SEM**	**Estim. $\pm$ SEM**
Intercept	-0.0013 $\pm$ 0.003	-0.0033 $\pm$ 0.003	-0.0039 $\pm$ 0.004	-0.0055 $\pm$ 0.005	-0.0049 $\pm$ 0.005	-0.0046 $\pm$ 0.006
dur_C	-0.081 $\pm$ 0.001	-0.080 $\pm$ 0.002	-0.078 $\pm$ 0.002	-0.077 $\pm$ 0.002	-0.076 $\pm$ 0.003	-0.075 $\pm$ 0.006
home=false	0.56 $\pm$ 0.008	0.55 $\pm$ 0.008	0.54 $\pm$ 0.009	0.53 $\pm$ 0.01	0.53 $\pm$ 0.01	-0.52 $\pm$ 0.01
dur_C and home=false	-0.39 $\pm$ 0.004	-0.39 $\pm$ 0.005	-0.38 $\pm$ 0.005	-0.38 $\pm$ 0.005	-0.38 $\pm$ 0.006	-0.37 $\pm$ 0.006
east_west=east and home=false	-0.59 $\pm$ 0.008	-0.59 $\pm$ 0.009	-0.58 $\pm$ 0.009	-0.58 $\pm$ 0.01	-0.58 $\pm$ 0.01	-0.58 $\pm$ 0.01

Table 8: Estimates of most important fixed effects (in terms of significance and effect size) with increasing number minimum of travel days required per user on weekdays

Fixed effect	**travel days $\geq$ 2**	**travel days $\geq$ 4**	**travel days $\geq$ 6**	**travel days $\geq$ 8**	**travel days $\geq$ 10**	**travel days $\geq$ 12**
	Estim. $\pm$ SEM	**Estim. $\pm$ SEM**	**Estim. $\pm$ SEM**	**Estim. $\pm$ SEM**	**Estim. $\pm$ SEM**	**Estim. $\pm$ SEM**
Intercept	-0.040 $\pm$ 0.005	-0.43 $\pm$ 0.006	-0.39 $\pm$ 0.007	-0.33 $\pm$ 0.009	-0.46 $\pm$ 0.009	-0.33 $\pm$ 0.01
dur_C	-0.093 $\pm$ 0.002	-0.090 $\pm$ 0.003	-0.083 $\pm$ 0.004	-0.081 $\pm$ 0.005	-0.078 $\pm$ 0.005	-0.080 $\pm$ 0.006
home=false	0.019 $\pm$ 0.01	0.033 $\pm$ 0.01	0.021 $\pm$ 0.02	0.015 $\pm$ 0.02	0.017 $\pm$ 0.02	0.015 $\pm$ 0.02
dur_C and home=false	-0.38 $\pm$ 0.007	-0.36 $\pm$ 0.008	-0.35 $\pm$ 0.009	-0.35 $\pm$ 0.01	-0.34 $\pm$ 0.01	-0.33 $\pm$ 0.01
east_west=east and home=false	-0.46 $\pm$ 0.01	-0.44 $\pm$ 0.02	-0.42 $\pm$ 0.02	-0.40 $\pm$ 0.02	-0.38 $\pm$ 0.02	-0.34 $\pm$ 0.03

Table 9: Estimates of most important fixed effects (in terms of significance and effect size) with increasing number minimum of travel days required per user on weekends

Quantifying complex patterns in high-resolution sleep activity data

Sigga Svala Jónasdóttir, Jari Saramaki and Sune Lehmann

February 27, 2021

Abstract

Until very recently, large-scale studies have been based on self-reported sleep estimates where key metrics were limited to quantities people could reasonably be expected to recall. Today, sleep recording technology has been revolutionized by wearable device, and sleep can easily be measured objectively *in situ* over long period of time. Consequently, the number studies drawing on such sleep recording devices is rapidly rising. However, it is not clear if methods and metrics of classical sleep epidemiology – developed for questionnaire research – are the right ones for detailed high-resolution sleep recordings. Thus, in this paper we ask whether there are salient features of sleep that the traditional metrics fail to capture? We begin to answer that question by first showing that sleep is multi-faceted process which manifests complex patterns across the population. We find wildly different sleep dynamics can appear similar in terms of traditional epidemiological sleep measures. Next, we introduce *the sleep portrait* that capture and visualise the richness of long-term patterns of individual sleep behavior. Finally, we propose a novel data-driven metric based on the skewness of sleep duration. The distribution of duration-skew across the population suggests that this new metric may allow us to estimate whether or not an individual's physiological sleep needs are met. Today, there is no method to quantify whether an individual is getting sufficient sleep, except comparing their typical sleep duration to official recommendations. Sleep duration continues to decrease across the world and sleep disorders become more prevalent, while the number of wearable device owners increases. We see a potential to accelerate understanding about the human sleep in modern society and conceivably improve health outcomes linked to poor sleep. Here we begin the process of identifying the appropriate tools and techniques for future studies of such data-sets.

Introduction

As humans have entered the modern era we have begun to collect data on a wide range of human behaviors [1, 2]. Our knowledge of sleep has developed from scattered mentions throughout historical sources to questionnaire data [3, 4] and to today's massive datasets [5–9]. While data collection on sleep behavior has been revolutionized over the past few years [5, 7, 10, 11], the commonly used metrics for analyzing sleep behavior [12–14] originate from a time when the key data source was questionnaire data.

Large scale sleep studies. There is no shortage of evidence that sleep is an essential part of human health but regardless, average sleep duration keeps decreasing across the world [3, 15–29]. For this reason, scientists have been interested in sleep at least since the days of Aristotle [30]. However, the study of quantitative population-level sleep patterns, *sleep epidemiology*, only dates back to around 1980, when the first documented studies were conducted [31, 32]. The most common ways to collect data about sleep at scale are i) sleep diaries, ii) sleep surveys, iii) sleep questionnaires, and the most recent technique, iv) wearable devices.

Bias in sleep epidemiology. In the case of sleep diaries and surveys, subjects self-report quantities and qualities regarding their sleep (e.g. bed-time, sleep duration and subjective tiredness). Studies have explored the extent to which self-reported sleep duration via sleep diaries and surveys corresponds to objectively measured estimates from wearable devices, and found them to correlate poorly, with systematic biases related to certain attributes [33–39]. Sleep questionnaires usually pose multiple questions to obtain a single score or estimate, intended to evaluate one aspect of an individual's sleep behavior. Examples of surveys are The Munich Chronotype Questionnaire (to quantify chronotype), Pittsburgh Sleep Quality Index (assesses sleep quality for the passed month) and the Karolinska Sleepiness scale (measure sleepiness) [4,12,14,40,41]. These questionnaires are validated and widely used, but only focus on one aspect of the multidimensional process of sleep [42–45].

Wearable devices for sleep epidemiology. Wearable devices are worn on the surface of the skin, most commonly around the wrist (wrist actigraphy) and monitor the users' movements to infer sleep. Wearables with built-in sleep tracking have become more common in the recent years, and thus have grown in importance for sleep research [5, 7–9, 46, 47]. Sleep trackers facilitate multi-night recordings *in-situ* and enable a more holistic examination of sleep [10, 48]. It is important to emphasize that sleep trackers are not perfect and are generally known to overestimate sleep duration and underestimate sleep fragmentation [10, 11]. However, as we argue below, despite these shortcomings consumer-grade sleep trackers are likely to provide better insights than self-reported single estimates of habitual sleep behavior. The dataset for this study stems from consumer wearable devices: an observational, global and large-scale sleep activity data from $\sim$ 15 000 users residing 149 countries, where each user has at least 56 recorded nights which amount to $\sim$ 5.5 million nights in total.

Brief overview of this paper. In this paper, we argue that the classic metrics from sleep epidemiology, while valuable, are not necessarily a perfect fit for data collected from consumer-grade sleep trackers. Specifically, we explore which aspects of sleep may remain hidden when focusing on classic measures of sleep, such as the habitual sleep duration, chronotype and social jetlag, as seen from the perspective of large-scale high-resolution sleep datasets. We introduce the idea of *sleep portraits* for quantifying sleep patterns and show how metrics derived from sleep portraits enable us to understand the factors behind the skewness of individual level sleep duration statistics. This skewness, apparent in our empirical electronic traces of sleep, is a feature that has been partially obscured by traditional measures of sleep. We explore the hypothesis that skewness of an individual's distribution of sleep duration is related to whether their sleep needs have been met.

Methods

Data collection & Pre-processing

The data analyzed here were collected with consumer wearable devices from 2015 to 2019 designed to track physical activity and sleep behavior. Users connect their devices to smartphones, and receive a visual instruction on how and where (wrist) to place the device. The devices use proprietary, internally validated algorithms based on movement registered by an internal accelerometer to infer sleep and wake states to 1-minute intervals (epochs). Epochs are aggregated into nights with sleep onset, offset and duration, and sleep fragmentation are quantified with wake after sleep onset (WASO) [49]. Measurements from the wristbands exhibit a high degree of validity since they i) replicate age-related sleep trends from previously published self-report studies, and ii) converge with country-level estimates of sleep measures from number of data-sets from other publications [7]. To reduce the risk of including sleep observations from those suffering from insomnia, shortened nights due to users ceasing wristband use in the middle of the resting period, observations from night-shift workers or any other possible data errors, outliers were removed. The details of the filtering process are described step-by-step in the SI, *Data Pre-Processing*. After the pre-processing, the final dataset used for analyses consists of $\sim$ 4 million weekday nights and $\sim$ 1.5 million weekend nights from approximately 15 000 users. By using these wristbands, we follow a growing trend of utilizing commercial devices in sleep research to study sleep behavior in naturalistic settings at large scales [6, 7, 38, 48].

Privacy & GDPR Users are anonymous and self-report their age, gender, height and weight. All data analysis was carried out in accordance with the EU's General Data Protection Regulation 2016/679 (GDPR) and the regulations set out by the Danish Data Protection Agency. The GDPR describes regulations for data protection and privacy in the European Union and the European Economic Area; it also addresses the transfer of personal data outside the EU and EEA areas.

Measuring skewness

A key element in our analysis below is *skew* or *skewness*, which measures the asymmetry of a probability distribution of random variable around its mean. We use the following expression for the skewness of a sample with n data points [50]:

$$skew = \frac{m_3}{s^3} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{[\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2]^{\frac{3}{2}}} \tag{1}$$

where m_3 is the sample third central moment, s is the sample standard deviation and n the number of datapoints. Note that real world data is almost never exactly symmetric (skew=0) due to noise, and therefore Bulmer *et al.* (1979) suggests the following rule of thumb to consider when working with empirical data [51]:

1. If skew is less than -1 or larger than 1 then the distribution is *highly skewed*
2. If skew is between -1 and -0.5 or between 0.5 and 1, then it is *moderately skewed*
3. If skew is measured between -0.5 and 0.5 then the distribution is approximately symmetric

Results

Sleep is wild and complex!

Summary of four sleep trajectories We begin our study by showing how a range behaviors can result in similar values for 'classical' measurements of sleep. In order to exemplify the wide range of behaviors observed when sleep is measured objectively *in-situ* over a long period of time, we choose four users (1-4) to analyze in detail. Table 1 exhibits their median bed-time (sleep onset), wake time (sleep offset) and sleep duration separately on weekdays and weekends, as well as their chronotype (MSF_{sc}) and social jetlag (SJ) [12,13]. Chronotype and social jetlag are widely used in sleep epidemiology and details on how they are derived is provided in section called '*Chronotype & Social jetlag*' in the SI. These summary statistics show a range of similarities and differences between the four users, which are further explored in Figure 1.

USER	**1**	**2**	**3**	**4**
Overall				
MSFsc [hh:mm]	04:26	04.35	05:13	03:42
Social jetlag [hrs]	0.27	0.88	1.45	0.13
Median duration [hrs]	9.10	4.82	7.30	7.4
Weekdays				
Median Onset [hh:mm]	23:27	01:17	23:56	23:24
Median Offset [hh:mm]	08:28	06:05	06:56	07:42
Median duration [hrs]	9.00	4.7	7.0	7.3
Weekends				
Median Onset [hh:mm]	23:29	02:00	00:12	23:34
Median Offset [hh:mm]	09:08	07:15	09:52	07:56
Median duration [hrs]	9.2	5.2	9.65	7.5

Table 1: Estimates of median bed and wake time (separately by weekdays and weekend), and estimated chronotype (MSF_{sc}), social jetlag (SJ) and overall median sleep duration for four selected users

Defining visual sleep trajectories. A convenient way to represent sleep recordings of an individual is shown in Figure 1C-F where the x-axis represents consecutive nights, and the y-axis shows the hours from 20:00 to 12:00. Sleep onsets and offset are indicated with blue and yellow dots respectively. Weekend-days are shown by a gray shading of the sleep-interval. We call these plots *visual sleep trajectories* and enable us to capture many aspects of sleep patterns, such as typical bed and wake-time, sleep regularity, particular habits on weekends, long-term trends, and more.

Analysis of four sleep trajectories: Figure illustrates how users 1 & 2 have nearly the same chronotype (both chosen within the range of dotted lines on Figure 1A), but their visual sleep trajectories are quite different; user 1 fluctuates in terms of sleep timing and does not exhibit a clear distinction between weekends and weekdays. In contrast, user 2 has consistent bed and wake time, but wake-time is shifted to later hours on weekends. Similarly, users 3 & 4 have nearly the same overall median sleep duration (7.3 and 7.4 hrs respectively, chosen withing the range of the dotted lines on Figure 1B), however their temporal patterns are quite different. User 3 is likely constrained by early morning work schedule (same wake-up time on weekdays), which is relieved on weekends. User 4 has an interesting behavior of falling asleep at almost the same time every night, all the while waking up at relatively random hours.

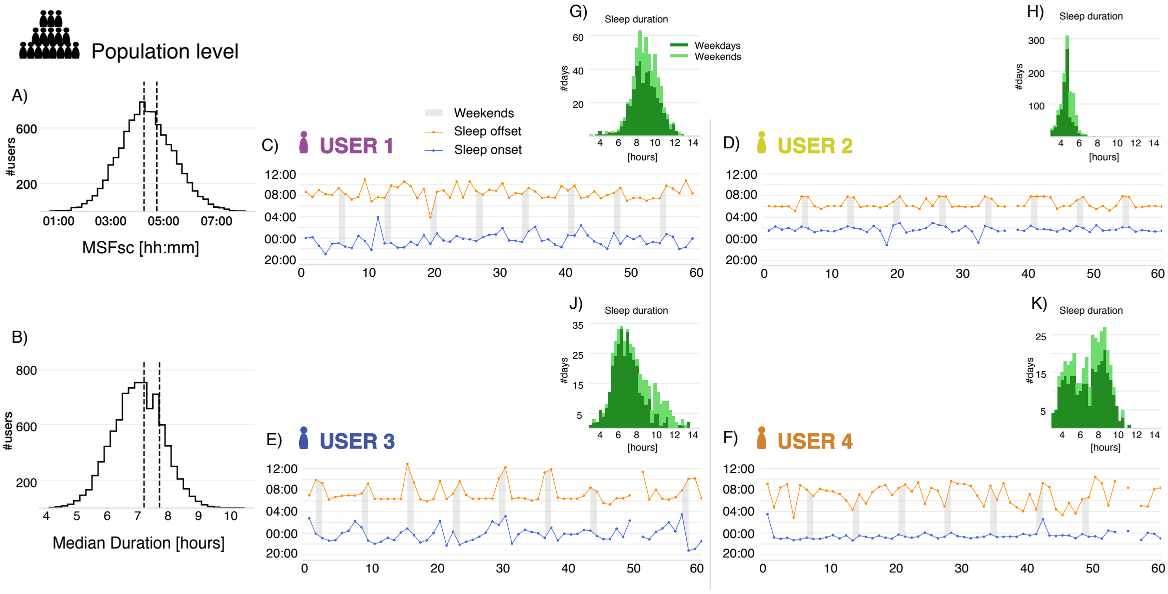


Figure 1: The multifaceted dynamics of sleep A: The distribution of chronotype (MSF_{sc}) across the population, and the dotted vertical line represent the range in which user 1 and 2 were selected from. **B:** The distribution of median sleep duration across the population, and the dotted vertical lines represent the range in which user 3 and 4 were selected from. **C-F:** Temporal patterns, or *sleep trajectories*, for 60 consecutive nights of bed and wake time. The vertical axis shows hours from 20:00 to 12:00 (next day). Sleep onset and offset are marked with blue and yellow dots respectively, and weekend nights are shown by a gray shading of the sleep-interval. Users 1 & 2 have nearly the same chronotype ($\sim$ 04:30) and users 3 & 4 have nearly the same overall medians sleep duration ($\sim$ 7.5 hrs). **G-K:** The stacked distributions of sleep duration (dark green color represents weekday-nights and lighter green color weekends) for users 1-4 respectively.

Same social jetlag but different weekend-weekday sleep duration difference. In Table 2 we provide examples of three additional users with nearly the same value of social jetlag (around 45 minutes), but different underlying sleep dynamics [13]. User 5 has similar wake time on weekends and weekdays, but bed-time is shifted to later hours on weekends, resulting in 1.3 hour shorter sleep duration on weekends. User 6 shifts both bed and wake time to later hours on weekends and therefore obtains the same amount of sleep both on weekday and weekend-nights. Lastly, user 7 goes earlier to bed on weekends compared to weekdays, and wakes up later resulting in more than three extra hours of sleep on weekends.

We believe that new metrics for sleep are needed. All of the examples of above (in Figure 1 and Table 2) illustrate that even though one aspect of sleep is measured to be similar across users, other (related) characteristics can be quite different. Users can have the same chronotype but have nearly two hour absolute difference in both bed and wake time (4 hrs in total, see median sleep onset and offset for user 1 and 2 in Table 1). Similarly, two users can change their mid-sleep marginal an equal amount between weekends and weekdays, but one might lose a lot of sleep by doing so, while the other gains substantial amount. Furthermore, we generalise these conclusions by illustrating the wide range of behaviors that distributions of different sleep metrics span for users with approximately the same chronotype ($\sim$04:30), social jetlag (45 minutes) and median sleep duration ($\sim$ 7.5 hrs) in Figures S2-S4. Taken together, these observations point out that existing measures and methods in sleep epidemiology may not be quantifying the complex patterns manifested during human sleep.

Users with SJ $\sim$ 45 minutes	**5**	**6**	**7**
Onset weekdays [hh:mm]	23:09	23:10	02:00
Offset weekdays [hh:mm]	08:15	06:00	07:37
Onset weekends [hh:mm]	01:16	23:53	00:48
Offset weekends [hh:mm]	08:06	06:59	10:28
Weekend-weekday duration difference [hrs]	-1.3	0.0	3.3

Table 2: Estimates of bed and wake-up time (separately on weekends and weekdays), and weekend-weekday median sleep duration difference for three selected users with the same social jetlag ($\sim$45 minutes)

Skew: Measuring the direction of preference

Introducing sleep-duration skew, Σ. Having discussed some shortcomings of the more traditional measures of sleep, we now move on to discuss possible new metrics driven by our exploration of empirical sleep patterns. Perhaps the most striking finding we have come across arises from investigating individual-level distributions of sleep duration over time. To see examples of distributions of an individuals' sleep duration over time, consider Figure 1C-F, which show the distribution of sleep duration for users 1-4 (see also Figure 2B for additional examples). User 1 is characterized by a relatively symmetric distribution except a thin left tail, user 2 has a narrow and balanced distribution, user 3 has an asymmetric distribution with more mass in the right-hand-side tail, consisting mainly of weekend nights (displayed in light green), and user 4 appears to have a bi-modal distributions which spans a broad range of behaviors.

Positive skew as a function of sleep duration. The symmetry or asymmetry of distributions of sleep duration (from now on Σ) may provide information about individual direction of preference, or sleep need. As a first observation, consider Table 3, which shows the percentage of users whose average sleep duration (μ) is larger than their median sleep duration (M). We show this fraction as a function of a user's *sleep group*, defined by rounding an individual's median sleep duration to the nearest half-hour bin.

Median sleep duration [hour]	**4.5**	**5.0**	**5.5**	**6.0**	**6.5**	**7.0**	**7.5**	**8.0**	**8.5**	**9.0**	**9.5**
% of users with $\mu >$ M	100	95	84	77	64	53	47	37	28	20	14

Table 3: Percentage of users with average (μ) larger than median (M) within sleep group (which are defined by rounding M sleep duration to the nearest half hour). Notice the monotonically decreasing trend.

The table clearly displays a monotonically decreasing trend in the fraction of users with the μ larger than M as a function of average sleep duration. Stated differently, the less someone sleeps, the more likely it is that they will experience an 'unusually' long (longer than median) sleep once in a while.

Skew and the direction of preference. We believe that a promising explanation (which we explore in more detail in the following) of this observation arises from the physiological process of *sleep-wake homeostasis* [52]. Sleep-wake homeostasis regulates sleep pressure and ensures that sleep takes place every night. For example, an individual who tends to sleep less than physiologically needed will build up sleep pressure from the last adequate sleep episode which can be eliminated by a long nighttime sleep (a 'catch-up' night) [52, 53]. In the case of insufficient sleep, these 'catch-up' nights result in a skewed distribution of sleep duration, with a disproportionately larger right tail or a positive skew. Similarly, we expect left-asymmetrical distributions for individuals who tend to have longer nighttime sleep than they can sustain.

Sleep skew across the population In order to study trends as a function of sleep duration, we estimate Σ_i for each individual i's distribution of sleep duration. Figure 2B shows examples of individual sleep duration distributions. Next, we aggregate individuals into the aforementioned sleep group, and calculate the average value for individuals in that group. The results are shown in Figure 2A, which shows that values of Σ tend to decrease as a function of sleep duration. In the Figure, error bars encode the standard error of the mean (SEM) Strikingly, we find that for all but the most extreme sleep duration (median sleep of less than 4.5 hours and 9 hours or more), we see a close-to-linear trend of average Σ for the sleep groups. We also observe that the sign of the average Σ changes between 7 and 7.5 hours of sleep coinciding with the recommended sleep duration [54–57]. The rightmost inset shows box-plots for the sleep groups and the leftmost inset shows shows that Σ is approximately normally distributed across the entire population.

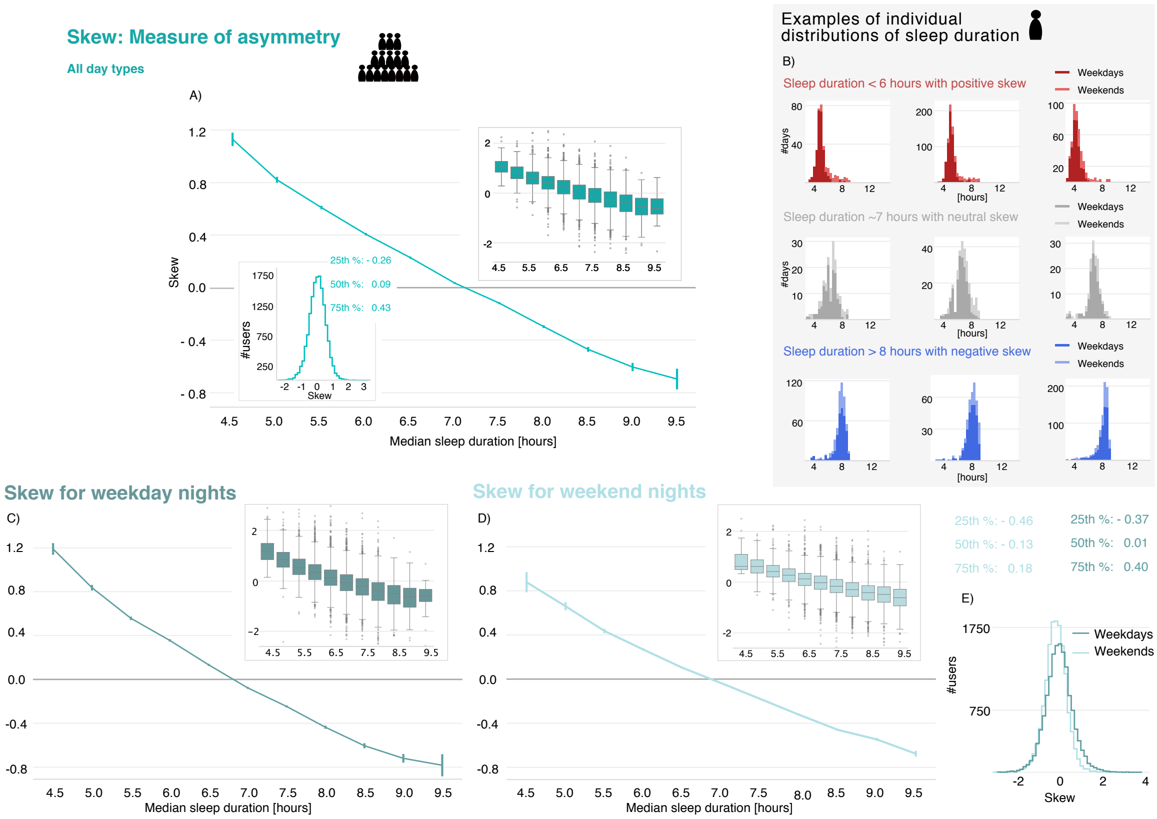


Figure 2: Systematic patterns of asymmetry for distribution of sleep duration across the population A: Aggregated skew, from the individual level for distribution of sleep duration, averaged by sleep group. Sleep groups are defined by rounding the median sleep duration to the next half hour bin and error estimates are marked with the standard error of the mean (SEM). The box plot in the upper right corner shows the underlying distribution behind the average for each sleep group, and the insert in the lower left corner illustrates how skew (for individual level distribution of sleep duration) distributes across the sample. **B:** Examples of individual level distributions of sleep duration for users with positive (red), neutral (grey) and negative (blue) skew. The distributions are stacked, where darker colors represent weekday nights and lighter color weekend nights. **C:** Aggregated skew, estimated for individual level for distribution of sleep duration for weekday nights, averaged by sleep group. Error estimates are marked with the standard error of the mean (SEM). The box plot in the upper right corner shows the underlying distribution behind the average for each sleep group. **D:** Aggregated skew, estimated for individual level for distribution of sleep duration for weekend nights, averaged by sleep group. Error estimates are marked with the standard error of the mean (SEM). The box plot in the upper right corner shows the underlying distribution behind the average for each sleep group. **E:** The distribution for skew (estimated for individual distribution of sleep duration) for weekday and weekend nights across the population

Considering weekday & weekend-night skew separately. Generally, it is considered important to measure sleep separately on both free (weekends) and work days (weekdays) since weekly social schedules are known to constrain the daily rhythm of rest and activity [12]. In terms of Σ, it also turns out to be interesting to consider the distribution of skew for weekday and weekend nights separately. Figure 2C and D illustrate the average skew estimated for sleep groups for weekdays and weekends, respectively. We observe same trend as before: skew decreases as median sleep duration increases. However, when we zero in on the absolute values of Σ, we see that the curve covers a wider range on weekdays compared to weekends. That difference can be explained by observing the comparison of the population-level distributions of skew (calculated from individual's distribution of sleep duration) separately for weekdays and weekends, shown in Figure 2E. The probability of having skew close to zero on weekends is substantially higher, and the distribution of weekday skew is slightly more symmetrical around zero, with the right tail shifted towards positive values. The observation that weekends are less constrained by social schedules might explain the difference in weekend and weekday skew [12]. On weekdays there is overall less time and flexibility, which might impose interruptions and alter typical behavioral patterns.

Low Σ for extreme median values of sleep duration. If we accept the hypothesis that Σ captures information about an individual's need for sleep by showing their direction of preference in terms of sleep homeostasis, it means that individuals with values of Σ close to zero are getting sufficient amounts of sleep. Therefore it is interesting to consider individuals who sleep very little (or a lot) and still have low values of Σ. This would potentially enable us to find people who require much less sleep than average – or individuals with an exceptionally high requirement for sleep. These individuals do exist: The box plot in Figure 2A illustrates the underlying distributions behind the average skew of each sleep group, and reveals that are individuals who have tend to have short sleep duration (<6.5 hrs) but a symmetric or negatively skewed distributions (disproportionate tendency for shorter nights). In exploring the characteristics of these 'out-of-the-ordinary' groups, we find that individuals that have a median sleep duration less than 6.5 hours and a negatively skewed or symmetric distribution (skew < 0.25) are likely to be older, from the East (Asia) and male – all demographic variables associated with short sleep duration. Similarly, those who tend to sleep for longer (> 8 hrs) and have positive or no skew (> 0.25) are more likely to be young, western (from Europe or North America) and female, all factors affiliated with long sleep duration. We provide a more in-depth exploration for these findings in the SI (*Short or long sleep duration but no skew?*). It is important to note here that there could be other explanations of low skew. For example, anyone sleeping according to a highly controlled schedule (e.g. going to bed at 1am and waking up at 6am every day), would also have short median sleep duration and low skew.

Capturing sleep complexity with sleep portraits

Introducing sleep portraits. So far, we have mainly discussed the strengths and weaknesses of existing, 'traditional' measures from sleep epidemiology in the context of high-resolution sleep recordings. We now turn our attention to novel measures, intended to capture the richness of empirical sleep trajectories. In the following, we will also aim to understand the phenomenon of skew in more detail and connect the different patterns observed in Figure 2. Figure 1 and 2. We begin by introducing *sleep portraits* that visualise the variation in bed and wake time as 2d-histograms (here using half-hour bins) separately for weekday and weekend nights. Figure 3 displays *sleep portraits* for the same users as in Figure 1. The color of each square represents the number of nights observed in that bin. To make the visualisation more interpretable, we have marked recommended sleep duration (7-9 hrs) with the grey step-lines [54–57]. We find that the sleep portraits reveal and highlight patterns that were not necessarily easy to see in the visual sleep trajectories (the illustrations from Figure 1 are shown above each sleep portrait in Figure 3).

Individual-level sleep patterns. The sleep portraits in Figure 3A reveal that user 1 has similar behavior on weekends and weekdays except that the wake time is slightly advanced to later hours on weekends. The large variability of bed and wake-times is clearly visible from the broad areas that the points cover in the sleep portrait. Furthermore, user 1 tends to sleep long, often >9 hrs. User 2 is highly regular, seen in the focused sleep portrait, where the data-points cover a very small area. While their sleep on/offset is shifted to later hours on weekends (the main mass of their sleep portrait is shifted to the right and up), this user consistently sleeps less than recommended both on weekdays and weekends, as all sleeps occur below the line of the shortest recommended sleep. The horizontally flat shape of user 3's data indicates that this person tends to wake up at approximately 7:30am on weekdays irrespective of sleep onset, which may be due to alarm clock on weekdays. On weekends we observe an 'explosion of freedom', and the user seems no longer constrained to wake up at 7:30. This behavior is common to many users in our dataset and, as we shall see below, part of what drives the positive skew of sleep durations discussed earlier. User 4's sleep portrait is flat in the vertical direction. This implies that user 4 tends to fall asleep at almost the same time all nights (midnight), while they wake up at highly variable hours. This unusual pattern is observed for both weekday and weekend nights.

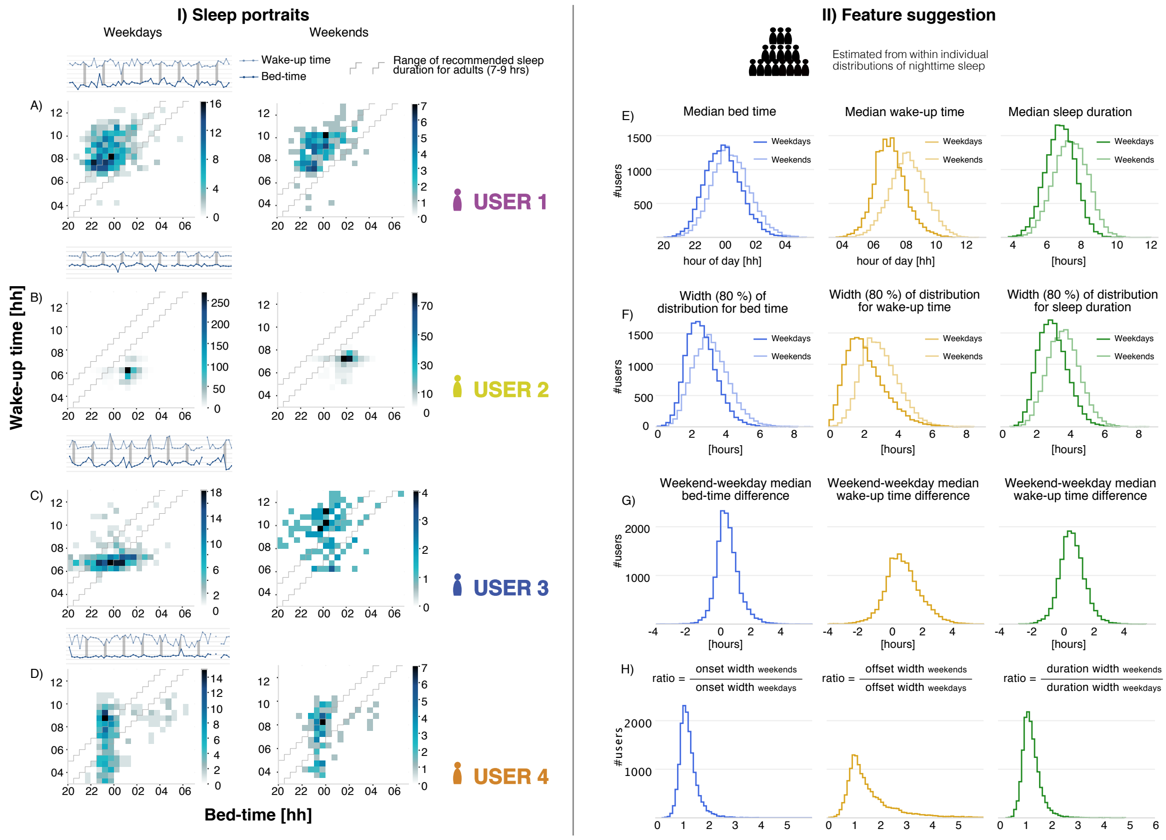


Figure 3: Sleep Portraits A-D: Examples of *the sleep portraits* for users 1-4 from Figure 1. The sleep portraits are 2d-histograms (1/2 hours bins) of bed time on the x-axis (range 22:00-07:00) and wake-time on the y-axis (range 03:00-13:00). For each user, the left plot represents weekdays and right plot weekends. The color-bar illustrates the number of nights taking place in each square. The grey stepped lines represent the area of recommended sleep duration (7-9 hrs).
Features suggestion E: The distribution of median bed-time, wake time and sleep duration separately for weekday and weekend-nights. **F:** The distribution for the width (10th percentile minus the 90th percentile) for bed-time, wake time and sleep duration separately for weekday and weekend-nights. **G:** The distribution for the weekend-weekday median difference for bed-time, wake time and sleep duration. **H:** Weekend-weekday width ratio for bed-time, wake time and sleep duration.

Large-scale descriptors based on sleep portraits

Based on the sleep portraits, we now introduce a set of sleep descriptors, which we argue are an interesting alternative set of measures for sleep epidemiology based on high-resolution data. As is clear from our discussion of the individual users above, the shape and location of the data displayed in the individual sleep portraits can tell us a lot about a person's sleeping habits.

The center, height, and width of the point cloud on the sleep portraits. The first descriptor is simply the center of the 2d-histogram or the 'point-cloud' (referred to as such henceforth) on the sleep portraits, which captures the typical sleep on/offset of a user. We use the median, which gives us a better estimate of typical behavior, since the mean tends to be highly influenced by the extreme values, which characterize many users (cf. our discussion of Σ above). The distributions for median sleep onset, offset and duration, separately by day type, are shown in Figure 3A. Next, as we argued above, the width of an individual's distribution of data in the sleep portrait contains important information about their sleep regularity. Typically, the standard deviation (std) is used to quantify sleep regularity [58–60], but we suggest using a measure based on quantiles. As is the case for the median, quantiles are less impacted by extreme events, and we also consider them to be more intuitively interpretable than, e.g., the standard deviation: We simply find it easier to understand the statement 'John wakes up 80 % of the time within a span of 15 minutes on weekdays' than 'John's has a standard deviation of 0.74 hours on weekdays'. As a specific measure, we suggest reporting the difference between the 90th and 10th percentile, which provides an estimate of where 80% of a person's sleep takes place. Therefore, we call this measure the *width at 80%*. We define a similar measure for the height (bed time) of the point-cloud. Other quantile-based measures could be equally meaningful. Finally, we do note that the width at 80% correlates well with std (ranging from 0.933 to 0.968, see Figure S5 in SI for more details). The distribution of the width at 80% for sleep onset, offset and duration on weekdays and weekends is presented in Figure 3B.

Weekdays and weekends. As is clear from the examples above, there is often a large difference in sleep behavior between weekdays and weekends, also established in previous research [61]. For this reason, we calculate portraits for weekdays and weekends separately, and also calculate measures that compare the weekday to weekend behavior, an analog to social jetlag Specifically, we measure the *weekend-weekday median difference* for sleep onset, offset and duration – distributions presented in Figure 3C. Lastly, we introduce a measure based on the 'explosion of freedom' observed in the sleep portrait for user 3 in Figure 3. To quantify this behavior, we propose to use the ratio between the width of a measure on weekends compared to weekdays, concretely defined as $ratio_{width} = \frac{width_{weekends}}{width_{weekdays}}$. The distribution for the weekend-weekday width ratio for sleep onset, offset and duration is illustrated in Figure 3D. We clearly observe the effect of alarm clocks on sleep offset, where the distribution is broader with a long right tail.

Connecting new measures to skew

Starting from a large scale dataset with high-resolution sleep recordings, our paper has three main parts. In the first part we pointed out skew as an important new metric for sleep, and reasoned that it might provide information about whether an individual is getting sufficient sleep or not. Our argument is two-fold; i) the observed correlation of skew with typical sleep duration across the population, and ii) a plausible physiological explanation relating to sleep/wake homeostasis. In the second part we proposed a number of new metrics for high-resolution sleep data (based on sleep portraits). However, what is still lacking is comparison between the richness of our novel metrics and traditional measures from sleep epidemiology. That is what we will investigate in the third part of the paper.

Setting up the prediction task To connect skew to other measures of sleep behavior, we define three user groups based on skew characteristics. For each of the groups, we select individuals with either the most positive, neutral and negative skewed distribution of sleep duration ($N = 2000$). To understand the relationship between classic and novel measures of sleep, we set up a prediction task for each of the skew group, where we add a random sample of 2000 individuals and then try predict whether users were originally selected to the group or not. The features used for the prediction task are novel sleep metrics proposed in Figure 3 alongside traditional measures of sleep from sleep epidemiology (chronotype, social jetlag as well as midsleep on weekdays). By considering *the feature importance* in this prediction task we can understand i) which metrics enable us to separate the skew groups from a set of random users, ii) what characterizes each of the skew group, and iii) allow us to compare the efficacy of the new metrics to traditional ones with respect to explaining skew. Specifically, we train a decision tree classifier to predict whether an individual belongs to the skew group or not (baseline 50% accuracy), results are summarised in Figure 4.

Duration plays a key role. Using all metrics introduced in Figure 3 alongside the traditional sleep epidemiology measures (chronotype, midsleep on weekdays and social jetlag), we are able to predict *positive skew* with 79 % accuracy for all day types, 76 % when skew is estimated for sleep duration on weekdays alone and with 73 % accuracy for weekend nights. The most important feature in all cases is the median sleep duration, which makes sense because of the close connection to sleep need. Accordingly, we also see a distinct difference between the groups on distributions of median sleep duration in Figure 4D-F. This feature is the primary explanation for the prediction task of positive skew on weekends (see Figure 4C and distributions in Figure S14). Furthermore, this skew group (positive skewed distribution for weekend-nights) also has significantly shorter median sleep duration on weekdays, thus overall sleeps less than the other groups (see Figure S14).

Characterizing positive skew. Another important characteristic for positively skewed individuals when estimated for all day types or weekdays alone, is that they are likelier to have higher weekend-weekday median sleep duration difference (see Figure 4G and K, as well as distributions in Figures S12 & S13 in the SI). As mentioned above, this characteristic describes those who are likely sleep deprived on weekdays due to the social clock [13, 61], and use the weekends to catch up. Furthermore, individuals in positive skew group have a broader distribution of sleep duration on weekends compared to weekdays, which is another indication of misalignment induced by the social clock.

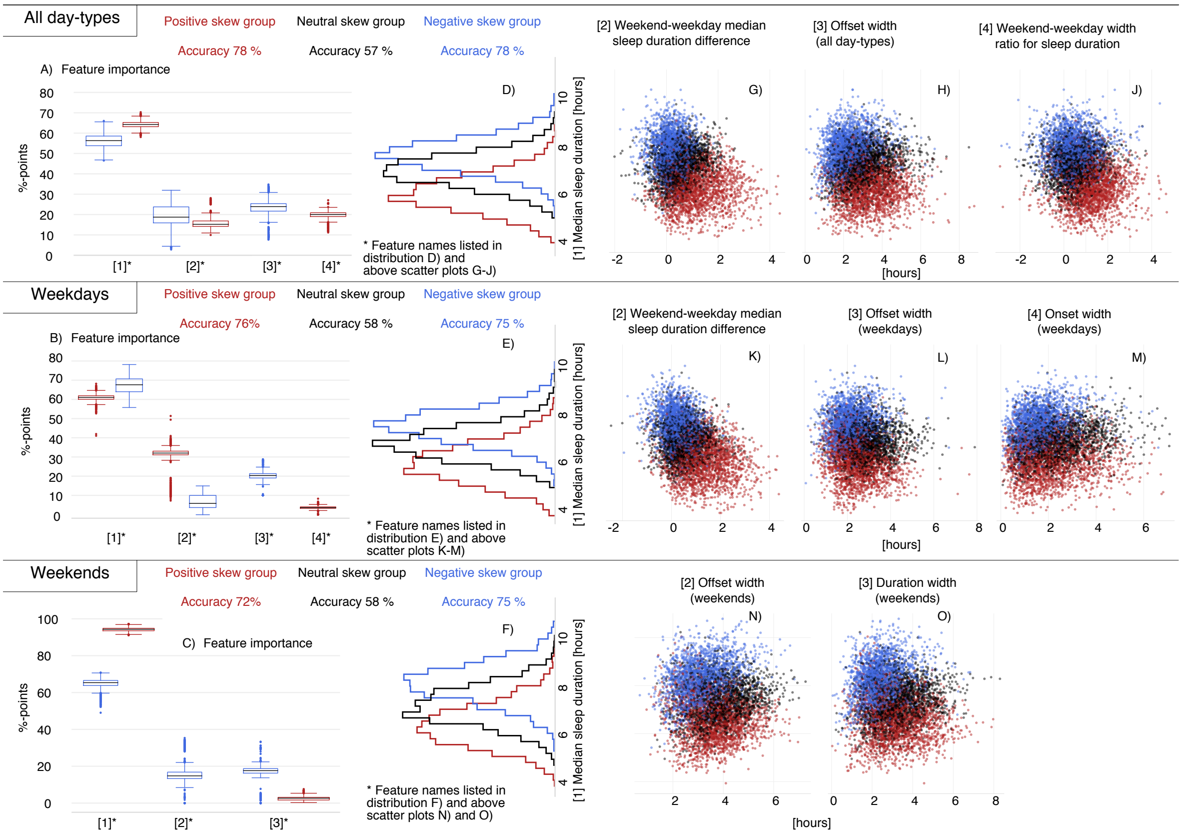


Figure 4: Information about skew groups from decision tree classifier A decision tree classifier is trained to predict whether an individual belongs to skew group or not. There are three groups, and each comprising 2000 individuals with either the most positive, neutral and negative skewed distribution of sleep duration, as well as a random selection of 2000 other individuals. **A-C:** Feature importance obtained from decision tree classifier from prediction of positive and negative skew group, where skew is estimated from the distribution of sleep duration for all day-types on **A**, weekday-nights on **B** and weekend-nights on **C**.

D-F: The distribution of median sleep duration for the individuals (N=2000) with the most positive, neutral and negative skewed distribution of sleep duration for all day-types on **D**, weekday-nights on **E** and weekend-nights on **F**.

G-J: Scatter plot of median sleep duration (estimated for all day-types) with weekend-weekday sleep duration difference, offset width (all day-types) and weekend-weekday width ratio for median sleep duration where points are colored by skew group; red, black and blue for positive, neutral and negative skewed distribution of sleep duration respectively. **K-M:** Scatter plot of median sleep duration (estimated for weekday nights) with weekend-weekday sleep duration difference, offset width (weekdays) and onset width (weekdays) where points are colored by skew group; red, black and blue for positive, neutral and negative skewed distribution of sleep duration respectively. **N-O:** Scatter plot of median sleep duration (estimated for weekend-nights) with offset width (weekends) and duration width (weekends) where points are colored by skew group; red, black and blue for positive, neutral and negative skewed distribution of sleep duration respectively.

Characterizing negative skew. People in the *negative skew* group can be predicted with 78 % accuracy if skew is estimated for all day types, 75 % accuracy if skew estimated for either weekday or weekend nights alone. As before, median sleep duration sets the user group apart from the rest with substantially longer overall nighttime sleep, clearly visible in the distributions displayed in Figure 4A-C. Individuals with negative skew, when classified by the distribution of sleep duration for all day types or just weekdays, tend to have shorter weekend-weekday median sleep duration difference than the other groups (see Figure 4G and K, as well as distributions in Figures S12 & S13 in the SI). Additionally, individuals in the negative skew group generally have more regular sleep patterns and span a narrower distributions of width at 80 % for sleep onset, offset and duration than the other groups (see Figure S14 in the SI). All of these characteristics are qualities of stable sleepers who sleep enough.

Characterizing neutral skew. The *neutral skew* group can only be predicted with slightly better accuracy than the baseline (57-58 % accuracy). The group is not characterised by any specific quality and their behavior falls between the positive and negative skew groups in all measured aspects of sleep (see Figure 4, and Figures S12-S14 in the SI).

Discussion

Drawing from a data-set of 5.5 million nights from approximately 15 000 users with high resolution objective recordings of sleep *in-situ*, we show that the detailed trajectories of nighttime sleep have complex and multifaceted patterns across the population. The most commonly used metrics to analyse sleep at large-scale were developed during a time where the key data source were self-reported estimates of sleep. However, wearable technologies are becoming more commonly used in sleep research and we argue that there is a need for new metrics and methods to study high-resolution sleep activity data-sets they produce [2, 5–8, 10]. We introduce a new visualisation method called *the sleep portrait* which illustrates the complexity of individual-level sleep behavior. Furthermore, we propose a novel data-driven metric, *skew*, to estimate whether individual direction of preference indicates an overall lack of sleep or adequate attainment.

Skewness of individual-level distribution of sleep duration quantifies an individual's tendency for long or short nighttime sleep, relative to typical behavior. We find skew to depend nearly linearly on median sleep duration. Those who tend to have short nighttime sleep ($<$ 6.5 hrs) are likely positively skewed and have a higher proportion of nights with longer than their median sleep duration. The prevalence of positive skew decreases as median sleep duration increases and the skew becomes negative for those who have long ($>$ 8 hrs) median sleep duration or a higher tendency for short nights (relative to their typical behavior). In this work we argue that a person's skew may reveal whether or not that person's physiological sleep needs are met.

The finding that skew may contain information about sleep need is important because of the ways in which we are currently able to measure the fraction of individuals that are getting sufficient sleep depends on self-reports. Information about sleep needs have been established by studying health outcomes for groups with different range of self-reported habitual nighttime sleep [54–57]. Recently, the US National Sleep Foundation conducted a scientifically rigorous update on the the topic of sleep need, where an expert committee concluded that median sleep duration from 4 up to 7 hours may be appropriate for some individuals [55]. These guidelines set the current standards, but it has been argued that they are generalized and may not apply to all [54–57]. With further validation, our metric may make it possible to provide a data-driven way to understand whether individual sleep needs are met directly from high-resolution recordings of sleep duration.

Sleep has been studied at large scale across the population for decades, but the research has been shaped by characteristics of data sources, self-reports via sleep diaries, surveys or questionnaires, all of which are known to be subject to recall biases [33–39,62,63]. Naturally, these data collection strategies have influenced the development of methods and metrics used in the field, where some of the main quantitative estimates for analysis of sleep across the population are retrospective, and refer to typical self-reported sleep duration, chronotype & social jetlag estimated with Munich Chronotype Questionnaire (MCQT) [3, 4, 12, 60, 61, 64, 65]. Our work argues that individuals with similar estimates of any of these three metrics, can differ largely with respect to other characteristics of sleep.

In terms of visualising multi-night recordings of sleep, there is no convention for any fixed approach in the literature, but we find both Walch *et al.* (2016) and Roenneberg *et al.* (2012) to visualise sleep start and end to span a horizontal line [59, 61]. The sleep portraits provide an alternative to that approach, since their purpose is to highlight characteristic patterns of individual sleep behavior. The new features we have introduced were partially inspired by the sleep portraits, but also built upon knowledge from previous sleep studies. In constructing these features we emphasized the importance of analysing sleep separately for weekend and weekday nights, considered the misalignment that the social clock might induce and regarded qualitative information that sleep regularity entails [58, 61, 66].

Our work has a number of limitations that are important to discuss. First, we underline that our proposed metric of skew has not been validated using subjective nor qualitative estimates of sleep. To mitigate these limitation we show that: **i)** skew depends approximately linearly on median sleep duration across the population, **ii)** the skew group with an average estimate closest to zero attains typically 7 hours of sleep per night which is in accordance with current recommendation [54–57], **iii)** individuals who are not skewed but obtain overall long or short nighttime sleep, are likely to have demographic characteristic which are associated with either long er short sleep duration, and **iv)** we find the group with the most positive skewed individuals to have higher weekend-weekday sleep offset and duration difference than other groups, which is behavior linked to unhealthy sleep patterns [13, 61, 67–70], and **v)** individuals with negative skew exhibit more sleep regularity – a behavior associated with good sleep hygiene [58, 66, 71].

Our sample of users may not be representative of the wider population due to potential unobserved factors also associated with wearable device ownership [72]. We also note that the wristbands have not been validated using the gold standard of polysomnography as recommended in the Sleep Research Society Workshop on wearable devices for the measurement of sleep [73]. However, we find that i) our data-set converges with country-level sleep measures from separate large-scale data-sets, ii) demonstrates consistency over the period of observation, and iii) replicates age-related sleep trends from previously published self-report studies, including changes in sleep duration and timing [7]. The devices have also been internally validated by the manufacturer. Due to the nature of the data sampling, we can not compare estimates of skew to subjective nor qualitative estimates of sleep, but we consider the most pressing matter for future studies.

Finally, we note that while we are hopeful that this study will spark debate about novel measures suitable for large-scale and high-resolution sleep recordings, we also acknowledge that this is only a first step. We are excited for a new paradigm of discoveries within sleep research driven by these data sources. In particular, while our sleep portraits and novel metric of skew capture some new patterns of the complex phenomenon that is sleep in the wild, we have not considered temporal correlations in the data, something which we believe holds a new richness of behavior to discover.

References

[1] Bruno Gonçalves and Nicola Perra. *Social phenomena: From data analysis to models*. Springer, 2015.

[2] Ignacio Perez-Pozuelo, Bing Zhai, Joao Palotti, Raghvendra Mall, Michaël Aupetit, Juan M Garcia-Gomez, Shahrad Taheri, Yu Guan, and Luis Fernandez-Luque. The future of sleep health: a data-driven revolution in sleep science and medicine. *NPJ digital medicine*, 3(1):1–15, 2020.

[3] Earl S Ford, Timothy J Cunningham, and Janet B Croft. Trends in self-reported sleep duration among us adults from 1985 to 2012. *Sleep*, 38(5):829–832, 2015.

[4] Till Roenneberg, Tim Kuehnle, Myriam Juda, Thomas Kantermann, Karla Allebrandt, Marijke Gordijn, and Martha Merrow. Epidemiology of the human circadian clock. *Sleep medicine reviews*, 11(6):429–438, 2007.

[5] Ju Lynn Ong, Jesisca Tandi, Amiya Patanaik, June C Lo, and Michael WL Chee. Large-scale data from wearables reveal regional disparities in sleep patterns that persist across age and sex. *Scientific reports*, 9(1):1–8, 2019.

[6] Stuti J Jaiswal, Giorgio Quer, Michael Galarnyk, Steven R Steinhubl, Eric J Topol, and Robert L Owens. Association of sleep duration and variability with body mass index: Sleep measurements in a large us population of wearable sensor users. *JAMA Internal Medicine*, 180(12):1694–1696, 2020.

[7] Sigga Svala Jonasdottir, Kelton Minor, and Sune Lehmann. Gender differences in night-time sleep patterns and variability across the adult lifespan: a global-scale wearables study. *Sleep*, 2020.

[8] L Kuula, M Gradisar, K Martinmäki, C Richardson, D Bonnar, K Bartel, C Lang, L Leinonen, and AK Pesonen. Using big data to explore worldwide trends in objective sleep in the transition to adulthood. *Sleep Medicine*, 62:69–76, 2019.

[9] Guy Fagherazzi, Douae El Fatouhi, Alice Bellicha, Amin El Gareh, Aurélie Affret, Courtney Dow, Lidia Delrieu, Matthieu Vegreville, Alexis Normand, Jean-Michel Oppert, et al. An international study on the determinants of poor sleep amongst 15,000 users of connected devices. *Journal of Medical Internet Research*, 19(10):e363, 2017.

[10] Avi Sadeh. The role and validity of actigraphy in sleep medicine: an update. *Sleep medicine reviews*, 15(4):259–267, 2011.

[11] Bhanu Prakash Kolla, Subir Mansukhani, and Meghna P Mansukhani. Consumer sleep tracking devices: a review of mechanisms, validity and utility. *Expert review of medical devices*, 13(5):497–506, 2016.

[12] Till Roenneberg, Anna Wirz-Justice, and Martha Merrow. Life between clocks: daily temporal patterns of human chronotypes. *Journal of biological rhythms*, 18(1):80–90, 2003.

[13] Marc Wittmann, Jenny Dinich, Martha Merrow, and Till Roenneberg. Social jetlag: misalignment of biological and social time. *Chronobiology international*, 23(1-2):497–509, 2006.

[14] Janet S Carpenter and Michael A Andrykowski. Psychometric evaluation of the pittsburgh sleep quality index. *Journal of psychosomatic research*, 45(1):5–13, 1998.

[15] Michael R Irwin. Why sleep is important for health: a psychoneuroimmunology perspective. *Annual review of psychology*, 66:143–172, 2015.

[16] Michael A Grandner, Nicholas J Jackson, Bilgay Izci-Balserak, Rebecca A Gallagher, Renee Murray-Bachmann, Natasha J Williams, Nirav P Patel, and Girardin Jean-Louis. Social and behavioral determinants of perceived insufficient sleep. *Frontiers in neurology*, 6:112, 2015.

[17] Rebecca A Bernert, Joanne S Kim, Naomi G Iwata, and Michael L Perlis. Sleep disturbances as an evidence-based suicide risk factor. *Current psychiatry reports*, 17(3):15, 2015.

[18] Andrea N Goldstein and Matthew P Walker. The role of sleep in emotional brain function. *Annual review of clinical psychology*, 10:679–708, 2014.

[19] Faith S Luyster, Patrick J Strollo, Phyllis C Zee, and James K Walsh. Sleep: a health imperative. *Sleep*, 35(6):727–734, 2012.

[20] Francesco P Cappuccio, Daniel Cooper, Lanfranco D'elia, Pasquale Strazzullo, and Michelle A Miller. Sleep duration predicts cardiovascular outcomes: a systematic review and meta-analysis of prospective studies. *European heart journal*, 32(12):1484–1492, 2011.

[21] Francesco P Cappuccio, Lanfranco D'Elia, Pasquale Strazzullo, and Michelle A Miller. Sleep duration and all-cause mortality: a systematic review and meta-analysis of prospective studies. *Sleep*, 33(5):585–592, 2010.

[22] Francesco P Cappuccio, Frances M Taggart, Ngianga-Bakwin Kandala, Andrew Currie, Ed Peile, Saverio Stranges, and Michelle A Miller. Meta-analysis of short sleep duration and obesity in children and adults. *Sleep*, 31(5):619–626, 2008.

[23] Christopher Ryan King, Kristen L Knutson, Paul J Rathouz, Steve Sidney, Kiang Liu, and Diane S Lauderdale. Short sleep duration and incident coronary artery calcification. *Jama*, 300(24):2859–2866, 2008.

[24] June J Pilcher and Allen I Huffcutt. Effects of sleep deprivation on performance: a meta-analysis. *Sleep*, 19(4):318–326, 1996.

[25] Matthew P Walker. A societal sleep prescription. *Neuron*, 103(4):559–562, 2019.

[26] Charles A Czeisler. Perspective: casting light on sleep deficiency. *Nature*, 497(7450):S13–S13, 2013.

[27] Saverio Stranges, William Tigbe, Francesc Xavier Gómez-Olivé, Margaret Thorogood, and Ngianga-Bakwin Kandala. 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*, 35(8):1173–1181, 2012.

[28] Erkki Kronholm, Timo Partonen, Tiina Laatikainen, Markku Peltonen, Mikko Härmä, Christer Hublin, Jaako Kaprio, Arja R Aro, Markku Partinen, Mikael Fogelholm, et al. Trends in self-reported sleep duration and insomnia-related symptoms in finland from 1972 to 2005: a comparative review and re-analysis of finnish population samples. *Journal of sleep research*, 17(1):54–62, 2008.

[29] Kristen L Knutson, Eve Van Cauter, Paul J Rathouz, Thomas DeLeire, and Diane S Lauderdale. Trends in the prevalence of short sleepers in the usa: 1975–2006. *Sleep*, 33(1):37–45, 2010.

[30] David Gallop. *Aristotle: On sleep and dreams*. Oxford University Press, 1996.

[31] Edward O Bixler, Anthony Kales, Constantine R Soldatos, Joyce D Kales, and S Healey. Prevalence of sleep disorders in the los angeles metropolitan area. *The American journal of psychiatry*, 1979.

[32] Peretz Lavie. Sleep habits and sleep disturbances in industrial workers in israel: main findings and some characteristics of workers complaining of excessive daytime sleepiness. *Sleep*, 4(2):147–158, 1981.

[33] Diane S Lauderdale, Kristen L Knutson, Lijing L Yan, Kiang Liu, and Paul J Rathouz. Sleep duration: how well do self-reports reflect objective measures? the cardia sleep study. *Epidemiology (Cambridge, Mass.)*, 19(6):838, 2008.

[34] Chandra L Jackson, Sanjay R Patel, W Braxton Jackson, Pamela L Lutsey, and Susan Redline. Agreement between self-reported and objectively measured sleep duration among white, black, hispanic, and chinese adults in the united states: Multi-ethnic study of atherosclerosis. *Sleep*, 41(6):zsy057, 2018.

[35] Elizabeth M Cespedes, Frank B Hu, Susan Redline, Bernard Rosner, Carmela Alcantara, Jianwen Cai, Martica H Hall, Jose S Loredo, Yasmin Mossavar-Rahmani, Alberto R

Ramos, et al. Comparison of self-reported sleep duration with actigraphy: results from the hispanic community health study/study of latinos sueño ancillary study. *American journal of epidemiology*, 183(6):561–573, 2016.

[36] Katarina Aili, Sofia Åström-Paulsson, Ulrich Stoetzer, Magnus Svartengren, and Lena Hillert. Reliability of actigraphy and subjective sleep measurements in adults: the design of sleep assessments. *Journal of Clinical Sleep Medicine*, 13(1):39–47, 2017.

[37] Marcela Zambrim Campanini, Esther Lopez-Garcia, Fernando Rodríguez-Artalejo, Alberto Durán González, Selma Maffei Andrade, and Arthur Eumann Mesas. Agreement between sleep diary and actigraphy in a highly educated brazilian population. *Sleep medicine*, 35:27–34, 2017.

[38] Jennifer Girschik, Lin Fritschi, Jane Heyworth, and Flavie Waters. Validation of self-reported sleep against actigraphy. *Journal of epidemiology*, 22(5):462–468, 2012.

[39] Erin O'Brien, Chantelle Hart, and Rena R Wing. Discrepancies between self-reported usual sleep duration and objective measures of total sleep time in treatment-seeking overweight and obese individuals. *Behavioral sleep medicine*, 14(5):539–549, 2016.

[40] Till Roenneberg, Luísa K Pilz, Giulia Zerbini, and Eva C Winnebeck. Chronotype and social jetlag: a (self-) critical review. *Biology*, 8(3):54, 2019.

[41] Torbjörn Åkerstedt and Mats Gillberg. Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience*, 52(1-2):29–37, 1990.

[42] Jason C Cole, Sarosh J Motivala, Daniel J Buysse, Michael N Oxman, Myron J Levin, and Michael R Irwin. Validation of a 3-factor scoring model for the pittsburgh sleep quality index in older adults. *Sleep*, 29(1):112–116, 2006.

[43] Sherry A Beaudreau, Adam P Spira, Anita Stewart, Eric J Kezirian, Li-Yung Lui, Kristine Ensrud, Susan Redline, Sonia Ancoli-Israel, Katie L Stone, Study of Osteoporotic Fractures, et al. Validation of the pittsburgh sleep quality index and the epworth sleepiness scale in older black and white women. *Sleep medicine*, 13(1):36–42, 2012.

[44] Michael A Grandner, Daniel F Kripke, In-Young Yoon, and Shawn D Youngstedt. Criterion validity of the pittsburgh sleep quality index: Investigation in a non-clinical sample. *Sleep and biological rhythms*, 4(2):129–136, 2006.

[45] Kosuke Kaida, Masaya Takahashi, Torbjörn Åkerstedt, Akinori Nakata, Yasumasa Otsuka, Takashi Haratani, and Kenji Fukasawa. Validation of the karolinska sleepiness scale against performance and eeg variables. *Clinical neurophysiology*, 117(7):1574–1581, 2006.

[46] Till Roenneberg, Lena K Keller, Dorothee Fischer, Joana L Matera, Céline Vetter, and Eva C Winnebeck. Human activity and rest in situ. In *Methods in enzymology*, volume 552, pages 257–283. Elsevier, 2015.

[47] X Xu, MP Conomos, O Manor, JE Rohwer, AT Magis, and JC Lovejoy. Habitual sleep duration and sleep duration variation are independently associated with body mass index. *International Journal of Obesity*, 42(4):794–800, 2018.

[48] Sonia Ancoli-Israel, Roger Cole, Cathy Alessi, Mark Chambers, William Moorcroft, and Charles P Pollak. The role of actigraphy in the study of sleep and circadian rhythms. *Sleep*, 26(3):342–392, 2003.

[49] Deepak Shrivastava, Syung Jung, Mohsen Saadat, Roopa Sirohi, and Keri Crewson. How to interpret the results of a sleep study. *Journal of community hospital internal medicine perspectives*, 4(5):24983, 2014.

[50] David P Doane and Lori E Seward. Measuring skewness: a forgotten statistic? *Journal of statistics education*, 19(2), 2011.

[51] Michael George Bulmer. *Principles of statistics*. Courier Corporation, 1979.

[52] Alexander A Borbély. A two process model of sleep regulation. *Hum neurobiol*, 1(3):195–204, 1982.

[53] Serge Daan, DG Beersma, and Alexander A Borbély. Timing of human sleep: recovery process gated by a circadian pacemaker. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 246(2):R161–R183, 1984.

[54] Consensus Conference Panel, Nathaniel F Watson, M Safwan Badr, Gregory Belenky, Donald L Bliwise, Orfeu M Buxton, Daniel Buysse, David F Dinges, James Gangwisch, Michael A Grandner, 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. *Journal of Clinical Sleep Medicine*, 11(6):591–592, 2015.

[55] Max Hirshkowitz, Kaitlyn Whiton, Steven M Albert, Cathy Alessi, Oliviero Bruni, Lydia DonCarlos, Nancy Hazen, John Herman, Eliot S Katz, Leila Kheirandish-Gozal, et al. National sleep foundation's sleep time duration recommendations: methodology and results summary. *Sleep health*, 1(1):40–43, 2015.

[56] Sutapa Mukherjee, Sanjay R Patel, Stefanos N Kales, Najib T Ayas, Kingman P Strohl, David Gozal, and Atul Malhotra. An official american thoracic society statement: the importance of healthy sleep. recommendations and future priorities. *American journal of respiratory and critical care medicine*, 191(12):1450–1458, 2015.

[57] Marie-Pierre St-Onge, Michael A Grandner, Devin Brown, Molly B Conroy, Girardin Jean-Louis, Michael Coons, and Deepak L Bhatt. Sleep duration and quality: impact on lifestyle behaviors and cardiometabolic health: a scientific statement from the american heart association. *Circulation*, 134(18):e367–e386, 2016.

[58] Bei Bei, Joshua F Wiley, John Trinder, and Rachel Manber. Beyond the mean: a systematic review on the correlates of daily intraindividual variability of sleep/wake patterns. *Sleep medicine reviews*, 28:108–124, 2016.

[59] Olivia J Walch, Amy Cochran, and Daniel B Forger. A global quantification of "normal" sleep schedules using smartphone data. *Science advances*, 2(5):e1501705, 2016.

[60] Dorothee Fischer, David A Lombardi, Helen Marucci-Wellman, and Till Roenneberg. Chronotypes in the us–influence of age and sex. *PloS one*, 12(6):e0178782, 2017.

[61] Till Roenneberg, Karla V Allebrandt, Martha Merrow, and Céline Vetter. Social jetlag and obesity. *Current Biology*, 22(10):939–943, 2012.

[62] Christopher B Miller, Christopher J Gordon, Leanne Toubia, Delwyn J Bartlett, Ronald R Grunstein, Angela L D'Rozario, and Nathaniel S Marshall. Agreement between simple questions about sleep duration and sleep diaries in a large online survey. *Sleep Health*, 1(2):133–137, 2015.

[63] Lianne M Kurina, Martha K McClintock, Jen-Hao Chen, Linda J Waite, Ronald A Thisted, and Diane S Lauderdale. Sleep duration and all-cause mortality: a critical review of measurement and associations. *Annals of epidemiology*, 23(6):361–370, 2013.

[64] Yoko Komada, Isa Okajima, Shingo Kitamura, and Yuichi Inoue. A survey on social jetlag in japan: a nationwide, cross-sectional internet survey. *Sleep and Biological Rhythms*, 17(4):417–422, 2019.

[65] Zhiwei Liu, Yingying Dong, Ying Xu, and Fei Zhou. Chronotype distribution in the chinese population. *Brain Science Advances*, 6(2):81–91, 2020.

[66] Danica C Slavish, Daniel J Taylor, and Kenneth L Lichstein. Intraindividual variability in sleep and comorbid medical and mental health conditions. *Sleep*, 42(6):zsz052, 2019.

[67] Anitra DM Koopman, Simone P Rauh, Esther van 't Riet, Lenka Groeneveld, Amber A Van Der Heijden, Petra J Elders, Jacqueline M Dekker, Giel Nijpels, Joline W Beulens, and Femke Rutters. The association between social jetlag, the metabolic syndrome, and type 2 diabetes mellitus in the general population: the new hoorn study. *Journal of biological rhythms*, 32(4):359–368, 2017.

[68] Susan Kohl Malone, Babette Zemel, Charlene Compher, Margaret Souders, Jesse Chittams, Aleda Leis Thompson, Allan Pack, and Terri H Lipman. Social jet lag, chronotype and body mass index in 14–17-year-old adolescents. *Chronobiology international*, 33(9):1255–1266, 2016.

[69] Michael J Parsons, Terrie E Moffitt, Alice M Gregory, Sidra Goldman-Mellor, Patrick M Nolan, Richie Poulton, and Avshalom Caspi. Social jetlag, obesity and metabolic disorder: investigation in a cohort study. *International Journal of Obesity*, 39(5):842–848, 2015.

[70] Femke Rutters, Sofie G Lemmens, Tanja C Adam, Marijke A Bremmer, Petra J Elders, Giel Nijpels, and Jacqueline M Dekker. Is social jetlag associated with an adverse endocrine, behavioral, and cardiovascular risk profile? *Journal of biological rhythms*, 29(5):377–383, 2014.

[71] Hylton E Molzof, Sarah E Emert, Joshua Tutek, Mazheruddin M Mulla, Kenneth L Lichstein, Daniel J Taylor, and Brant W Riedel. Intraindividual sleep variability and its association with insomnia identity and poor sleep. *Sleep medicine*, 52:58–66, 2018.

[72] Soultana Macridis, Nora Johnston, Steven Johnson, and Jeff K Vallance. Consumer physical activity tracking device ownership and use among a population-based sample of adults. *PloS one*, 13(1):e0189298, 2018.

[73] Christopher M Depner, Philip C Cheng, Jaime K Devine, Seema Khosla, Massimiliano de Zambotti, Rébecca Robillard, Andrew Vakulin, and Sean PA Drummond. Wearable technologies for developing sleep and circadian biomarkers: a summary of workshop discussions. *sleep*, 43(2):zsz254, 2020.

[74] Tim Althoff, Eric Horvitz, Ryen W White, and Jamie Zeitzer. Harnessing the web for population-scale physiological sensing: A case study of sleep and performance. In *Proceedings of the 26th international conference on World Wide Web*, pages 113–122, 2017.

Supplementary Information

Data Pre-Processing

The raw data consists 1-minute epochs of sleep activity which are aggregated into nights with sleep onset, offset, duration and wake-time after sleep onset (WASO). For each night a user can wake up multiple times but each awakening can only last for 60 minutes or less. For each the day, the longest sleep was considered "nighttime sleep". Users were then required to include six weeks of data (12 weekend nights and 30 weekday) resulting in a dataframe with $\sim$ 10.8 million nights. We then observed the distribution of bed-time, wake-time and sleep duration, separately for weekdays and weekends (see Figure 1), which reveal outliers both in terms of sleep timing and duration.

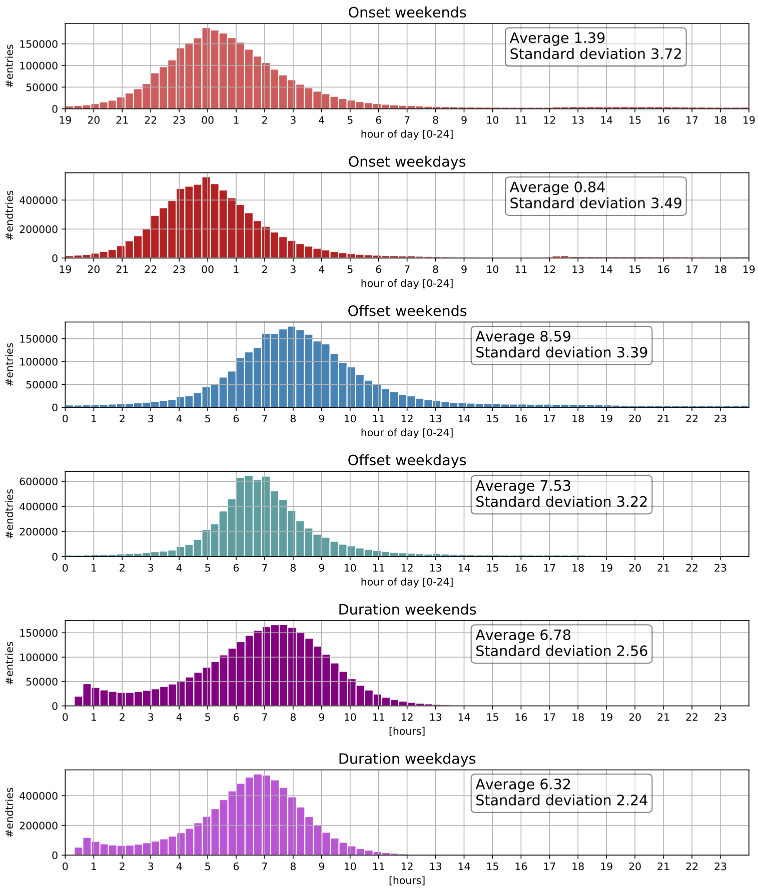


Figure 1: Distribution of sleep onset, offset and duration plotted separately for weekdays and weekends

In order to obtain nighttime sleep (exclude day-time naps and remove outliers) we apply a data-driven filtering. Entries with sleep onset and offset (separately on weekends and weekends) greater than one and a half standard deviations away from the sample average are removed, which eliminates approximate $\sim$ 11% of the entries. The filters were:

- 19:36 $\leq$ onset weekdays $\leq$ 06:05
- 19:49 $\leq$ onset weekends $\leq$ 06:59
- 02:42 $\leq$ offset weekdays $\leq$ 12:21
- 03:30 $\leq$ offset weekends $\leq$ 13:40

If the same approach would be employed for sleep duration, the filters would be [2.96 : 9.67] on weekdays and [2.93 : 10.67]. We rather choose to apply standard filters from the literature (since the both the distributions are oddly shaped at the left tail), where sleep duration is required to be 3 hours or more, but less or equal to 13 hours (3 $\leq$ duration $\leq$ 13). These filters are more inclusive (by 2 hours) than those used by Walch et al. [59] and Althoff et al [74] (4 $<$ duration $<$ 12). After all data filtering has been implemented, the data contains $\sim$ 9.3 million nights. However, for our in-depth analysis we choose to apply an even stricter requirement in terms of number of nights and data coverage. Users need to have 8 weeks (56 days) of data with 70 % cover ratio (meaning the span of their recording must include at least 70 % of the days).

Chronotype & Social jetlag

The Munich Chronotype Questionnaire was created by Roenneberg *et al.* (2003) and poses 17 questions about individual's sleep behavior [12]. Answers are used to estimate corrected mid-sleep on work-free days, or chronotype (MSF_{sc}), calculated as:

$$MSF_{sc} = \begin{cases} SO_{free\,days} + \frac{SD_{free\,days}}{2} & \text{if } SD_{free\,days} \leq SD_{work\,days} \\ MSF - \frac{SD_{free\,days} - SD_{week}}{2} = SO_{free\,days} + \frac{SD_{week}}{2} & \text{if } SD_{free\,days} > SD_{work\,days} \end{cases} \tag{1}$$

Where MSF is midsleep on free days while MSF_{sc} is corrected midsleep on free days. SD refers sleep duration where the subscript 'week' denotes weekly average sleep duration, 'free days' weekend averages and 'work days' weekday averages. SO refers to sleep onset (point in time where people fall asleep).

Wittman et. al. (2006) developed a concept to describe this misalignment between the biological and social clock called *social jetlag*, and is estimated by calculating the difference between midsleep on free days and work days [13].

$$Social\ jetlag = MSF - MSW \tag{2}$$

where MSF denotes midsleep on free days (weekends) and MSW midsleep on work days (weekdays).

Underlying differences across users with same estimates of chronotype, social jetlag and median sleep duration

On Figure 2-4 we illustrate how different metrics of sleep (sleep duration, bed and waketime, weekend-weekday differences, width of behaviors, and the weekend-weekday width differences) distributes across users who have one sleep epidemiological metric approximately the same.

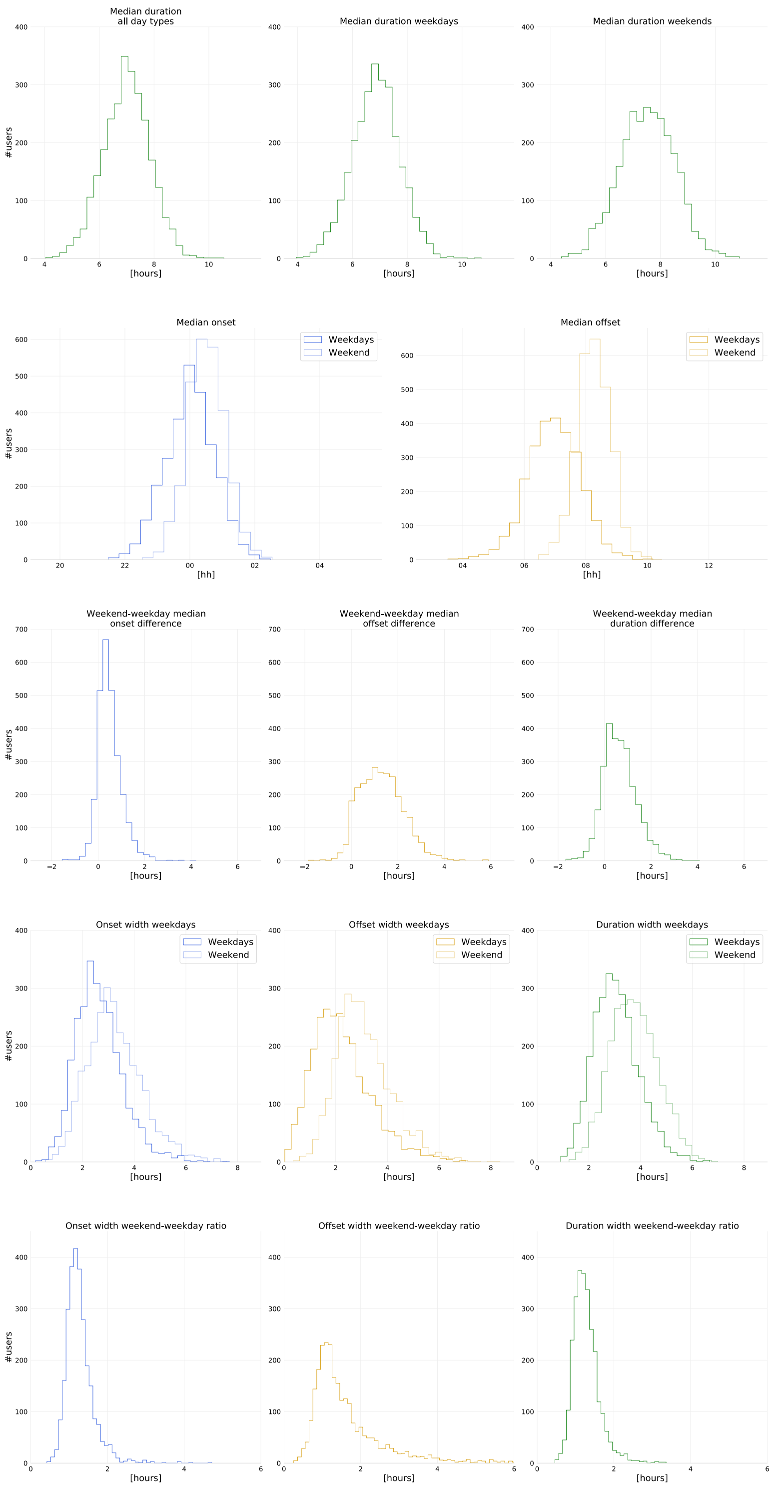


Figure 2: Distribution of different metrics of sleep for users with approximately the same chronotype (∼04:30 and $N_{group} = 2719$)

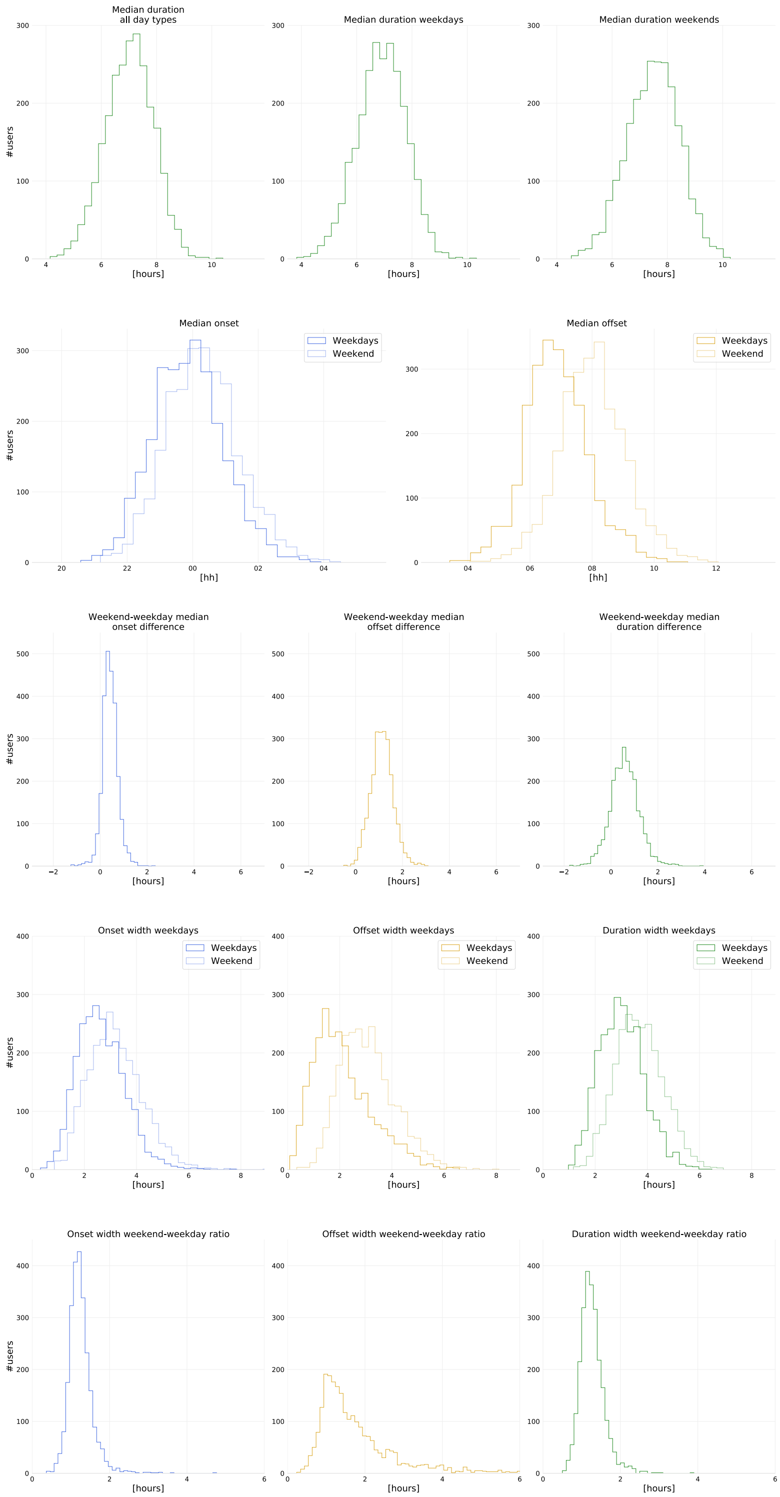


Figure 3: Distribution of different metrics of sleep for users with approximately the same social jetlag (∼45 minutes and $N_{group} = 2479$)

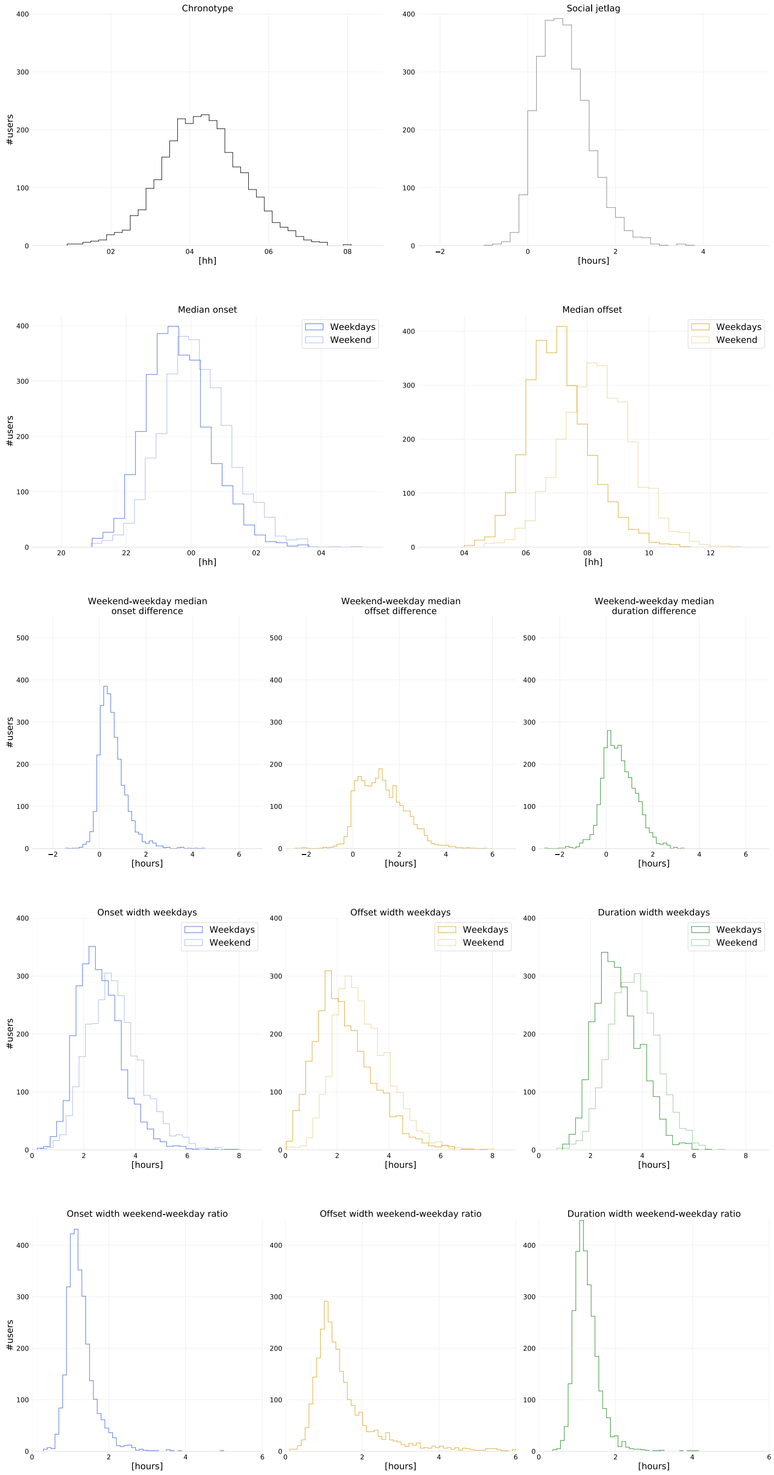


Figure 4: Distribution of different metrics of sleep for users with approximately the same overall median sleep duration (∼7.5 hours and $N_{group} = 2860$)

New measure for sleep regularity

In our analysis we present a new measure to quantify sleep variability. Typically variability is measured as a standard deviation (std) of a distribution for measure, but we rather suggest using the difference between the 90th and 10th percentile. We consider it a more intuitive than the std and provide an example to demonstrate; imagine the scenario where one has to comprehend either "John has 0.74 hours std in wake-up time on weekdays" or "John wakes up 80 % of the time within a span of 15 minutes on weekdays". The two measures correlate nearly perfectly as demonstrated on Figure 5.

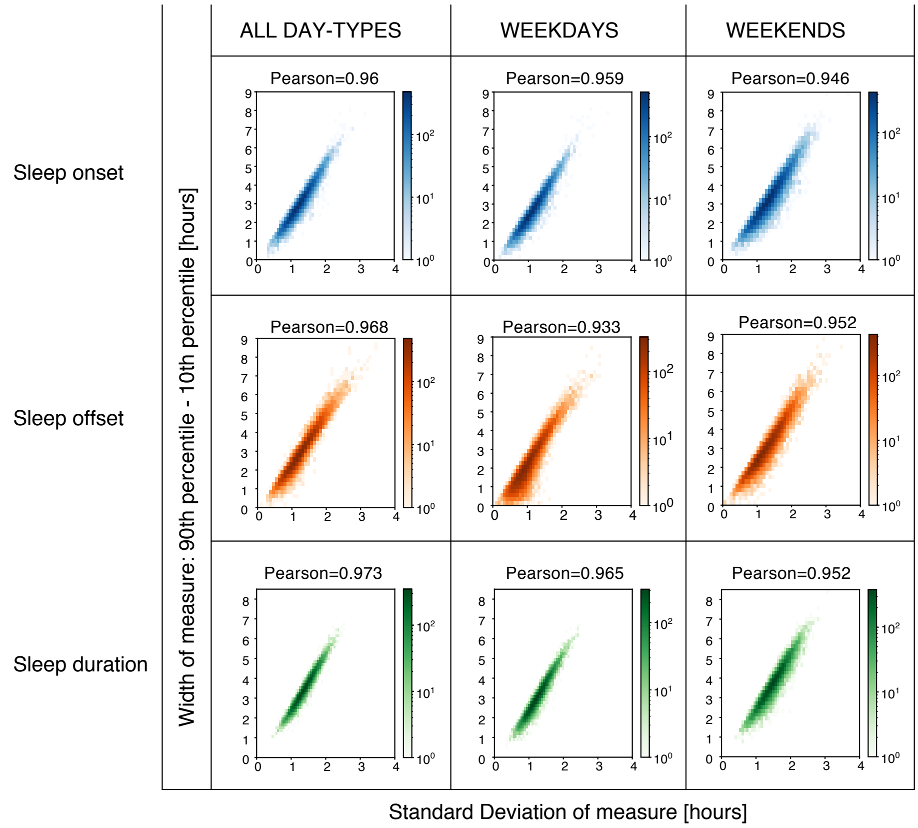


Figure 5: Illustrates how width of a distribution (calculated as the 90th percentile minus 10th percentile) correlates with its standard deviation for sleep onset, offset and duration by all day-types, weekdays and weekends separately

Short or long sleep duration but no skew?

We estimate the probability for a randomly selected individual to belong to different demographic groups (`gender` [female/male], `age group` [19-24/25-29/30-34/35-39/40-44/45-49/50-54/55-59/60-67] and `region of residence` (east=Asia/west=Europe & N-America) in the full data-set. Thereafter, we filter the data by criteria to obtain two dataframes; i) short median sleep duration (<6.5 hrs) with no or negative skew (skew<0.25) and ii) long median sleep duration (>8 hrs) with positive or no skew (>-0.25). For these two different dataframes we look again at the probability of belonging to demographic groups (all listed before) and obtain the relative probability by dividing with the probability of a member of the demographic group in the full data-set.

Overall

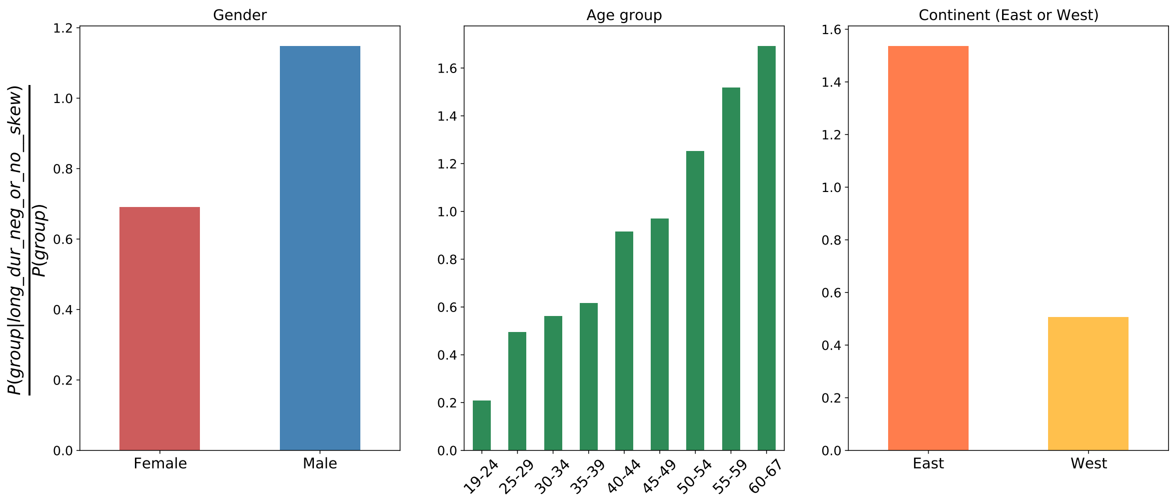


Figure 6: The relative probability that individual belongs to a demographic group in the filtered data with short sleep duration (<6.5 hrs) and no or negative skew (skew<0.25). Skew is estimated for the distribution of sleep duration for all day-types.

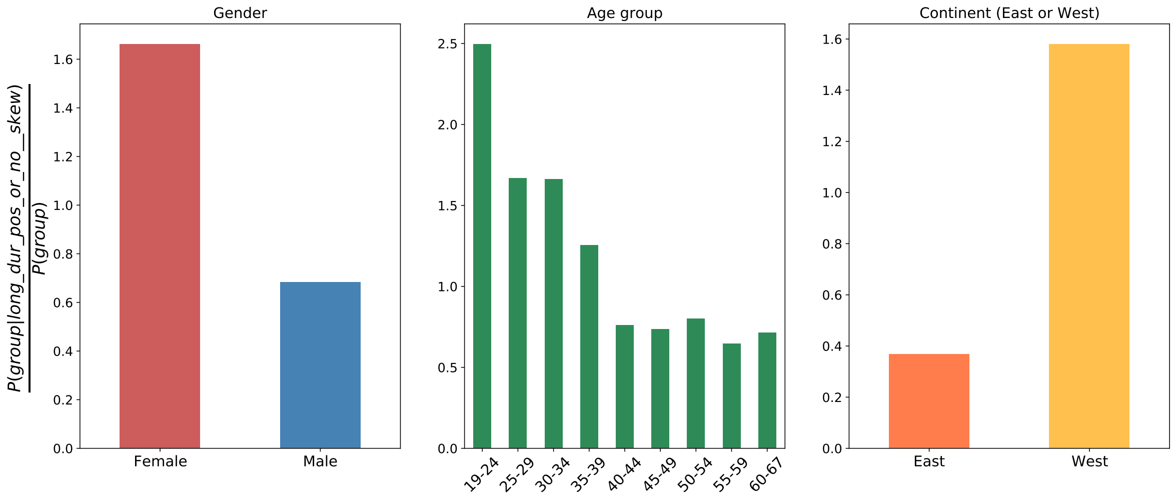


Figure 7: The relative probability that individual belongs to a demographic group in the filtered data with long sleep duration (>8.0 hrs) and positive or no skew (skew>-0.25). Skew is estimated for the distribution of sleep duration for all day-types.

Weekdays

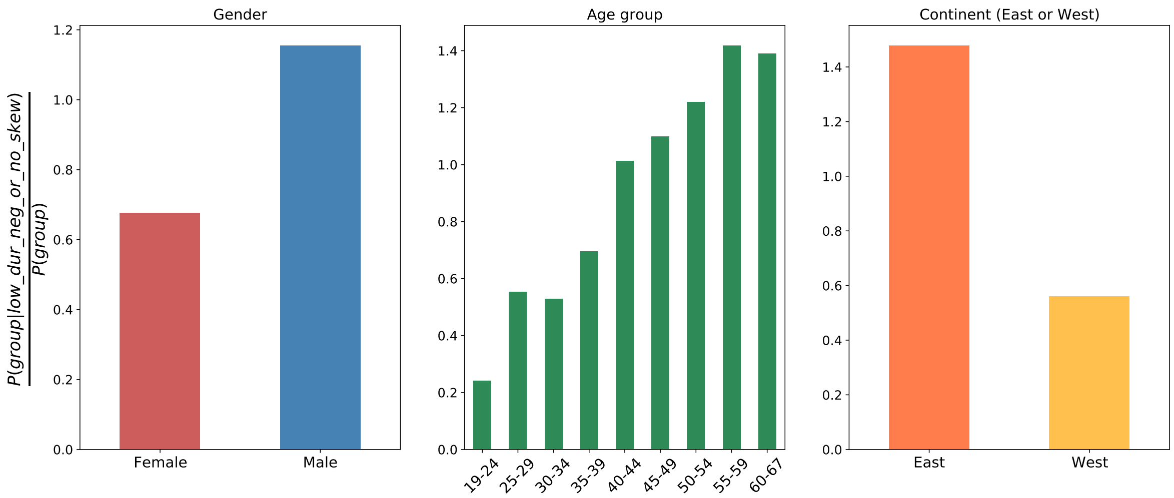


Figure 8: The relative probability that individual belongs to a demographic group in the filtered data with short sleep duration (<6.5 hrs) and no or negative skew (skew<0.25). Skew is estimated for the distribution of sleep duration for weekday-nights.

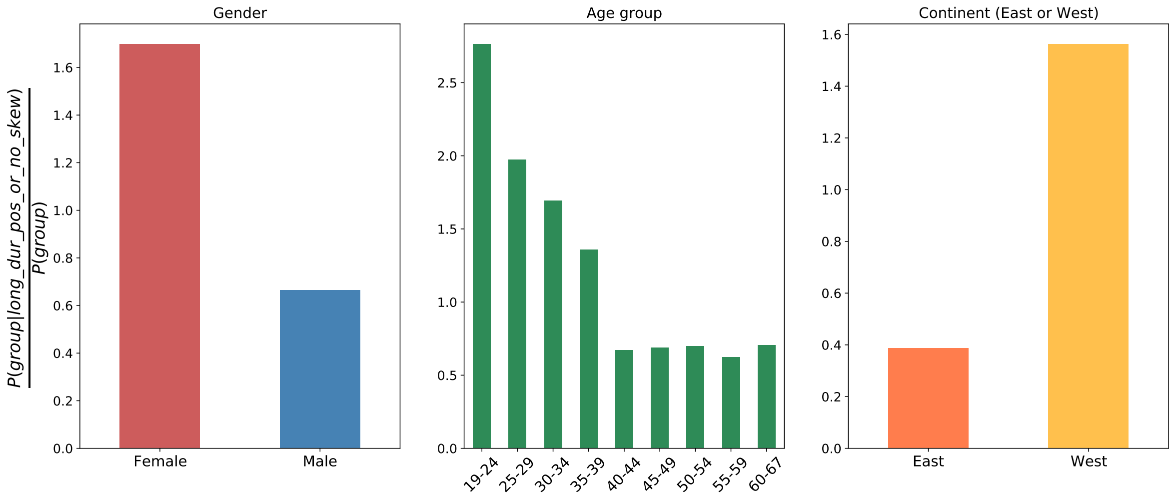


Figure 9: The relative probability that individual belongs to a demographic group in the filtered data with long sleep duration (>8.0 hrs) and positive or no skew (skew>-0.25). Skew is estimated for the distribution of sleep duration for weekday-nights.

Weekends

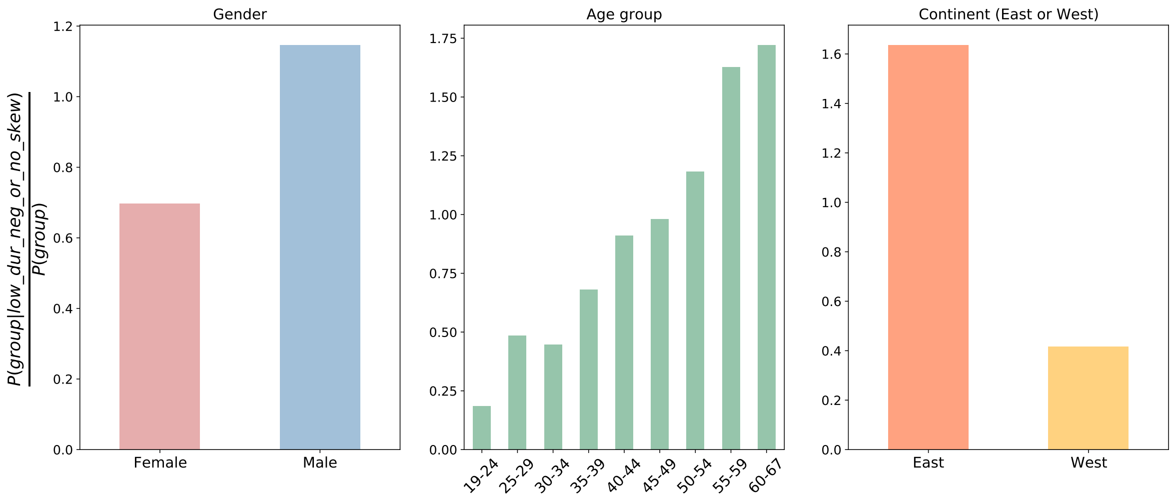


Figure 10: The relative probability that individual belongs to a demographic group in the filtered data with short sleep duration (<6.5 hrs) and negative or no skew (skew<0.25). Skew is estimated for the distribution of sleep duration for weekend-nights

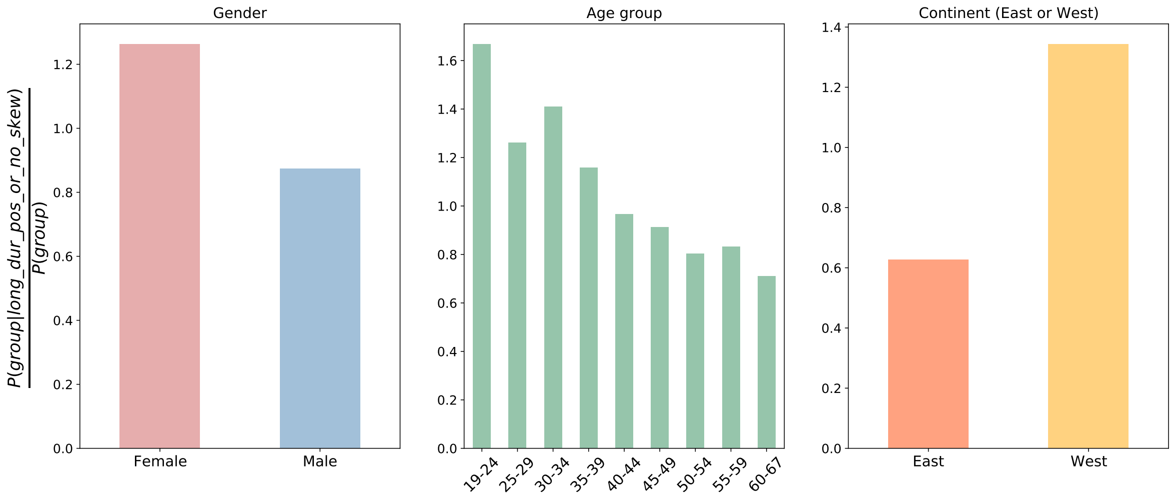


Figure 11: The relative probability that individual belongs to a demographic group in the filtered data with long sleep duration (>8.0 hrs) and positive or no skew (skew>-0.25). Skew is estimated for the distribution of sleep duration for weekend-nights

Skew group characteristics

We plot the distribution of all sleep metrics, by the three skew groups; the 2000 individuals with either the most positive, neutral and negative skewed distribution of sleep duration for all day-types, weekday-nights or weekend-nights respectively Figures 12–14.

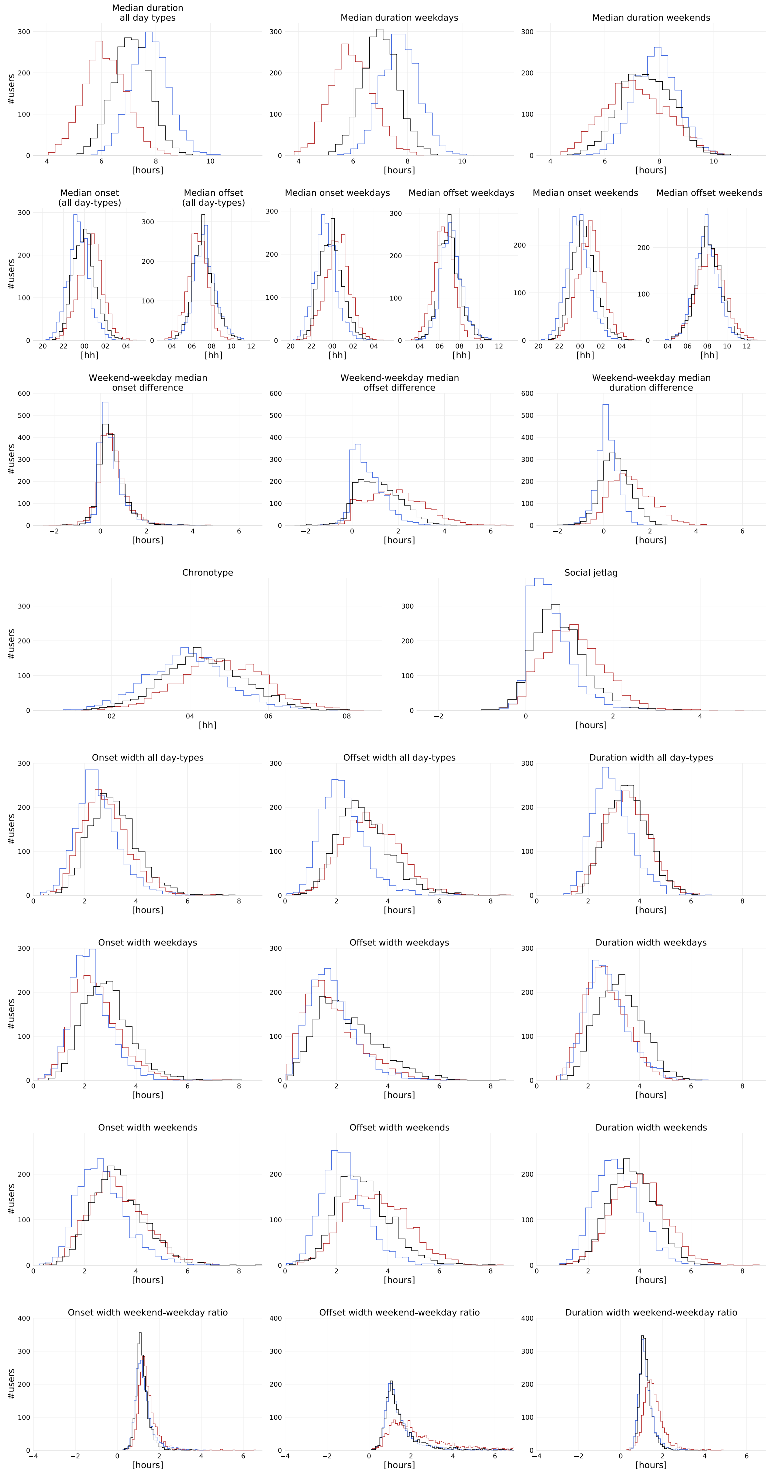


Figure 12: Distributions of multiple features for the 2000 most positively (red), neutrally (black) and negatively (blue) skewed individuals. Skew is estimated for the distribution of sleep duration for all day-types.

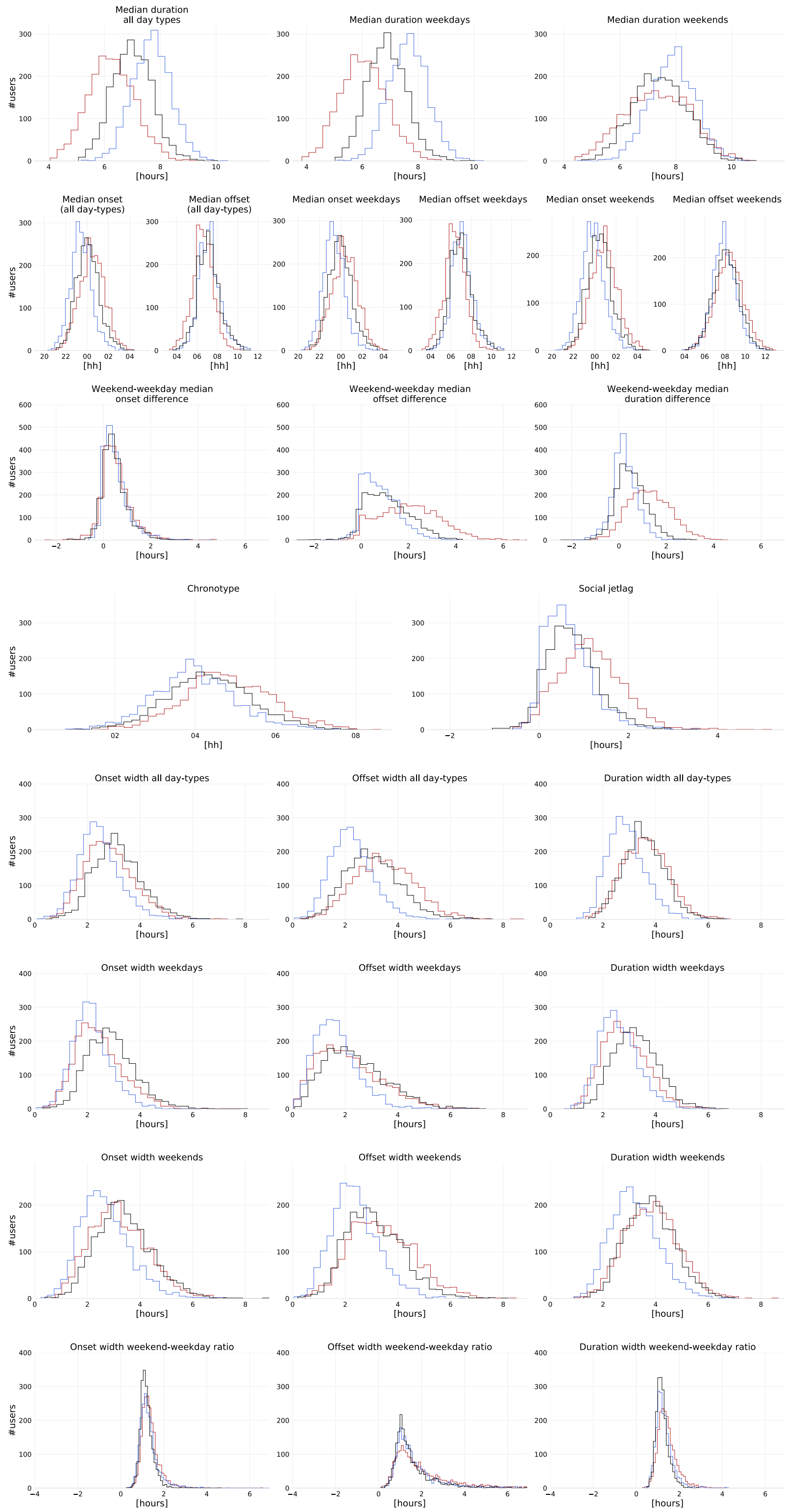


Figure 13: Distributions of multiple features for the 2000 most positively (red), neutrally (black) and negatively (blue) skewed individuals. Skew is estimated for the distribution of sleep duration on weekdays.

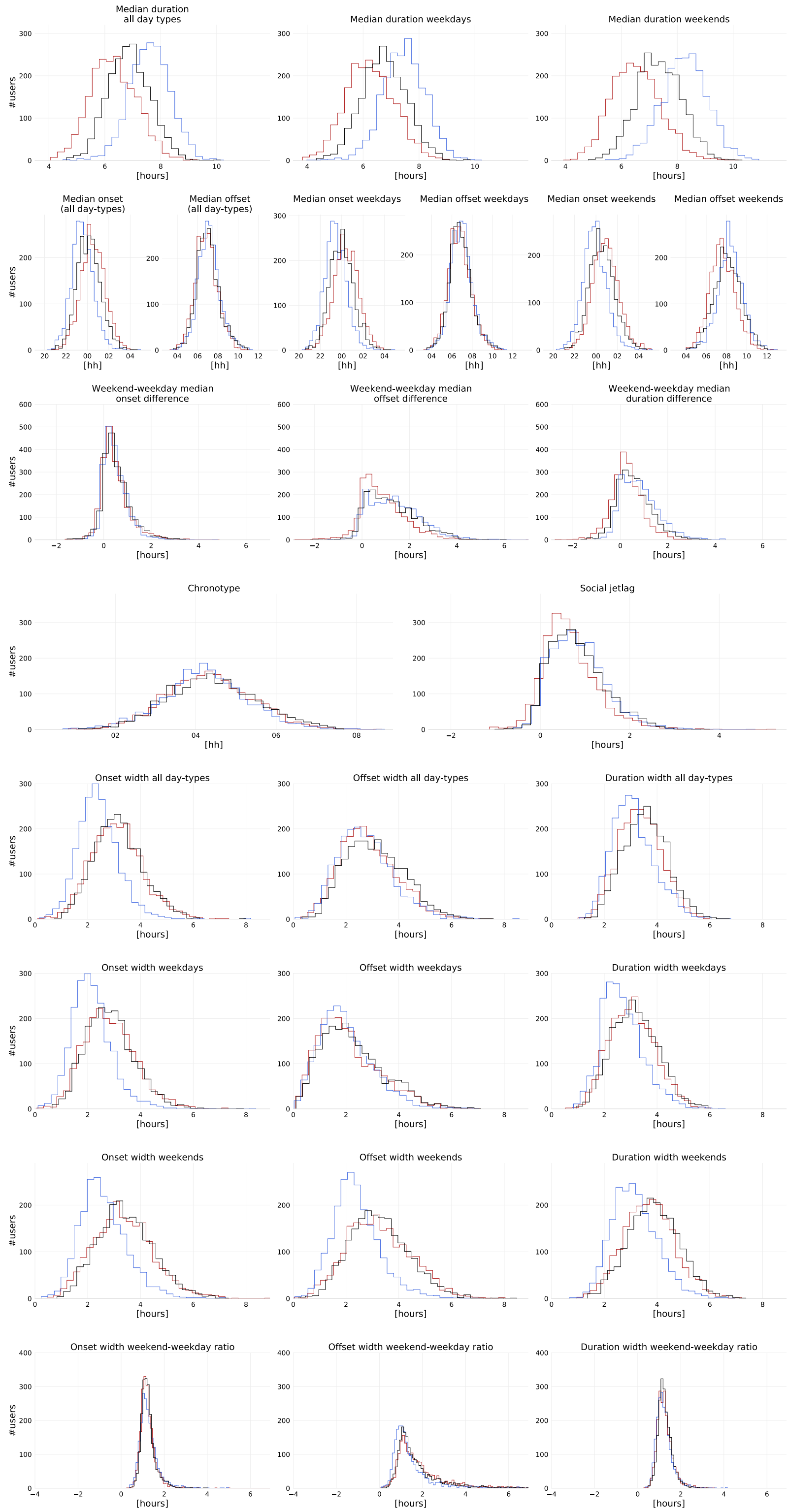


Figure 14: Distributions of multiple features for the 2000 most positively (red), neutrally (black) and negatively (blue) skewed individuals. Skew is estimated for the distribution of sleep duration on weekends.