



26th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2022)

## Initial evaluation of substituting a sleep diary by smartwatch measurement

Maksym Gaiduk<sup>a\*</sup>, Ralf Seepold<sup>a</sup>, Natividad Martínez Madrid<sup>b</sup>, Simone Orcioni<sup>c</sup>, Massimo Conti<sup>c</sup>, Juan Antonio Ortega<sup>d</sup>

<sup>a</sup>HTWG Konstanz, Alfred-Wachtel-Str.8, Konstanz 78462, Germany

<sup>b</sup>Reutlingen University, Alteburgstraße 150, Reutlingen 72762, Germany

<sup>c</sup>Università Politecnica delle Marche, Via Brecce Bianche 12, Ancona 60131, Italy

<sup>d</sup>University of Seville, Avda. Reina Mercedes s/n, Seville 41012, Spain

### Abstract

Healthy sleep is required for sufficient restoration of the human body and brain. Therefore, in the case of sleep disorders, appropriate therapy should be applied timely, which requires a prompt diagnosis. Traditionally, a sleep diary is a part of diagnosis and therapy monitoring for some sleep disorders, such as cognitive behaviour therapy for insomnia. To automatise sleep monitoring and make it more comfortable for users, substituting a sleep diary with a smartwatch measurement could be considered. With the aim of providing accurate results, a study with a total of 30 night recordings was conducted. Objective sleep measurement with a Samsung Galaxy Watch 4 was compared with a subjective approach (sleep diary), evaluating the four relevant sleep characteristics: time of getting asleep, wake up time, sleep efficiency (SE), and total sleep time (TST). The performed analysis has demonstrated that the median difference between both measurement approaches was equal to 7 and 3 minutes for a time of getting asleep and wake up time correspondingly, which allows substituting a subjective measurement with a smartwatch. The SE was determined with a median difference between the two measurement methods of 5.22%. This result also implicates a possibility of substitution. Some single recordings have indicated a higher variance between the two approaches. Therefore, the conclusion can be made that a substitution provides reliable results primarily in the case of long-term monitoring. The results of the evaluation of the TST measurement do not allow to recommend substitution of the measurement method.

© 2022 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0>)

Peer-review under responsibility of the scientific committee of the 26th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2022)

*Keywords:* sleep diary; sleep study; wearable

\* Corresponding author. Tel.: +49 7531 206 703.

E-mail address: maksym.gaiduk@htwg-konstanz.de

## 1. Introduction

Sleep is a significant part of our lives and greatly impacts our health and well-being, which was confirmed in numerous scientific studies [1, 2]. Consequently, it is essential to keep sleep healthy, which means maintaining its duration and quality at a sufficient level. Herewith, the restoration of a human's mental state, brain, and body can be ensured [3].

In the case of existing problems with sleep, it is crucial to identify the reason and subsequently apply an appropriate treatment. There are two main groups of methods for the analysis of sleep with the aim of recognising possible sleep disorders – objective and subjective [4]. Objective measurement is based on the use of electronic devices measuring physiological signals of the human body with the following examination [5]. A classic example of an objective approach is polysomnography (PSG) which is accurate though time- and cost-intensive [6]. In the case of subjective measurement, the perceived values are being measured. Typically, it is done with the help of questionnaires or sleep diaries [7].

Both approaches have their advantages and disadvantages, but what is more critical are typically applied to detect different sleep disorders. For instance, for the diagnosis of sleep apnoea syndrome, objective measurement (PSG or polygraphy) is often used [8, 9]. To identify some other sleep disorders, such as insomnia, subjective measurement is usually applied [10]. A sleep diary is one of the commonly used methods. It can also be used to monitor and adjust the therapy, for instance, cognitive behaviour therapy for insomnia (CBT-I) [11].

Due to the necessity of manually filling out a questionnaire, subjective measurement requires daily action by a user. This, consequently, can lead to the reduction of days when the data is being collected, as known from other research [12]. Therefore, it could be advantageous to automatise the collection of necessary data by employing an electronic device, keeping the accuracy of a questionnaire, as it is a standard approach with years of experience and numerous datasets collected for diagnosing some sleep disorders like insomnia.

Several research studies have addressed the substitution of subjective measurement in sleep medicine for objective one and vice versa [13, 14]. Some of them have used a device placed under the mattress to evaluate the measurement results [15]. Some others have used wearables as a device to be compared [16]. Due to the fact that various devices use different types of sensors, have different placements, and the development constantly provides new devices with new characteristics and features, the generalisation of findings cannot provide the exact answer if some specific device could be used for an appropriate substitution of subjective measurement in a particular area of sleep knowledge. Furthermore, the possibility of substitution depends on the art of subjective measure, which means that the specific questionnaire with particular questions must be determined to provide an accurate evaluation. Consequently, not generalisation but individual assessment of devices and their field of application could facilitate precise results. Performed scientific reviews have also pointed out that further research is still necessary to provide reliable and accurate results [17].

This research aims to perform an initial evaluation of the possibility of substituting a sleep diary by means of smartwatch measurement on the example of the Samsung Galaxy Watch 4.

## 2. Methods

According to the defined research aim, the investigation on the selection of appropriate methods was performed with the following implementation of them. This section presents a detailed description of the elements relevant to the evaluation procedure. For better readability, the chapter “Methods” is split into several subsections in accordance with the aspect addressed.

### 2.1. Sleep diary

One of the commonly used methods for subjective sleep evaluation is the application of a sleep diary [18]. Several versions exist, including the adapted German version recommended by the German Sleep Society [19]. Divers questions are being asked to be completed in the evenings and mornings. Some examples of them are according to [18]:

- What time did you get into bed?
- What time did you try to go to sleep?
- How long did it take you to fall asleep?
- How many times did you wake up, not counting your final awakening?
- In total, how long did these awakenings last?
- What time was your final awakening?
- What time did you get out of bed for the day?
- How would you rate the quality of your sleep (very poor/poor/fair/good/very good)?

Evaluating the answers, several additional sleep relevant parameters can be determined, such as total sleep time (TST), total time in bed (TIB), and sleep efficiency (SE). Results of sleep diary completion conjoined with the derived parameters are of high relevance in diagnosing and observing sleep disorders, for instance, in the case of insomnia [20].

## 2.2. Smartwatch

Various devices may be applied for the measurement of sleep parameters. To ensure its comfortable use, the device should disturb a user during sleep as less as possible. To achieve this, several approaches may be applied. A device placed under a mattress may be one of the options [15, 21]. Another commonly used approach is the use of devices that can be comfortably placed on the human body – wearables [22]. There are several kinds of wearable devices, one of them is a smartwatch [23]. For the performed study Samsung Galaxy Watch 4<sup>1</sup> was selected.

The choice is based on several reasons. Anterior models of the Samsung Watch were evaluated to measure physiological parameters with acceptable results [24], including the sleep monitoring application [25]. Therefore, it can be assumed that the newest available version of the smartwatch would provide similar or better results due to using the latest sensors and software system. Furthermore, the relevant sleep parameters such as times of getting asleep/waking up, sleep efficiency, and sleep profile can be measured with this device. The measurement results can be read off directly from the corresponding mobile app or downloaded as a part of the stored private data using the menu option in the mobile application. The adjustment can be made using the mobile app, and no additional action is required before going to bed. The watch can be worn on a wrist during the day and night and has about 30-40 hours of battery life.

## 2.3. Sleep characteristics

Several parameters are typically analysed as a part of subjective sleep measurement. They are either extracted directly from a questionnaire/sleep diary or calculated using this data, as described in Section “Sleep diary”. For the evaluation of the possibility of substitution of sleep diary by objective measurement, several relevant parameters were selected:

- Time of getting asleep
- Wake up time
- Total sleep time (TST)
- Sleep efficiency (SE)

These parameters are either directly measured by the used smartwatch or can be calculated from the available measurement values. Thus, the set of sleep characteristics used for the initial evaluation was determined.

---

<sup>1</sup> <https://www.samsung.com/>

According to [26], there are several opinions regarding the calculation of SE, and to exclude possible misunderstanding by reading this article, we provide the used formula for its determinations. SE for subjective measurement was calculated using the following formula (1):

$$SE = \frac{TST}{(\text{Wake up Time} - \text{Time trying go to Sleep})} * 100\% \quad (1)$$

For the initial evaluation of the possibility of substituting subjective measurement with an objective one, the differences between these two kinds of measurement for every sleep characteristic of interest are to be calculated and analysed.

#### 2.4. Experiment design

A research team defined the design of the experiment performed for the evaluation, which was then followed by study participants. On the first day, the experiment process was introduced to test persons, and smartwatches, as well as sleep diaries, were presented and explained. All participants' questions were addressed, and the devices were tried to be used.

Test persons were asked to wear smartwatches during the day and night or at least put them on 20 minutes before going to bed. They were not allowed to take them off earlier than 20 minutes after getting out of bed. Smartwatches had to be charged whenever a battery level was lower than 35% to ensure their functioning during the entire night. In case of any problems with the device, test persons were to contact the study organisers and stop using it up to the clarification.

Every evening, directly after going to bed, participants were asked to fill out the evening part of a sleep diary. Immediately after getting out of bed in the morning, they were required to fill out another part of a sleep diary with further questions.

### 3. Results

The experiment was performed for the projected evaluation according to the defined design. To facilitate reliable results, the total number of night recordings was aimed to be equal to at least 30, and this number was achieved during the experiment.

The age group of participants was defined as 30-40 years old to avoid possible age-specific deviations. Both male and female participants were included in the study course to obtain more generalised and not gender-specific outcomes. No significant disorders were known to be diagnosed by test subjects.

As proposed in the “Sleep characteristics” section, the absolute differences between the objective and subjective measurements for every night were calculated. After that, the estimation of relevant statistical parameters was performed with the following graphical representation for better comprehensibility of the results. For that, a box plot was selected to represent significant values of measurement distribution.

Three of the four sleep characteristics (asleep time, wakeup time, and TST) determined to be analysed as a part of the experiment can be presented in minutes as the unit of measurement. Therefore, we visualise them combined in Figure 1 to facilitate good comparability. A box plot for every selected sleep characteristic can be seen in this figure, representing relevant statistical measurements, including lower and upper quartiles, median and mean values, and outliers.

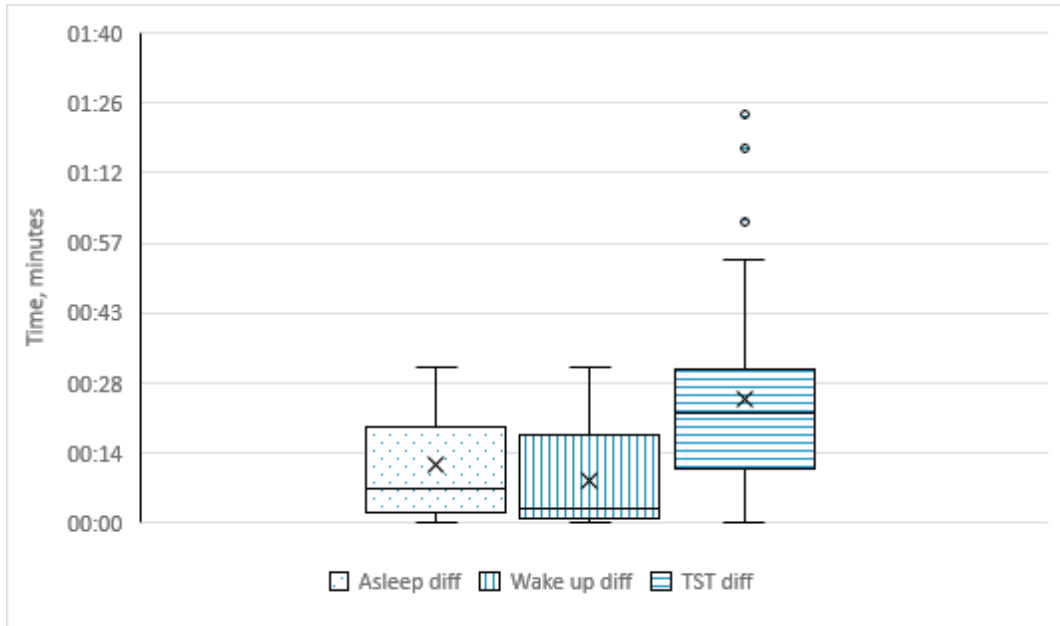


Fig. 1. Absolute differences between objective and subjective measurement for three relevant sleep characteristics for 30 test subjects.

The difference in SE between subjective measurement with a sleep diary and objective measurement with a smartwatch is presented in Figure 2. One can see that both mean and median values of differences are situated between 5% and 6%.

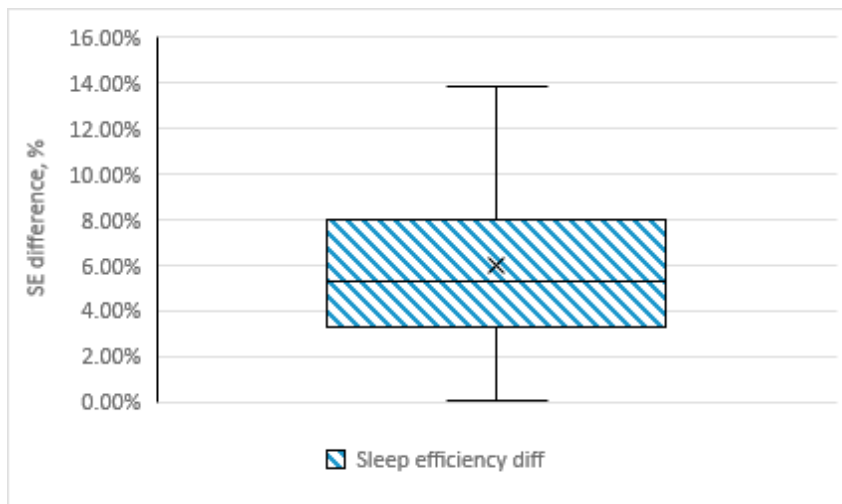


Fig. 2. Absolute differences between objective and subjective measurement of sleep efficiency (SE) for 30 test subjects.

To summarise the main findings in validating the possibility of substituting a sleep diary with a measurement employing the Samsung Galaxy Watch 4, the relevant statistical values for all four sleep characteristics are presented in Table 1.

Table 1. Statistical analysis of the differences between the objective and the subjective measurement.

Sleep characteristic	Mean	Median	Standard deviation
Getting asleep time (minutes)	11	7	10
Waking up time (minutes)	08	3	10
Total sleep time (minutes)	25	22	21
Sleep efficiency (%)	5.99%	5.30%	3.50%

By analysing the results presented in Figure 1 and Table 1, one can see that the difference between objective and subjective measurement approaches in determining “Time of getting asleep” and “Wake up time” is relatively low. Therefore, the substitution of one measurement method with another may be considered. Nevertheless, some outliers are possible. This issue should be considered if only a low number of nights is being analysed, which increases the effect of single outliers on a general pattern.

In the case of TST, the difference between the two measurement approaches is higher, and the number of outliers also increases. This can be partially explained by the fact that the objective measurement identifies more “Awake” periods during the night and assigns them a longer duration than the subjective approach. Furthermore, the errors of measurement of parameters “Time of getting asleep” and “Wake up time” can be added up by determination of TST. Summing up the above points, the TST measurement is not eligible for substituting subjective measurement through objective if a high level of precision is required.

Interpreting the statistical values corresponding to SE measurement and demonstrated in Figure 2 and Table 1, a good level of agreement between the two measurement methods can be recognised. This leads us to the implication of the possibility of substitution of the subjective measurement approach for the objective one.

#### 4. Conclusions

Considering the fact that an automatic measurement of sleep characteristics would increase the comfort for users and avoid possible missing data, as no daily action would be required, the evaluation of the possibility of substitution of sleep diary by an electronic device appears to be meaningful. As different devices can use different kinds of sensors and software systems, so a specific device should be undergone evaluation to provide an accurate conclusion regarding the possibility of the mentioned substitution.

In this work, Samsung Galaxy Watch 4 was considered, being one of the modern smartwatches with up-to-date hardware and software components. Following inferences always refer to this specific smartwatch and cannot be directly generalised. To summarise the obtained results, the following statements can be presented as the main findings:

- The time of getting asleep can be determined using a smartwatch with a high level of accuracy, having the subjective measurement as a reference. This implicates the possibility of substitution of sleep diary measurement by the use of a smartwatch in case of analysing this sleep characteristic.
- The waking up time measurement performed using an electronic device strongly correlates with the subjective measurement. For this reason, it can be concluded that the substitution of these two measurement approaches can be reasonable for this parameter.
- Evaluating the TST as one of the significant sleep characteristics has led us to conclude that substituting the subjective measurement with the objective one and vice versa is not reliable if a high level of precision is required. The reason is a high variance in the distribution of differences between the two mentioned measurement approaches being observed during the performed study.
- Another analysed parameter – SE, appears to be eligible for substituting a sleep diary as a measurement tool by a Samsung Galaxy Watch as the difference between them in the calculated value is, on average, lower than 6%. However, in particular cases, the difference may reach up to 14%, and therefore the substitution may be imprecise, evaluating only a single a minor number of nights.

The results of the performed initial evaluation demonstrate that the substitution of a sleep diary measurement by a Samsung Galaxy Watch 4 may be performed for several relevant sleep characteristics. This can increase comfort and decrease the required users' time. Furthermore, due to the automation of the process, the amount of missing data may be reduced, especially in the case of long-term monitoring.

There are several limitations of the performed study to be mentioned:

- One specific device was evaluated to obtain accurate data that can be reproduced by the same kind of device. A generalisation of results to all smartwatches is not possible due to the diverse hardware and software components used in different wearables.
- For the evaluation, the experiment with 30 nights was performed. Following the traditional assumption, this number of recordings can provide statistically relevant outcomes. However, with the increasing sample size, the transferability of the results to the entire population would be more exact.
- The age of the participants was limited to the range between 30 and 40 years old to obtain statistically significant results for this age group. The obtained outcomes are limited to this subject group, and it cannot be stated that the findings can be transferred to other age groups.
- The performed evaluation was aimed to analyse four sleep characteristics. The obtained within the performed research results cannot be extrapolated to others and are limited to the evaluated ones due to the particularities of measurement of different parameters and variable accuracy depending on the characteristic. As a matter of course, several more parameters can be measured both objectively and subjectively.

According to the results, no general recommendation for substituting sleep diary measurement with an objective approach can be currently expressed.

To overcome the listed study limitations and obtain more comprehensive results, several directions of future work were defined:

- Further studies with the increased size of participants are being planned. Furthermore, other age groups are to be included in the experiment.
- The set of sleep characteristics for the evaluation is projected to be enlarged. Among others, total time in bed, time of getting into bed, and sleep quality should be included in future studies to facilitate comprehensive research outcomes.

It is planned to perform a study with additional wearable devices (including smartwatches from other manufacturers) to compare the measurement results. Furthermore, the aim is to determine if there are significant differences in the findings or if there are meaningful similarities, and some generalisation of outcomes and conclusions regarding the possibility of substitution may be made.

## Acknowledgments

This research was partially funded by Carl Zeiss Foundation: Project “Non-invasive system for measuring parameters relevant to sleep quality” (project number: P2019-03-003) and by the EU Interreg V-Program “Alpenrhein-Bodensee-Hochrhein”: Project “IBH Living Lab Active and Assisted Living”, grants ABH040, ABH041 and ABH066.

## References

- [1] Worley Susan L. (2018). "The Extraordinary Importance of Sleep: The Detrimental Effects of Inadequate Sleep on Health and Public Safety Drive an Explosion of Sleep Research." *P T* **43** (12): 758–63.
- [2] Knutson Kristen L., Phelan Julie, and Paskow Michael J., et al. (2017). "The National Sleep Foundation's Sleep Health Index." *Sleep Health* **3** (4): 234–40.
- [3] Grandner Michael A. (2017). "Sleep, Health, and Society." *Sleep Med Clin* **12** (1): 1–22.
- [4] Ibáñez Vanessa, Silva Josep, and Cauli Omar (2018). "A survey on sleep assessment methods." *PeerJ* **6**: e4849.

- [5] Gaiduk Maksym, Seepold Ralf, and Ortega Juan A., et al. (2020). "Comparison of sleep characteristics measurements: a case study with a population aged 65 and above." *Procedia Computer Science* **176**: 2341–9.
- [6] Hirshkowitz Max (2016). "Polysomnography Challenges." *Sleep Med Clin* **11** (4): 403–11.
- [7] Ibáñez Vanessa, Silva Josep, and Cauli Omar (2018). "A survey on sleep questionnaires and diaries." *Sleep Med* **42**: 90–6.
- [8] Masa Juan F., Corral Jaime, and Pereira Ricardo, et al. (2011). "Effectiveness of home respiratory polygraphy for the diagnosis of sleep apnoea and hypopnoea syndrome." *Thorax* **66** (7): 567–73.
- [9] Berry RB, Quan SF, Abreu A.R., et al. (2020). *The AASM manual for the scoring of sleep and associated events: Rules, terminology and technical specifications Version 2.6*. Darien, Illinois: American Academy of Sleep Medicine.
- [10] Riemann Dieter, Baglioni Chiara, and Bassetti Claudio, et al. (2017). "European guideline for the diagnosis and treatment of insomnia." *J Sleep Res* **26** (6): 675–700.
- [11] Baglioni Chiara, Altena Ellemarije, and Bjorvatn Bjørn, et al. (2020). "The European Academy for Cognitive Behavioural Therapy for Insomnia: An initiative of the European Insomnia Network to promote implementation and dissemination of treatment." *J Sleep Res* **29** (2): e12967.
- [12] Gaiduk Maksym, Seepold Ralf, and Martínez Madrid Natividad, et al. (2021). "Digital Health and Care Study on Elderly Monitoring." *Sustainability* **13** (23): 13376.
- [13] Jungquist Carla R., Pender John J., and Klingman Karen J., et al. (2015). "Validation of Capturing Sleep Diary Data via a Wrist-Worn Device." *Sleep Disord* **2015**: 758937.
- [14] Tonetti Lorenzo, Mingozi Roberto, and Natale Vincenzo (2016). "Comparison between paper and electronic sleep diary." *Biological Rhythm Research* **47** (5): 743–53.
- [15] Gaiduk Maksym, Seepold Ralf, and Martínez Madrid Natividad, et al. (2021). "A Comparison of Objective and Subjective Sleep Quality Measurement in a Group of Elderly Persons in a Home Environment. In: Saponara S, Gloria A de, editors. *Applications in Electronics Pervading Industry, Environment and Society*. Cham: Springer International Publishing, p. 286–291.
- [16] van Hees Vincent Theodoor, Sabia S., and Jones S. E., et al. (2018). "Estimating sleep parameters using an accelerometer without sleep diary." *Sci Rep* **8** (1): 12975.
- [17] King Christine E., and Sarrafzadeh Majid (2018). "A SURVEY OF SMARTWATCHES IN REMOTE HEALTH MONITORING." *J Healthc Inform Res* **2** (1-2): 1–24.
- [18] Carney Colleen E., Buysse Daniel J., and Ancoli-Israel Sonia, et al. (2012). "The consensus sleep diary: standardising prospective sleep self-monitoring." *Sleep* **35** (2): 287–302.
- [19] Hoffmann R. M., Müller T., and Hajak G., et al. (1997). "Abend-Morgenprotokolle in Schlafforschung und Schlafmedizin—Ein Standardinstrument für den deutschsprachigen Raum." *Somnologie* **1** (3): 103–9.
- [20] Maich Kristin H. G., Lachowski Angela M., and Carney Colleen E. (2018). "Psychometric Properties of the Consensus Sleep Diary in Those With Insomnia Disorder." *Behav Sleep Med* **16** (2): 117–34.
- [21] Gaiduk Maksym, Seepold Ralf, and Martínez Madrid Natividad, et al. (2020). "Recognising Breathing Rate and Movement While Sleeping in Home Environment. In: Saponara S, Gloria A de, editors. *Applications in Electronics Pervading Industry, Environment and Society*. Cham: Springer International Publishing, p. 333–339.
- [22] Penzel Thomas, Schöbel Christoph, and Fietze Ingo (2018). "New technology to assess sleep apnea: wearables, smartphones, and accessories." *F1000Res* **7**: 413.
- [23] Roomkham Sirinthip, Lovell David, and Cheung Joseph, et al. (2018). "Promises and Challenges in the Use of Consumer-Grade Devices for Sleep Monitoring." *IEEE Rev Biomed Eng* **11**: 53–67.
- [24] Nissen Michael, Slim Syrine, and Jäger Katharina, et al. (2022). "Heart Rate Measurement Accuracy of Fitbit Charge 4 and Samsung Galaxy Watch Active2: Device Evaluation Study." *JMIR Form Res* **6** (3): e33635.
- [25] Zhou Yi, Price Blaine, and Li Jinfeng, et al. (2020). "Sleep Tracking of a Commercially Available Smart Ring and Smartwatch Against Medical-Grade Actigraphy in Everyday Settings: Instrument Validation Study." *JMIR Mhealth Uhealth* **8** (10).
- [26] Reed David L., and Sacco William P. (2016). "Measuring Sleep Efficiency: What Should the Denominator Be?" *J Clin Sleep Med* **12** (2): 263–6.