Rising like a Phoenix: Emerging from the Pandemic and Reshaping Human Endeavors with Digital Technologies ICIS 2023

IS in Healthcare Addressing the needs of post-pandemic digital healthcare

Dec 11th, 12:00 AM

Two Sides of the Same Pillow: Unfolding the Relationship between Objective and Subjective Sleep Quality with Unsupervised Learning

Luka Biedebach
Reykjavik University, lukab@ru.is

María Óskarsdóttir
Reykjavik University, mariaoskars@ru.is

Erna Sif Arnardottir
Reykjavik University, ernasifa@ru.is

Anna Sigríður Islind
Reykjavik University, islind@ru.is

Two Sides of the Same Pillow: Unfolding the Relationship between Objective and Subjective Sleep Quality with Unsupervised Learning

Complete Research Paper

Luka Biedebach
Reykjavik University
Menntavegur 1, 102 Reykjavik, Iceland
lukab@ru.is

María Óskarsdóttir
Reykjavik University
Menntavegur 1, 102 Reykjavik, Iceland
mariaoskars@ru.is

Erna Sif Arnardóttir
Reykjavik University
Menntavegur 1, 102 Reykjavik, Iceland
ernasifa@ru.is

Anna Sigridur Islind
Reykjavik University
Menntavegur 1, 102 Reykjavik, Iceland
islind@ru.is

Abstract

Advances in digital health allow us to take an active part in monitoring and improving our sleep quality. Both, objectively recorded and subjectively perceived sleep quality impacts our general health and well-being. This research shows how these two dimensions of sleep quality can be captured with smartwatches and digital symptom trackers. We contribute to the gap in the literature on how recorded values from wearables and user-generated content from mobile applications can elevate each other. Analysing the recorded and reported sleep quality in a longitudinal sleep study (n=45) shows differences in how participants perceive their sleep. We address this need for personalization, by creating clusters of participants with a similar perception of sleep using unsupervised machine learning. Analysing these clusters provides us with a more wholesome understanding of their sleep quality and raises awareness for the uniqueness of individuals in digital health.

Keywords: sleep quality, unsupervised machine learning, wearables, mobile application, clustering

Introduction

Our well-being depends on good sleep (Luyster et al., 2012), but the notion of sleeping well is still not fully understood (Buysse, 2014). Just like for any objectively and subjectively captured phenomenon, there are two sides of the same coin – or the same pillow – respectively for sleep quality. Bad sleep can have several negative effects on the body including daytime sleepiness, memory impairment (Orzeł-Gryglewska, 2010) and decreased neurobehavioral performance (Belenky et al., 2003). Chronically bad sleep increases the risk of inflammatory diseases (Irwin, 2015), cardiovascular diseases (Hoevenaar-Blom et al., 2011) and obesity (Beccuti & Pannain, 2011). There is no generally accepted definition of what sleep quality entails, or how to capture it. It can be described by characteristics of a person's sleep that can be measured or quantified by parameters such as the total sleep time, the sleep onset latency and sleep efficiency (Krystal & Edinger, 2008). The rise of digital health and the transition of digital health solutions from the clinic to the home allows more people to track both their subjective and objective sleep quality. Objective sleep quality can be derived from tracking the body with sensors during sleep. The gold standard of sleep measurement is polysomnography, which continuously records the activity of the brain, the heart, the muscles and the respiratory system through multiple sensors during the night (Bruyneel et al., 2011). It is to date, seen as the the most reliable form of sleep measurement but it is not feasible for measuring sleep over an extended period of time due to its high effort. Wearable devices such as smartwatches measure the heart rate and movement, to give an estimation of a person's bed and rise times, the number of awakenings during the night and the duration of light sleep and deep sleep (Sadeh, 2011). Even though wearables measure sleep less reliably than a full polysomnography, their strength is to collect longitudinal data, the so-called 'free-living sleep' (Arnardottir et al., 2021), which is highly relevant in sleep research (Óskarsdóttir et al., 2022). However, sleep quality is a multidimensional construct that cannot be explained by sleep parameters only. The way we subjectively perceive our sleep, i.e. if we feel rested or fatigued, is essentially equally or more important than any measurement (Bin, 2016; Hoevenaar-Blom et al., 2011). Walsh et al. (2022) demonstrated the relevance of subjective sleep quality in the example of upper respiratory tract infections. Subjective sleep quality can for instance be measured as a Likert-style rating by the participant (Krystal & Edinger, 2008). On a more granular level, symptom trackers are used to split the rating into questions regarding the ease of falling asleep, the comfort during sleep or the feeling when waking up. This can be done digitally in the form of an entry in a digital symptom tracker application or in an analogue way via a manually filled-in sleep diary (Schmitz et al., 2022). Buysse (2014) introduced the highly relevant concept of sleep health, where objective and subjective sleep quality act combined as a wholesome indicator of sleep health. If we understand the underlying correlations, we can leverage the usage of wearables in combination with digital symptom trackers to objectively and subjectively monitor our health and well-being in general, and sleep in particular (Islind et al., 2022; Vallo Hult et al., 2022).

Digital health platforms and wearables allow patients to take part in the care process more actively and become co-creators of their own health data, which makes more individualized approaches for patient care possible (Topol, 2014). This leads to a shift towards precision medicine, which adjusts the medical treatment to the individual needs of the participants (National Research Council, 2011) and individualized, data-driven healthcare (Lillie et al., 2011). For sleep-disordered patients, the continuous form of health data collection enables personalized remote care by health professionals (Grisot et al., 2019). Another recent development in health information systems is self-monitoring, which enables individuals usually through wearables or symptom trackers to keep track of their symptoms and behaviors, review the self-recorded data and act accordingly. This may include adjusting their behaviour, applying treatment or seeking the help of a professional (Jiang & Cameron, 2020). Common motivations for self-tracking are curiosity, chronic disease or improving personal performance (Baumgart & Wiewiorra, 2016). Self-monitoring of sleep quality can help to increase the individuals awareness of sleep hygiene (Mairs & Mullan, 2015) and adjust their behaviour accordingly (Berryhill et al., 2020). However, there has been no research yet on how to combine objective and subjective measures for self-monitoring in sleep. For this reason, we aim to answer the following research question:

RQ: *How can the information from wearable devices and digital symptom trackers be combined to achieve more individualized sleep analysis?*

By answering this research question we contribute to the field of information systems, individualized healthcare and self-monitoring by showing the need for an individual-level analysis of sleep quality and proposing a method to address this need. We contribute to the field of digital health, especially regarding wearable devices and their health outcomes as well as data-driven approaches for digital subjective symptom trackers. Our main contributions are the analysis of subjective and objective sleep parameters and the proposed method of clustering participants based on their sleep perception for more individualized sleep analysis. The rest of this paper is organized as follows. In the next section, we review the related literature on the relationship between objective and subjective sleep quality and individualized healthcare. We then introduce the longitudinal sleep study in more detail and explain the methodological structure of this paper, including clustering and cluster analysis. We will present the results of the cluster analysis and based on the identified characteristics and correlations define *sleep quality types*. Finally, we will discuss the implications of our results in the context of digital health.

Related Work

Mobile digital health application and wearable devices have become a widespread and accepted way for tracking ones health (Farivar et al., 2020). The pandemic led to an increased speed in the adoption of digital health applications and devices which showed that the future of public health is becoming increasingly digital (Budd et al., 2020). Whether we model the health status of individuals through their continuous flow of data from digital health applications to personalize medical care or aggregate their data to generate new information about populations, there is potential for improving public health through digital health solutions (Kamel Boulos & Zhang, 2021). These digital health solutions could be mobile health applications (i.e., apps) which have been widely used for symptom tracking, implementing behavioural changes and remote care (Ghose et al., 2021). Moreover, wearable devices such as smartwatches have shown success in continuous long-term health monitoring (Dunn et al., 2018) and health education (Sultan, 2015). Additionally, there is existing research on combining the data from mobile apps and wearable devices. Sigurðardóttir et al. (2022) showed that leveraging objective data from smartwatches and subjective data from digital symptom trackers can enhance the clinical decision-making process for healthcare professionals for determining the ebb and flow in symptoms when treating patients with schizophrenia or bipolar disorders. Bremer et al. (2017) used machine learning methods on digital symptom trackers that produce diary data to predict the mood level of patients with depression. Their work showed how this kind of data can contribute to gaining insights on a subjective and therefore hard-to-measure condition with clinical relevance. It also showed how this subjective data can support personalized interventions. Another subjective phenomena discussed in digital health is subjective well-being. Hu et al. (2023) use a fitness health application to analyse the user's subjective well-being. Similar to subjective sleep quality it cannot be tied to an objective measurement, as it represents an individual's emotional responses and their general life satisfaction (Aboelmaged et al., 2021). The existing work in this area shows the potential of digital health for public health and possible solutions for handling the discrepancy between objective and subjective data. However, sleep - an essential prerequisite for health and well being - has not been approached with a combined analysis of objective and subjective longitudinal data as presented in this research. According to Bin (2016) there is a need to analyse effects of sleep quality on both physical and subjective health to understand the contribution of sleep as a whole to public health.

Previous research has attempted to analyse the relationship between objective and subjective sleep quality in general populations and found little correlation between the objective sleep parameters and subjective sleep quality (Baker et al., 1999). Zhang and Zhao (2007) raised the question whether subjective sleep quality relates to any objective sleep parameters, arguing that it is a combination represented by more than one parameter. The most important drivers of subjective sleep quality have been found to be the recorded total sleep time, awakenings, sleep efficiency (Åkerstedt et al., 2016) and reported sleep parameters (Goelema et al., 2019). Åkerstedt et al. (2016) showed that there are varying relationships between objective and subjective sleep quality in different age groups. Kaplan et al. (2017) approached the relationship between objective sleep parameters and subjective sleep quality with machine learning. They divided the population into age groups and then used a supervised classifier to predict the subjective sleep quality and analysed the feature importance. They showed that the correlations between objective sleep parameters and subjective sleep quality change throughout the age groups. On a more granular level, previous research did compare objective sleep parameters with the subjective reporting of sleep parameters. This gap, which is referred to as *sleep perception*, was previously considered as an effect of sleep disorders and medication as shown by Baker et al. (1999). Other researchers, such as Pinto et al. (2009) compared sleep perception of healthy and sleep disordered individuals and concluded that the way individuals perceive their sleep may be a relevant marker for the evolution of disorders and the effectiveness of treatment. Means et al. (2003) show that sleep perception varies amongst individuals and that distinctive subgroups with similar sleep perceptions can be identified with clustering. We aim to improve their approach by using an extended time period of sleep tracking through wearables instead of single nights from a traditional polysomnography recording. Based on these findings, we propose a method which analyses sleep in more homogeneous groups as proposed by Åkerstedt et al. (2016). In contrast to their work, we do not create those groups based on demographics such as age or gender but based on the gap between their objective and subjective parameters as proposed by Means et al. (2003).

Methods

The study was conducted within the research project Sleep Revolution (Arnardottir et al., 2022) at Reykjavik University in Iceland. The data for this research were collected by two different means simultaneously. Participants were asked to i) wear a Withings smartwatch (Issy-les-Moulineaux, France) and ii) fill out a digital sleep diary in the Sleep Revolution app (Reykjavik, Iceland), for 90 consecutive days. Collecting data from heterogeneous data sources is a challenging task. The data from the smartwatch and the digital sleep diary app were transformed into a homogeneous data format and stored in a digital platform as proposed by Sveinbjarnarson et al. (2023). The study is covered by ethical approval of the National Bioethics Committee of Iceland (21-070) and was approved by the Data Protection Agency of Iceland and includes a written consent by each participant. The participants were recruited through online and offline campaigns in the general population and were selected based on their age, BMI, gender and health status. The study aimed to include individuals with a wide range of age and body mass index and have a equal distribution of male and female participants as well as healthy and sleep disordered participants. Only participants who wore the smartwatch and filled out the digital sleep diary app for more than two weeks were selected for this analysis, which included 45 of 63 total participants. The participants wore the smartwatch on average for 75 days and filled out the digital sleep diary app on average for 40 days. This resulted in 2259 nights with information about both objective and subjective sleep quality in total. The population was gender balanced with 53.3% women. The average BMI was 28.2 and the average age was 48.9. There were 5 participants younger than 30, 32 participants between 30 and 60 and there were 8 participants older than 60. Most participants had a high educational level, as 64,4% of them have a university degree, 13,3% of them did vocational training or a technical degree and 22,2% have no degree in higher education. 3 of the 45 participants work in shift work or do night shifts. 80% of them were married or living with a partner and 93.3% were employed or studying. An overview of demographic information about the participants can be seen in Table 1.

Variable	Mean ±Standard Deviation
Age [years]	48.9 ±14.6
Body Mass Index [kg/m^2]	28.2 ±4.7

Table 1. Demographic Information

Objective Data: Smartwatch and Self-Applied Somnography

The smartwatch provides raw measurements, such as heart rate, oxygen saturation as well as aggregated information about the participant's sleep for each night. The aggregated data included the time the participant spent in light sleep and in deep sleep. The REM sleep, the sleep stage in which dreams are experienced, is not captured by this smartwatch even though it accounts for 20% to 25% of sleep (Carskadon, Dement, et al., 2005). The smartwatch also captures the time spent awake during the night the number of awakenings during the night. It measures the sleep onset latency, i.e the duration to fall asleep, and the duration to wake up. Being awake does not necessarily include standing up during the night, it also includes periods of laying in bed awake. Hence, awakenings can be several minutes or even hours, but in most cases only lasts for seconds and may not be remembered in the morning by the individual. The average sleep duration among all participants is 8.2 hours with a standard deviation of 1 hour. Other sleep parameters have been calculated from the available data. This includes sleep efficiency, which is the time asleep during the night divided by the total time spent in bed. Sleep variability was calculated, by taking the standard deviation of the sleep duration of the previous 3 nights. During the night, the heart rate was tracked and aggregated as minimum, maximum and average heart rate. The smartwatch tracked the participants not only during the night but also tracked their activity during the day. It captured the participant's number of steps, distance and elevation. The average distance among all participants is 3.1 km per day with a standard deviation of 1.8. We transformed all time-related features into seconds and brought all features to a uniform scale using the scikit-learn StandardScaler. An overview of all smartwatch variables can be found in Table 2.

Additionally, each participant participated in a self-applied somnography. Somnography is a simplified version of a polysomnography, which is designed for measuring sleep outside of the hospital but with a similar reliability to the traditional measurement (Kainulainen et al., 2021). The measurement included an electroencephalogram (EEG), electrooculogram (EOG), electrocardiogram (ECG), RIP belts for the thorax and abdomen, a finger probe pulse oximeter, a microphone, a nasal cannula, electrodermal activity (EDA) and accelerometry measuring the movement and body position. The somnography provides a more reliable measurement of the participants than the smartwatch but was, due to its high measurement effort, only performed on up to three consecutive nights whereas the smartwatch was worn for the 90-day period. There are 3 participants with no somnography recorded nights, 3 participants with one recorded night, 11 participants with two recorded nights and 28 participants with three recorded nights. All recordings have been manually scored by sleep technologists according to the rules of the American Academy of Sleep Medicine (Berry et al., 2018).

Subjective Data: Digital Sleep Diary from a Mobile Application

The subjective sleep quality was captured with a digital sleep diary in the form of the custom-made app developed by Sleep Revolution. The sleep diary was co-designed and developed, according to the digital sleep diary standards proposed by Schmitz et al. (2022). The design aimed to increase compliance with the digital sleep diary and avoid memory bias. The sleep diary was filled out by the participant two times per day. In the morning diary, the participants described their nocturnal sleep. Additionally, they reported sleep parameters such as the total sleep time, the sleep onset latency, the number of awakenings and the awake time during the night. They were reminded to fill out the digital sleep diary through nudging, delivered by push notifications. In the evening diary, the participants filled in information about their day, in particular about the factors which can impact sleep. This included their stress level, daytime sleepiness, naps, the number of caffeine and alcohol units they consumed and the duration of exercise. This data was used to measure the participants' subjective sleep quality. The average sleep quality rating among all participants was 3.3 out of 5 with a standard deviation of 0.5.

Source	Variables
Smartwatch	Sleep hour, wake-up hour, variability, efficiency, regularity, light sleep, deep sleep, duration to wake up, heart rate (avg, min, max), steps, distance, elevation
Sleep Diary	Exercise duration, work day, stress level, nap count, nap duration, drug use, alcohol count, caffeine count
Both	Sleep duration, awake time, awakenings, sleep onset latency

Table 2. Variables from the Smartwatch and Sleep Diary

Data Analysis

The methods are structured in three parts as can be seen in Figure 1. First, the agreement between recorded and reported sleep parameters was analysed. We calculated the agreement of the objective sleep parameters recorded with the smartwatch and the subjective sleep parameters reported in the sleep diary for each participant. The results from this step described how the individual participants perceived their own sleep. Based on this information, the second step was to cluster the participants into groups with similar sleep perceptions, using the unsupervised learning algorithm K-Means. Lastly, we analysed these clusters of participants with different sleep perceptions. We identified characteristics of the clusters using an analysis of variance. This allowed us to define sleep quality types. Finally, we analysed the correlations between factors that can impact sleep quality in order to understand which areas of improvement are relevant for which sleep quality type.

Comparison of Reported and Recorded Sleep Parameters

There are multiple sleep parameters that were captured both objectively and subjectively. They were recorded by the smartwatch and reported in the digital sleep diary app. This included sleep duration, the sleep onset

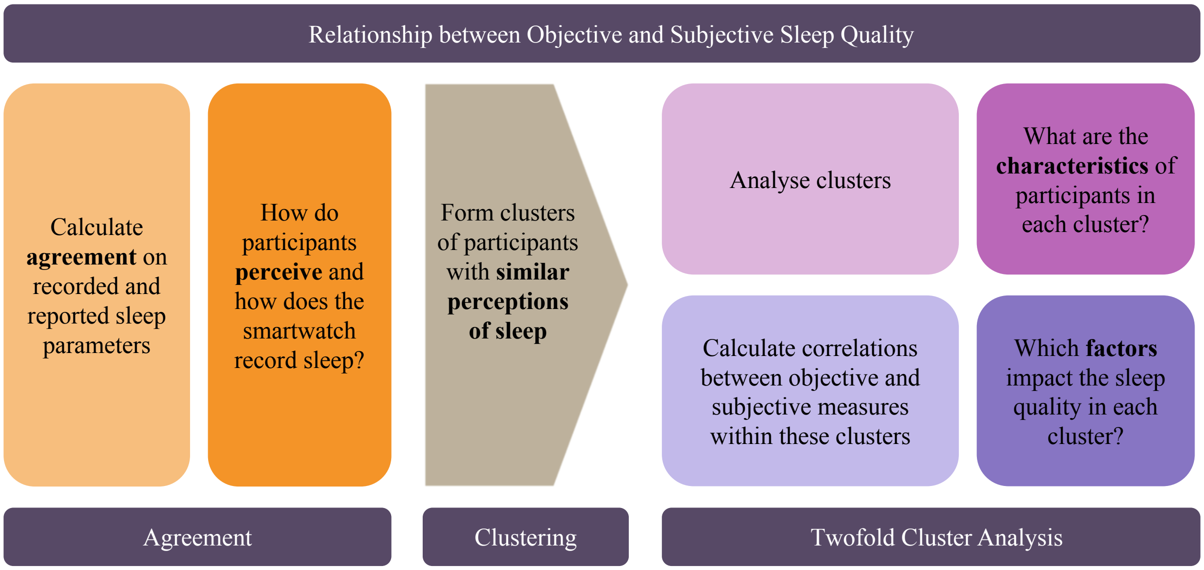

Figure 1. Methodological Structure

latency, the number of awakenings and the awake time after sleep onset as can be seen in Table 2. We directly calculated the agreement of the reported and recorded parameters by taking the absolute difference for each of the four overlapping sleep parameters for each night. This comparison showed us the gap between objective recording and subjective reporting of each participant. It remains unknown whether this gap arises from unreliable recording or reporting, but it indicates how the recorded sleep is connected to the individual perception of sleep of the individual participants. We took neither the smartwatch recording, nor the participants' reporting as ground truth, instead this comparison gave us information about how related or unrelated these two measures generally are. An overview of the availability of both recorded and reported sleep parameters of all participants can be seen in Figure 2. It shows that most participants filled out the sleep diary regularly, with their rating represented as colorful dots. Moreover, it shows that the rating behavior of participants varies, as some participants gave more extreme ratings such as 1 and 5, while others mainly rated medium sleep quality between 2-4.

Additionally, we reviewed the general reliability of the recorded and reported sleep parameters with somnography recordings, which is seen as a more accurate tool for sleep measurement, although not longitudinal. The somnography sleep parameters were manually scored by sleep technologists. The participants had up to three nights of somnography recording, smartwatch recording and digital sleep diary reporting simultaneously. This review gave us important information about the reliability of the recorded and reported sleep parameters, but ultimately the aim of this proposed method is to rely only on recorded data from wearable devices in combination with reported data from a digital symptom tracker.

Clustering

Using unsupervised machine learning, we aimed to identify clusters among participants to find out what impacts the perception of their sleep quality. Clustering allows a more individual-level analysis of the data by creating clusters of participants in a way that participants within a group are more similar to each other than to participants in the other clusters. This strategy can be used to e.g. characterize clinical phenotypes (An et al., 2020). We use the partitional clustering algorithm *K-Means* to identify similar groups of participants. We chose K-means, because it is works well for balanced cluster sizes and small cluster numbers. It furthermore has a low computation time and is easily understandable. The K-means algorithm assigns each observation to the cluster with the closest mean over multiple iterations (MacQueen, 1967). The value of K defines the number of clusters. The optimal value for K is chosen by the steepest descent in an elbow plot, which shows the cluster impurity by the number of clusters. We performed the clustering on the calculated agreement between the recorded and reported sleep parameters and the subjective sleep quality.

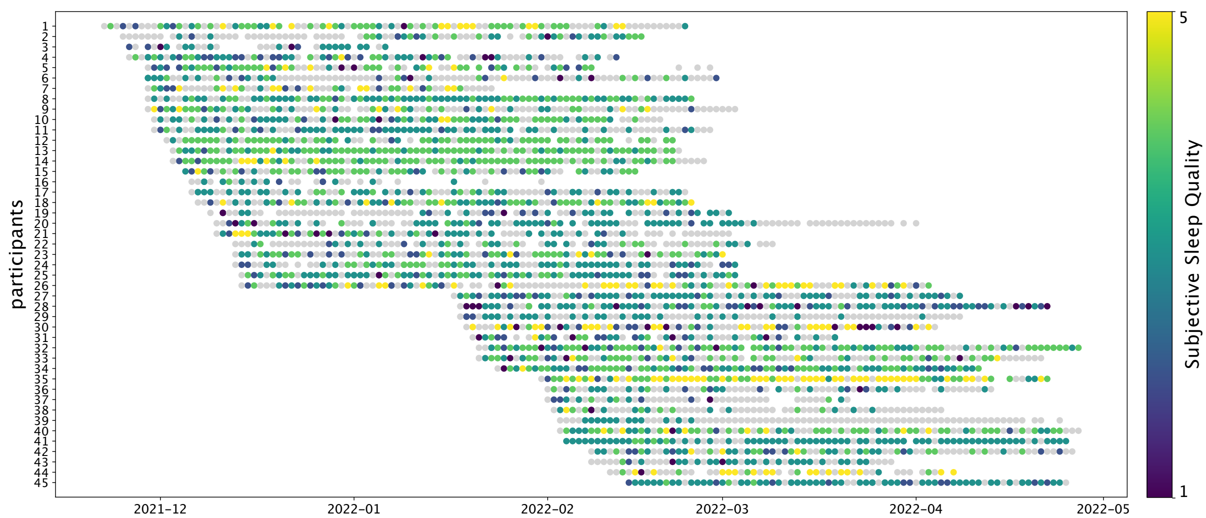

Figure 2. Study Coherence by the Participants over Time (every dot represents a night in which the smartwatch was worn, colored dots describe the subjective sleep quality from the sleep app and grey dots represent recorded nights without a sleep quality rating)

The cluster analysis was twofold, i) we identified common characteristics of the participants within each cluster and ii) we calculated the correlations between the variables and the subjective sleep quality within each cluster. This allowed us to define sleep quality types and improve the understanding of the different factors impacting sleep quality within the clusters. We first analysed the clusters by comparing the median and standard deviation of each cluster. The aim of this analysis was to identify common characteristics of participants within the clusters. Additionally, an analysis of variance (ANOVA) showed, whether the difference between the groups was significant. We performed the ANOVA for each feature to determine the F statistic and p-value. As a second step, we calculated Spearman's rank correlation coefficient between possible impacting factors and the subjective sleep quality over the whole study duration within each cluster. This method resulted in a *rho* value between -1 and 1 for each variable, representing the positive or negative correlation with the subjective sleep quality. Additionally, it resulted in a p-value for each correlation, representing the significance of the correlation. In this analysis, we only considered correlations with a p-value lower than 0.05 as significant.

Results

In the following section, we review the results of our analysis according to the different steps of our method. We first show the gap between the reported and recorded sleep parameters, then introduce the four clusters that were identified in the clustering and then show characteristics of the clusters and review the individual correlations to subjective sleep quality in each cluster.

Sleep Parameter Agreement

In order to understand the relationship between objective and subjective sleep quality, we first compared sleep parameters that were simultaneously captured with the smartwatch, sleep diary and somnography recording for up to three consecutive nights. This comparison showed that there is a varying agreement of objective and subjective sleep quality among the participants. Figure 3 shows the sleep duration of the smartwatch, digital sleep diary app and somnography in 15 exemplary participants. We can see that each participant has a different degree of agreement between the three reported sleep parameters. Considering somnography as the most reliable measurement, we can see that for 30% of the participants the reported sleep duration is more accurate than the recorded sleep duration, while for 70% the recorded sleep duration is more accurate.

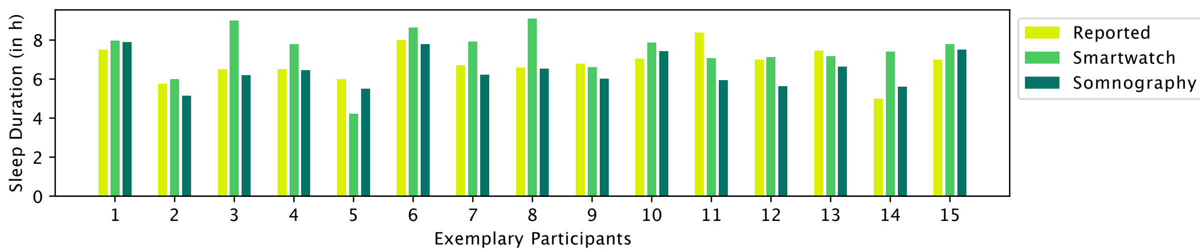

Figure 3. Comparison of Reported Sleep Duration to the Smartwatch and Somnography

Figure 4 shows the difference between the reported and recorded sleep duration. All data points above the vertical line are nights where the participant reported a higher sleep duration than the smartwatch. All nights under the vertical line have a lower reported sleep duration than recorded sleep duration. The color of the data points reflects the subjective sleep quality assigned to the night by the participant. Here, we can see that the subjective sleep quality is usually lower when the reported sleep duration is low regardless of the recorded sleep duration. The reported sleep duration by the participants is on average 57 minutes lower than the recorded sleep duration by the smartwatch. The participant with the highest difference between recorded and reported sleep duration reports on average two hours less than captured by the smartwatch.

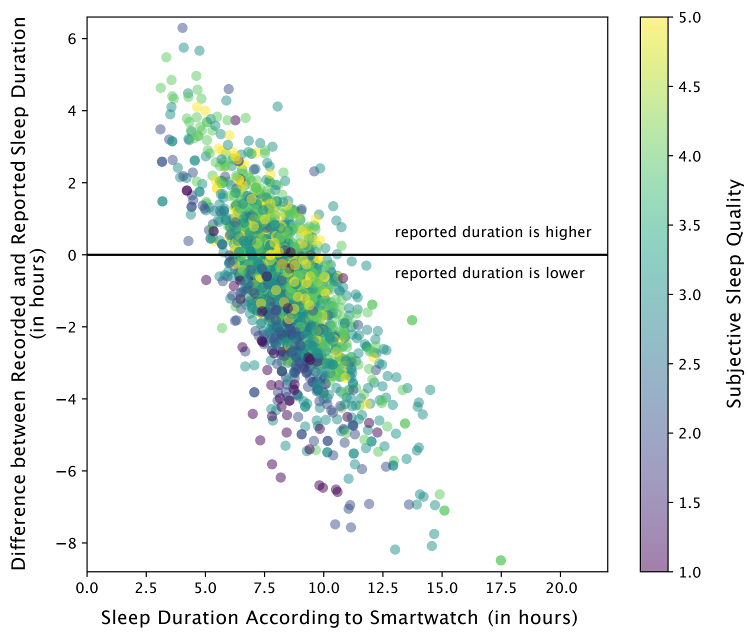

Figure 4. Difference Between Recorded and Reported Sleep Duration

The average number of recorded awakenings during the night is 1.6 with a standard deviation of 1.7. The reported number of awakenings is on average 0.26 lower than the recorded value. An increased number of recorded awakenings leads to a more extreme difference between the recorded and the reported value. Comparing the number of awakenings between the smartwatch and the somnography shows that the recorded value by the smartwatch is usually too high. The sleep onset latency has the proportionally lowest agreement. A comparison of recorded and reported values show that the smartwatch usually records a lower sleep onset latency than the participants report. The average sleep onset latency recorded by the watch is 3 minutes, while the average reported time by the participant is 18 minutes.

Sleep Quality Types

We found the optimal number of clusters by creating an elbow plot, that shows the intra-cluster similarity by the number of clusters. The steepest drop of the curve can be observed at four clusters, which is why we chose K=4. The first cluster has 11, the second cluster 8, the third cluster 11 and the fourth cluster has 15 participants. We use the identified clusters to define sleep quality types based on their agreement, characteristics and correlations. In the 1970s, Horne and Östberg designed a questionnaire to assess an individual's sleep type according to their circadian rhythm (Horne & Östberg, 1976). Based on this research, sleep chronotypes have been developed, which provide personal guidance on the optimal bed and rise times, as well as the optimal time of productivity during the day (Roenneberg et al., 2003). In contrast to the chronotypes, the sleep quality types proposed in this research aim to show influencing factors for the personal perception of sleep quality. The defined sleep quality types can be seen in Figure 5.

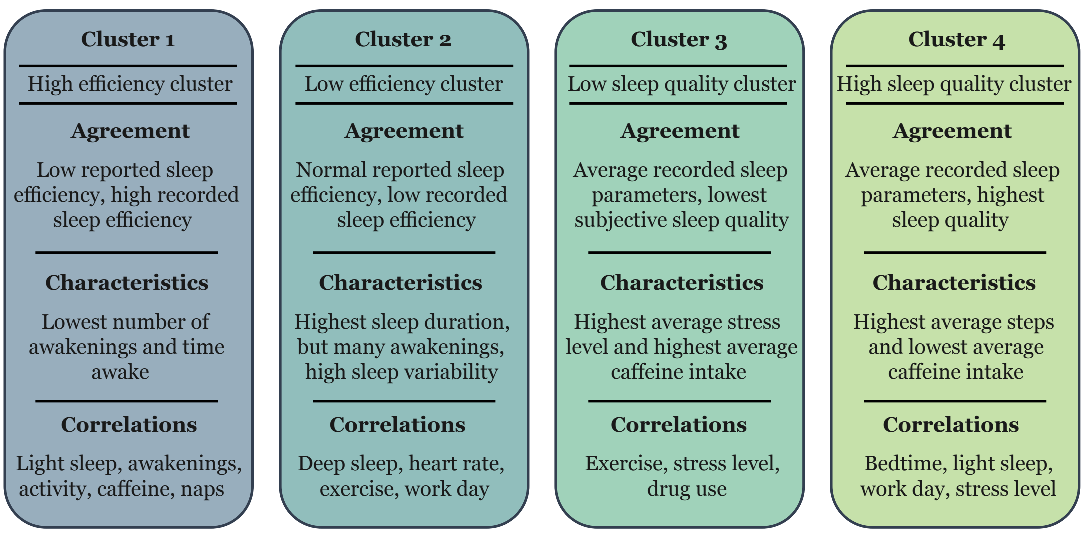

Figure 5. The Four Identified Sleep Quality Types

Identified Characteristics within the Clusters

We use box plots to visually identify characteristics of the different clusters. We show the distributions in each cluster and additionally compare them to the population mean. There is an even distribution of age, body mass index (BMI) and gender throughout the clusters. The main differences we identified between the clusters were found in their relationship between the recorded sleep parameters and sleep quality in general and the relationship between reported and recorded sleep efficiency. Since sleep efficiency describes the sleep duration relative to the total time in bed, a low sleep onset latency and a low duration of time awake during the night contribute to a high sleep efficiency regardless of the total sleep duration. Therefore, sleep efficiency is a combination of all 4 sleep parameters. The sleep quality types in Figure 5 show how the main difference between cluster 1 and 2 are the opposing gaps between reported and recorded sleep efficiency.

Figure 6 shows how the clusters significantly differ in their median subjective sleep quality. Cluster 4 reported on average a higher sleep quality than all other clusters, even though its median sleep duration is close to the population median of 8.2 h. What distinguishes cluster 4 is the relatively high reported sleep duration in comparison to the recorded sleep duration. Cluster 1 has an almost identical distribution of recorded and reported sleep duration but does not show an increased subjective sleep quality.

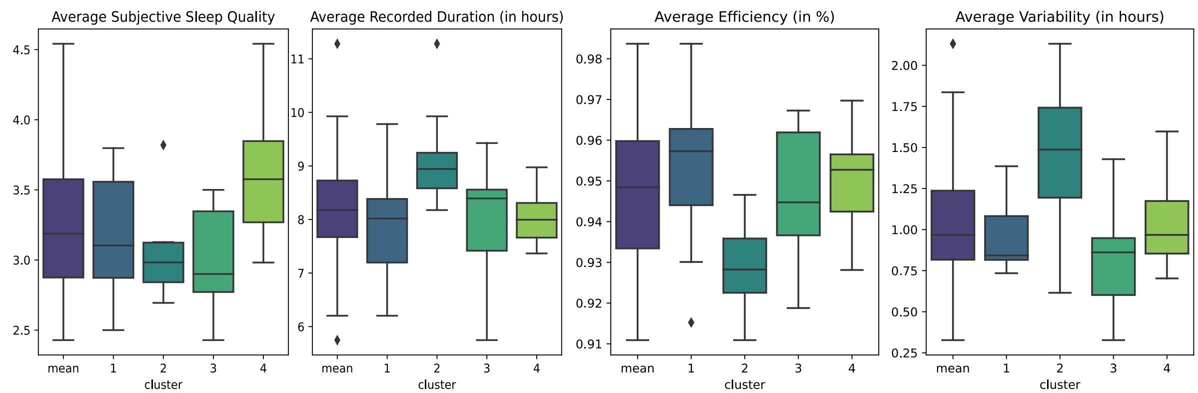

Figure 6. Distribution of Reported Sleep Quality, Recorded Sleep Duration, Recorded Sleep Efficiency and Recorded Sleep Variability among the Clusters

Cluster 1 and two show differences both in the reported and recorded sleep efficiency. Cluster 1 shows a high recorded sleep efficiency, which does not match the high number of reported awakenings and awake time after sleep onset. Contrarily, has cluster 2 the lowest median recorded sleep efficiency, even though the reported sleep parameters indicate a normal sleep efficiency. This cluster has the highest median sleep duration of 9 hours but still has a below-average subjective sleep quality. We conclude that the subjective sleep quality of cluster 2 may be mainly impacted by sleep efficiency.

Avoiding stress, being active, avoiding caffeine after midday and avoiding alcohol can be beneficial for your sleep. Cluster 4, the high subjective sleep quality cluster, shows the highest median recorded activity during the day and the lowest median daily caffeine intake. Cluster 3, the cluster with the lowest sleep median subjective sleep quality, has the highest median stress level and highest median daily caffeine intake. Figure 6 shows that the second cluster has the highest variability of all clusters with a median of 1.5 hours. The participant's medical background gives us a further understanding of the identified characteristics among the clusters. Figure 7 shows the prevalence of moderate and severe insomnia within each cluster according to the insomnia severity index (Morin et al., 2011). The cluster with low sleep efficiency and high sleep variability has a high percentage of insomnia patients, while the cluster with the highest sleep quality has the lowest percentage of insomnia patients.

Identified Correlations within the Clusters

Figure 8 shows the correlations of each variable to the subjective sleep quality for the participants in each cluster. Warm colors indicate a positive correlation and cold colors indicate a negative correlation. Black indicates no correlation between the variable and the subjective sleep quality. We exclude all non-significant correlations with a p-value higher than 0.05 from the analysis and set them to zero. All clusters show the strongest correlation between subjective sleep quality and the reported sleep parameters. These variables have been reported in the same way and at the same time as the subjective sleep quality, which may explain the high correlation. Figure 8 shows that all clusters have different correlations between the given variables and the subjective sleep quality. Cluster 1 shows positive correlations with both sleep duration and wake-up time. This means, that a later bedtime in the morning and a longer total sleep duration are associated with higher subjective sleep quality. Cluster 1 is the only cluster that shows a negative correlation between exercise and subjective sleep quality. Surprisingly, we can see a positive correlation with caffeine intake. Cluster 2 has a negative correlation between subjective sleep quality and heart rate. A high average heart rate during the night is associated with lower sleep quality in this cluster. Additionally, cluster 2 shows the highest positive correlation between exercise duration and subjective sleep quality. Therefore, more exercise is associated with higher subjective sleep quality in this cluster.

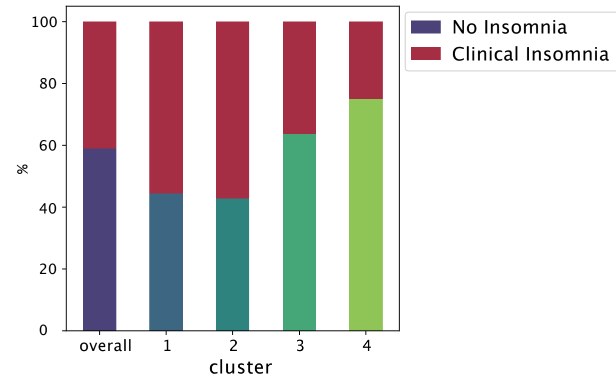

Figure 7. Insomnia Severity among the Clusters

Cluster 3 has almost no correlation with any variable from the smartwatch. The most relevant correlations to subjective sleep quality in this cluster are the four estimated sleep measures. This cluster shows higher correlations of this group of variables than all other clusters. This could indicate that a smartwatch is not a suitable measurement device for this cluster. Cluster 4, the high sleep quality cluster, has a positive correlation between bedtime and subjective sleep quality and has, with 12 PM, the latest median bedtime hour. Furthermore, we can see that a higher percentage of deep sleep correlates with higher subjective sleep quality. Similar to cluster 2, work days and stress are correlated with low subjective sleep quality.

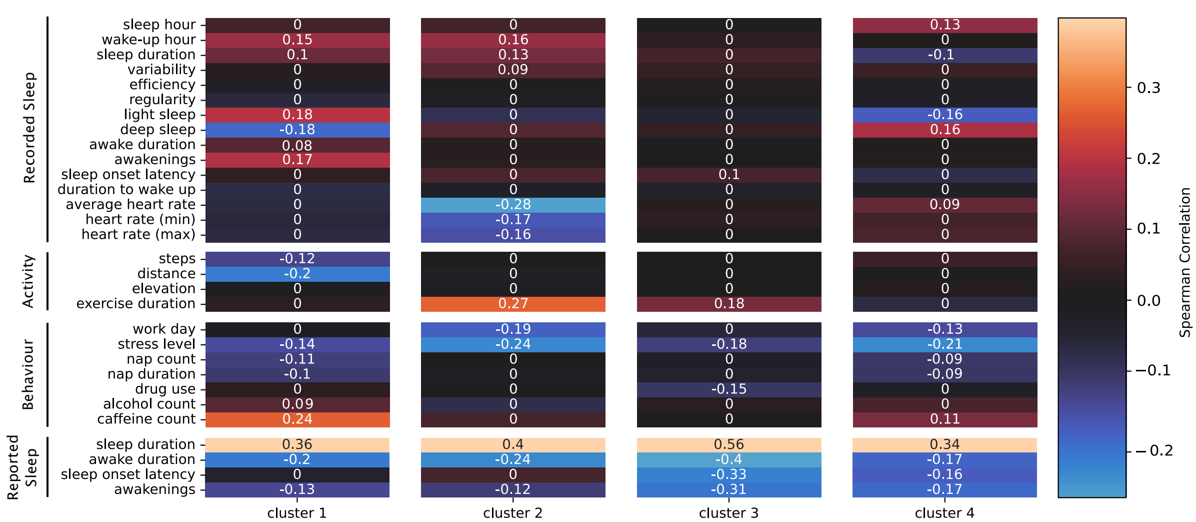

Figure 8. Correlations Between the Objective Attributes and the Subjective Sleep Quality within Each Cluster (non-significant correlations are displayed as zero)

If we take this personalized analysis one step further and look at the data of each participant individually even more extreme correlations become visible. Figure 9 shows a correlation matrix with each row representing one participant and each column representing one variable. The top row represents the full study population, in which only slight correlations are visible. The participants show individual correlations. We can see that most variables both a positive correlation with subjective sleep quality in some participants and a negative correlation in others. It further shows, that some participants have multiple strong correlations with subjective sleep quality while other participants show none. This perspective supports our results on the individuality of the relationship between objective and subjective sleep quality.

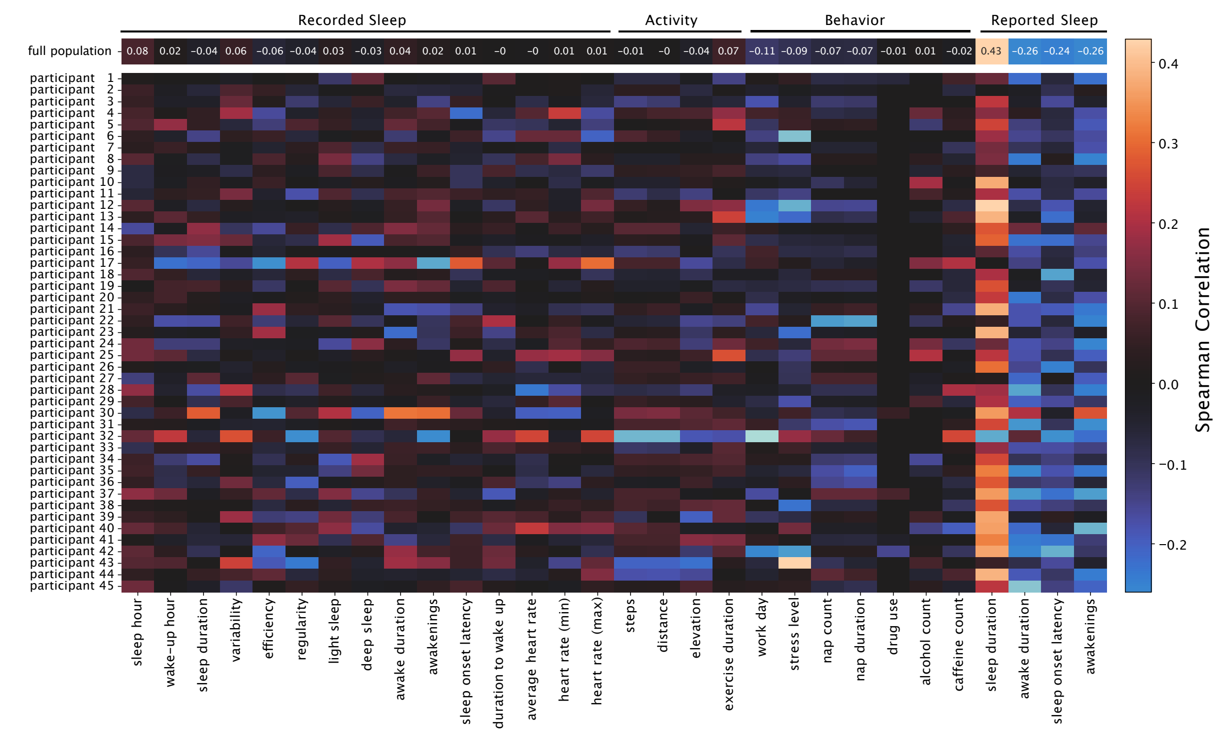

Figure 9. Correlations to Subjective Sleep Quality on the Population vs Individual Level

Discussion

The results presented in this paper showed that there are gaps between the recorded and reported sleep parameters. Moreover, our findings illustrate that some individuals have higher gaps between their reported and recorded sleep duration compared to others. This result could either be explained by measurement difficulties attributed to the smartwatch or to a form of insomnia, as affected individuals have been shown to have a strong discrepancy between objective and subjective sleep duration (Rezaie et al., 2018). The results furthermore show that there are big gaps between the recorded and reported sleep onset latency. When analysing that, we need to consider, that falling asleep is a continuous process where the participant is shifting between levels of consciousness, which makes it challenging both for the participant to give an estimation of the time and for the smartwatch to determine the sleep onset latency. The agreement of the number of awakenings is low across most participants. For all participants, the smartwatch records a higher number of awakenings than reported by the participants. Most awakenings during the night are short and happen during low consciousness, which makes it difficult to remember them. However, a smartwatch has a limited ability to capture awakenings as well (Gruwez et al., 2019). Overall, comparing the sleep parameters showed that both the reported sleep parameters from digital sleep diary app and the smartwatch deviate from the measurements captured by the somnography. This confirms our assumption that neither the recorded nor the reported sleep parameters can be treated as a ground truth. However, for longitudinal sleep measurements, the combination of both, is a viable option.

We showed that the information inherent in these gaps can be used to create clusters of participants with similar sleep perception. The participants within these clusters showed commonalities regarding their median sleep quality, sleep duration, sleep variability and sleep efficiency. This result showed, that even though two clusters show a similar median sleep duration there are still differences in the subjective sleep quality. This is in line with the assumption of Bin (2016), who maintains that more dimensions of sleep quality than sleep duration need to be considered. Analysing these clusters of participants with similar sleep perceptions made correlations between exercise and subjective sleep quality visible. We observed that exercise only has a positive correlation in two clusters. Generally, exercise is beneficial for sleep, except when done shortly before sleep (American Academy of Sleep Medicine, 2005). Even though Hynynen et al. (2010) observe an increased nocturnal heart rate after moderate or heavy exercise during the day, Myllymäki et al. (2012) do not observe a negative impact on subjective sleep quality. As we only observe these correlations in one cluster, this effect might be overlooked when studying the general population. The regularity of sleep does not have a positive or negative correlation in any cluster, even though a positive correlation would be expected. Keeping to a regular sleep duration is beneficial, as high sleep variability has a negative effect on subjective well-being (Lemola et al., 2013) and increases the risk of weight gain (Kobayashi et al., 2013). In two clusters, the caffeine intake shows a positive correlation. This outlines a counterintuitive result, as caffeine is typically considered to have a negative effect on sleep. However, O'Callaghan et al. (2018) describe the complex cyclic relationship between caffeine consumption and sleep deprivation. This correlation might therefore arise from interactions, which are not included in this model, e.g., we do not include the time of consumption or the effects of caffeine withdrawal from *caffeine natives* (O'Callaghan et al., 2018). Finally, when looking at the correlations of each individual participant, the results show an even stronger differences across individuals. This shows that sleep quality is highly individual and based on that, we would like to emphasize the relevance of personalized sleep analysis for future research endeavours.

This paper contributes to the field of information systems in several ways. Firstly, by outlining a method for personalized digital health monitoring utilizing a combination of recorded data from wearable devices and reported data from digital self-tracking applications for the general population. Secondly, we show that the user-generated health data from smartwatches and digital symptom trackers does not necessarily reflect the sleep parameters derived from the somnography, but still gives us valuable information about the participant's sleep health. Thirdly, our paper shows that by combining both objective and subjective measurements and learning about the participant's sleep perception through the agreement between the two, we can develop personalized sleep interventions and go from short term monitoring to reliable, dimensional longitudinal data collection. Improving sleep monitoring with the combination of objective data from wearable devices and subjective data from self-tracking apps allows individuals to actively partake in their own health through technological interfaces, which Petrakaki (2017) refers to as *technological self-care*. Additionally, our paper makes contributions to the field of sleep research by comparing subjective and objective sleep quality in a longitudinal study. To the best of our knowledge, there is no comparable study that captures both the recorded and reported sleep parameters over an extended period of time. It showed that the sleep duration recorded by the smartwatch is on average higher than the reported sleep duration. Our results go in line with the results by Rupp and Balkin (2011) when comparing different wearables to a polysomnography. However, we additionally showed that participants tend to report lower sleep parameters than the smartwatch when they experience low sleep quality. As stated earlier in this paper, the potential of digital symptom trackers is seemingly large. Our research confirms that including information from the digital sleep diary app to the smartwatch measurements enhances sleep analysis. More specifically, we illustrate an added value in including both smartwatch and user-generated health data, since both objective and subjective sleep quality have clinical relevance. Finally, we show, that the interaction of these two dimensions of sleep quality creates an additional value in itself. The agreement between objective and subjective sleep quality contains information about the individual's sleep perception; the combination of the two, outlines two sides of the same pillow.

Practical Implications

The practical implications of our research are two-fold by showing: i) the value of combining objective data and user-generated health data and ii) the need for personalized sleep analysis to cater to the fact that there are individual differences that are important to consider. Current methods of sleep tracking usually either favor subjective or objective sleep data. As a contrast to that, we proposed that both are combined in order to gain an in-depth view of an individual's sleep. This method could be transferred to other areas of information systems and data-driven healthcare for contexts such as chronic disease management, where various types of health data exist, but a coherent integration framework between them is needed (Bardhan et al., 2020). In addition to that, we have illustrated interesting correlations between objective sleep parameters and subjective sleep quality, which may be relevant in clinical practice. We showed that the relationship between objective and subjective sleep quality varies between participants and because of that, we proposed to analyse sleep on in a more personalized manner, through our individual based method. Moreover, our findings show that general assumptions about sleep quality may not apply to all individuals and that issue could be addressed by analysing clusters of individuals with similar sleep perceptions.

Limitations and Future Research

One limitation of this study is the low reliability of the smartwatch, as they can only give estimations of the actual sleep, due to the placement of the sensor on the wrist, and the infancy of the technology to date. Therefore, measurement errors might prevent us from estimating the true sleep parameters. They have shown to be suitable to assess bed and rise times but show a low agreement with awakenings during the night or sleep efficiency (Sadeh, 2011). Kang et al. (2017) showed that smartwatches are less reliable for individuals with sleep insomnia. Furthermore, this work is based on data that is rarely collected over a long period of time. We hope that similar studies can be conducted in the future, as one limitation of the current research is the small number of participants in the study. It limits our ability to observe patterns across the study population. In future research, applying the clustering method to a larger study population may result in more significant characteristics of the clusters. Moreover, this method could be extended by taking the temporal dimension into account and learning from the changes in the participants' sleep over time. Similar to Liang et al. (2016), who proposed a method for calculating individual markers for good sleep quality, future research could do a dynamic analysis of sleep quality based on the previous nights.

Conclusion

This research showed, that the relationship between objective and subjective sleep quality is different for every individual. Comparing the sleep parameters resulting from the smartwatch and the digital sleep diary to the somnography showed that the reliability of both the recorded and reported parameters varied among the participants. We used this difference between reported and recorded sleep parameters to assess the participant's perception of sleep. Clustering participants with similar sleep perceptions allowed us to perform a more personalized analyse of their sleep quality. Based on the identified commonalities in sleep, activity and daytime behavior of these clusters we defined four sleep quality types. All clusters show different correlations between objective and subjective sleep parameters and subjective sleep quality. Hence, objective and subjective sleep quality, are two sides of the same *pillow*, and looking at both sides is vital for assessing sleep quality. Based on that, we propose to combine the data from both smartwatches and digital symptom trackers to outline an individual's perception of sleep. This allows for a more personalized sleep monitoring and ultimately more individual, data-driven patient care in the future.

Acknowledgements

We want to thank everyone who was involved in the set-up of the study. A special thanks to Dimitri Ferretti, Elena Richert, Lisa Schmitz, Bjarki Freyr Sveinbjarnarson, Emil Harðarson, Kolfinna Þórisdóttir, Jacopo Piccini and Katrín Ýr Friðgeirsdóttir. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 965417. The corresponding author is Luka Biedebach and the senior author of this paper is Anna Sigríður Islind.

References

Aboelmaged, M., Hashem, G., & Mouakket, S. (2021). Predicting subjective well-being among mHealth users: a readiness–value model. *International Journal of Information Management* (56), 102247.

Åkerstedt, T., Schwarz, J., Gruber, G., Lindberg, E., & Theorell-Haglöw, J. (2016). The relation between polysomnography and subjective sleep and its dependence on age–poor sleep may become good sleep. *Journal of Sleep Research* (25:5), 565–570.

American Academy of Sleep Medicine (2005). International classification of sleep disorders. *Diagnostic and coding manual*, 51–55.

An, H.-J., Baek, S.-H., Kim, S.-W., Kim, S.-J., & Park, Y.-G. (2020). Clustering-based characterization of clinical phenotypes in obstructive sleep apnoea using severity, obesity, and craniofacial pattern. *European journal of orthodontics* (42:1), 93–100.

Arnardottir, E. S., Islind, A. S., Óskarsdóttir, M., Ólafsdottir, K. A., August, E., Jónasdóttir, L., Hrubos-Strøm, H., Saavedra, J. M., Grote, L., Hedner, J., et al. (2022). The Sleep Revolution project: the concept and objectives. *Journal of sleep research* (31:4), e13630.

Arnardottir, E. S., Islind, A. S., & Óskarsdóttir, M. (2021). The future of sleep measurements: a review and perspective. *Sleep medicine clinics* (16:3), 447–464.

Baker, F. C., Maloney, S., & Driver, H. S. (1999). A comparison of subjective estimates of sleep with objective polysomnographic data in healthy men and women. *Journal of psychosomatic research* (47:4), 335–341.

Bardhan, I., Chen, H., & Karahanna, E. (2020). Connecting systems, data, and people: A multidisciplinary research roadmap for chronic disease management. *MIS Quarterly* (44:1), 185–200.

Baumgart, R. & Wiewiorra, L. (2016). The Role of Self-Control in Self-Tracking. *Proceedings of the International Conference on Information Systems (ICIS), Dublin, Ireland, December 11-14, 2016*.

Beccuti, G. & Pannain, S. (2011). Sleep and obesity. *Current opinion in clinical nutrition and metabolic care* (14:4), 402.

Belenky, G., Wesensten, N. J., Thorne, D. R., Thomas, M. L., Sing, H. C., Redmond, D. P., Russo, M. B., & Balkin, T. J. (2003). Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: A sleep dose-response study. *Journal of Sleep Research* (12:1), 1–12.

Berry, R. B., Brooks, R., Gamaldo, C. E., Harding, S. M., Lloyd, R. M., Quan, S. F., Troester, M. M., & Vaughn, B. V. (2018). *AASM Manual for the Scoring of Sleep and Associated Events*. Tech. rep. Amer. Acad. Sleep Med., Darien, IL, USA: American Academy of Sleep Medicine, Version 2.5.

Berryhill, S., Christopher, J., Dean, A., Provencio-Dean, N., Patel, S. I., Estep, L., Combs, D., Mashaqi, S., Gerald, L. B., et al. (2020). Effect of wearables on sleep in healthy individuals: a randomized crossover trial and validation study. *Journal of Clinical Sleep Medicine* (16:5), 775–783.

Bin, Y. S. (2016). Is sleep quality more important than sleep duration for public health? *Sleep* (39:9), 1629–1630.

Bremer, V., Becker, D., Funk, B., & Lehr, D. (2017). Predicting the individual mood level based on diary data. *Proceedings of the 25th European Conference on Information Systems (ECIS), Guimarães, Portugal, June 5-10*, 1161–1177.

Bruyneel, M., Sanida, C., Art, G., Libert, W., Cuvelier, L., Paesmans, M., Sergysels, R., & Ninane, V. (2011). Sleep efficiency during sleep studies: results of a prospective study comparing home-based and in-hospital polysomnography. *Journal of sleep research* (20:1pt2), 201–206.

Budd, J., Miller, B. S., Manning, E. M., Lampos, V., Zhuang, M., Edelstein, M., Rees, G., Emery, V. C., Stevens, M. M., Keegan, N., et al. (2020). Digital technologies in the public-health response to COVID-19. *Nature medicine* (26:8), 1183–1192.

Buysse, D. J. (2014). Sleep health: can we define it? Does it matter? *Sleep* (37:1), 9–17.

Carskadon, M. A., Dement, W. C., et al. (2005). Normal human sleep: an overview. *Principles and practice of sleep medicine* (4:1), 13–23.

Dunn, J., Runge, R., & Snyder, M. (2018). Wearables and the medical revolution. *Personalized medicine* (15:5), 429–448.

Farivar, S., Abouzahra, M., & Ghasemaghaei, M. (2020). Wearable device adoption among older adults: A mixed-methods study. *International Journal of Information Management* (55), 102209.

Ghose, A., Guo, X., Li, B., & Dang, Y. (2021). Empowering Patients Using Smart Mobile Health Platforms: Evidence from a Randomized Field Experiment. *Forthcoming at MIS Quarterly, NYU Stern School of Business Forthcoming*.

Goelema, M., Regis, M., Haakma, R., Van Den Heuvel, E., Markopoulos, P., & Overeem, S. (2019). Determinants of perceived sleep quality in normal sleepers. *Behavioral sleep medicine*, 388–397.

Grisot, M., Moltubakk Kempton, A., Hagen, L., & Aanestad, M. (2019). Data-work for personalized care: Examining nurses' practices in remote monitoring of chronic patients. *Health informatics journal* (25:3), 608–616.

Gruwez, A., Bruyneel, A.-V., & Bruyneel, M. (2019). The validity of two commercially-available sleep trackers and actigraphy for assessment of sleep parameters in obstructive sleep apnea patients. *PLoS One* (14:1), e0210569.

Hoevenaar-Blom, M. P., Spijkerman, A. M., Kromhout, D., van den Berg, J. F., & Verschuren, W. (2011). Sleep duration and sleep quality in relation to 12-year cardiovascular disease incidence: the MORGEN study. *Sleep* (34:11), 1487–1492.

Horne, J. A. & Östberg, O. (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International journal of chronobiology*.

Hu, J., He, W., Zhang, J., & Song, J. (2023). Examining the impacts of fitness app features on user well-being. *Information & Management* (60:5), 103796.

Hynynen, E., Vesterinen, V., Rusko, H., & Nummela, A. (2010). Effects of moderate and heavy endurance exercise on nocturnal HRV. *International Journal of Sports Medicine* (31:06), 428–432.

Irwin, M. R. (2015). Why sleep is important for health: a psychoneuroimmunology perspective. *Annual Review of Psychology* (66), 143–172.

Islind, A. S., Hult, H. V., Rydenman, K., & Wekell, P. (2022). Co-creating a Digital Symptom Tracker: An App as a Boundary Object in the Context of Pediatric Care. *International Working Conference on Transfer and Diffusion of IT,* Springer, 79–93.

Jiang, J. & Cameron, A.-F. (2020). IT-Enabled Self-Monitoring for Chronic Disease Self-Management: An Interdisciplinary Review. *MIS quarterly* (44:1).

Kainulainen, S., Korkalainen, H., Sigurðardóttir, S., Myllymaa, S., Serwatko, M., Sigurðardóttir, S. Þ., Clausen, M., Leppänen, T., & Arnardóttir, E. S. (2021). Comparison of EEG Signal Characteristics Between Polysomnography and Self Applied Somnography Setup in a Pediatric Cohort. *IEEE Access* (9), 110916–110926.

Kamel Boulos, M. N. & Zhang, P. (2021). Digital twins: from personalised medicine to precision public health. *Journal of personalized medicine* (11:8), 745.

Kang, S.-G., Kang, J. M., Ko, K.-P., Park, S.-C., Mariani, S., & Weng, J. (2017). Validity of a commercial wearable sleep tracker in adult insomnia disorder patients and good sleepers. *Journal of Psychosomatic Research* (97), 38–44.

Kaplan, K. A., Hardas, P. P., Redline, S., Zeitzer, J. M., Group, S. H. H. S. R., et al. (2017). Correlates of sleep quality in midlife and beyond: a machine learning analysis. *Sleep medicine* (34), 162–167.

Kobayashi, D., Takahashi, O., Shimbo, T., Okubo, T., Arioka, H., & Fukui, T. (2013). High sleep duration variability is an independent risk factor for weight gain. *Sleep and Breathing* (17:1), 167–172.

Krystal, A. D. & Edinger, J. D. (2008). Measuring sleep quality. *Sleep medicine* (9), S10–S17.

Lemola, S., Ledermann, T., & Friedman, E. M. (2013). Variability of sleep duration is related to subjective sleep quality and subjective well-being: an actigraphy study. *PloS one* (8:8), e71292.

Liang, Z., Martell, M. A. C., & Nishimura, T. (2016). A personalized approach for detecting unusual sleep from time series sleep-tracking data. in *2016 IEEE International Conference on Healthcare Informatics (ICHI),* IEEE, 18–23.

Lillie, E. O., Patay, B., Diamant, J., Issell, B., Topol, E. J., & Schork, N. J. (2011). The n-of-1 clinical trial: the ultimate strategy for individualizing medicine? *Personalized medicine* (8:2), 161–173.

Luyster, F. S., Strollo, P. J., Zee, P. C., & Walsh, J. K. (2012). Sleep: a health imperative. *Sleep* (35:6), 727–734.

MacQueen, J. (1967). Classification and analysis of multivariate observations. in *5th Berkeley Symp. Math. Statist. Probability,* University of California Los Angeles LA USA, 281–297.

Mairs, L. & Mullan, B. (2015). Self-monitoring vs. implementation intentions: a comparison of behaviour change techniques to improve sleep hygiene and sleep outcomes in students. *International journal of behavioral medicine* (22), 635–644.

Means, M. K., Edinger, J. D., Glenn, D. M., & Fins, A. I. (2003). Accuracy of sleep perceptions among insomnia sufferers and normal sleepers. *Sleep medicine* (4:4), 285–296.

Morin, C. M., Belleville, G., Bélanger, L., & Ivers, H. (2011). The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. *Sleep* (34:5), 601–608.

Myllymäki, T., Rusko, H., Syväoja, H., Juuti, T., Kinnunen, M.-L., & Kyröläinen, H. (2012). Effects of exercise intensity and duration on nocturnal heart rate variability and sleep quality. *European journal of applied physiology* (112:3), 801–809.

National Research Council (2011). *Toward precision medicine: building a knowledge network for biomedical research and a new taxonomy of disease*.

O'Callaghan, F., Muurlink, O., & Reid, N. (2018). Effects of caffeine on sleep quality and daytime functioning. *Risk management and healthcare policy* (11), 263.

Orzeł-Gryglewska, J. (2010). Consequences of sleep deprivation. *International journal of occupational medicine and environmental health*.

Óskarsdóttir, M., Islind, A. S., August, E., Arnardóttir, E. S., Patou, F., Maier, A. M., et al. (2022). Importance of Getting Enough Sleep and Daily Activity Data to Assess Variability: Longitudinal Observational Study. *JMIR Formative Research* (6:2), e31807.

Petrakaki, D. (2017). Producing communal health through self care: the emergence of digital patient activism. *Proceedings of the 25th European Conference on Information Systems (ECIS), Guimarães, Portugal, June 5-10*, 815–827.

Pinto, L. R., Pinto, M. C. R., Goulart, L. I., Truksinas, E., Rossi, M. V., Morin, C. M., & Tufik, S. (2009). Sleep perception in insomniacs, sleep-disordered breathing patients, and healthy volunteers–an important biologic parameter of sleep. *Sleep Medicine* (10:8), 865–868.

Rezaie, L., Fobian, A. D., McCall, W. V., & Khazaie, H. (2018). Paradoxical insomnia and subjective–objective sleep discrepancy: A review. *Sleep medicine reviews* (40), 196–202.

Roenneberg, T., Wirz-Justice, A., & Merrow, M. (2003). Life between clocks: daily temporal patterns of human chronotypes. *Journal of biological rhythms* (18:1), 80–90.

Rupp, T. L. & Balkin, T. J. (2011). Comparison of Motionlogger Watch and Actiwatch actigraphs to polysomnography for sleep/wake estimation in healthy young adults. *Behavior research methods* (43), 1152–1160.

Sadeh, A. (2011). The role and validity of actigraphy in sleep medicine: an update. *Sleep Medicine Reviews* (15:4), 259–267.

Schmitz, L., Sveinbjarnarson, B. F., Gunnarsson, G. N., Davidsson, Ó. A., Davidsson, Þ. B., Arnardottir, E. S., Óskarsdóttir, M., & Islind, A. S. (2022). Towards a Digital Sleep Diary Standard. *Proceedings of the Americas Conference on Information Systems (AMCIS), Minneappolis, August 9-13*.

Sigurðardóttir, S. G., Islind, A. S., & Óskarsdóttir, M. (2022). "Collecting Data from a Mobile App and a Smartwatch Supports Treatment of Schizophrenia and Bipolar Disorder," in *Challenges of Trustable AI and Added-Value on Health,* IOS Press, pp. 239–243.

Sultan, N. (2015). Reflective thoughts on the potential and challenges of wearable technology for healthcare provision and medical education. *International Journal of Information Management* (35:5).

Sveinbjarnarson, B. F., Schmitz, L., Arnardottir, E. S., & Islind, A. S. (2023). The Sleep Revolution Platform: a Dynamic Data Source Pipeline and Digital Platform Architecture for Complex Sleep Data. *Current Sleep Medicine Reports*, 1–10.

Topol, E. J. (2014). Individualized medicine from prewomb to tomb. *Cell* (157:1), 241–253.

Vallo Hult, H., Islind, A. S., Rydenman, K., & Hällsjö Wekell, P. (2022). "Decreased Memory Bias via a Mobile Application: A Symptom Tracker to Monitor Children's Periodic Fever," in *Challenges of Trustable AI and Added-Value on Health,* IOS Press, pp. 915–919.

Walsh, N. P., Kashi, D. S., Edwards, J. P., Richmond, C., Oliver, S. J., Roberts, R., Izard, R. M., Jackson, S., & Greeves, J. P. (2022). Good perceived sleep quality protects against the raised risk of respiratory infection during sleep restriction in young adults. *Sleep*.

Zhang, L. & Zhao, Z.-X. (2007). Objective and subjective measures for sleep disorders. *Neuroscience bulletin* (23:4), 236.