A Sleep Analysis Method Based on Heart Rate Variability

Firstbeat Technologies Oy
This white paper has been produced to describe a sleep analysis method based on heart rate variability and acceleration data developed by Firstbeat Technolgies Ltd. Parts of this paper may have been published elsewhere and are referred to in this document.

TABLE of CONTENTS

KEY TERMS

- **Polysomnography, PSG** = The gold-standard resource for sleep study which typically includes measuring brain activity (EEG), eye movements (EOG), chin muscles (EMG), cardiovascular system (ECG) and respiratory system.
- **Sleep stage** = a period of sleep having certain scientifically accepted polysomnographic characteristics.
- **Sleep efficiency** = percentage of time spent asleep from the overall time between going to bed and awakening.
- **Sleep latency** = time required to fall into sleep after going to bed.
- **Heart rate variability, HRV** = A measure of dynamical control of heart rate associated with autonomic nervous system activity, stress, and recovery.

SUMMARY

- Sleep is the most important time for recovery. Good sleep replenishes physical and mental resources and is a cornerstone of healthy lifestyles.
- Good sleep is a combined function of getting enough sleep and the restorative quality of sleep.
- Age-based recommendations for healthy sleep durations exist, e.g. 7-9 hours for adults, and first statements on recommended sleep structure have been made and published by experts in the field.
- Important sleep quality metrics include sleeping long enough, falling asleep in less than 30 minutes, waking up no more than once per night, and being awake for 20 min or less after initially falling asleep.
- A certain spectrum of sleep stages should also be present: descriptive recommendations exist, but there is no full consensus regarding ideal sleep structure.
- Autonomic Nervous System (ANS) plays a key role in recovery and sleep, as ANS activity reflects the restoration of physiological systems and is also associated with sleep onset and sleep stages.
- Firstbeat has developed a unique sleep detection and assessment methodology combining various aspects of sleep from sleep stages to the ANS function using HRV to measure physiological recovery during sleep, the ultimate effect of sleep.

SLEEP AND HEALTH

The Role of Sleep

Sleep is a fundamental part of life, and about 1/3 of our lives is spent sleeping. It is a huge biological investment associated with growth, repair, and maintenance of bodily functions [1]. Sleep supports proper functioning of the immune system. And it is also linked to the removal of waste products and reorganization of the central nervous system networks [1]. Sleep affects almost every type of tissue and bodily system from the brain to cardiovascular, pulmonary, metabolic/hormonal, and immune system activity. Sleep also effects heart rate and heart rate variability, respiratory rate, blood pressure, and body temperature [2-3].

Although all the functions of sleep are not fully understood, sleep is recognized to be an active and dynamic process of physical and mental recovery [4-5]. Sleep is a reversible behavioral state marked by unresponsiveness and perceptual disengagement from the environment [6]. It is comprised of different stages that change cyclically and are identifiable based on variation in brain and neuronal activity documented through polysomnography (PSG), a combination of electroencephalogram (EEG), the electrooculogram (EOG) and electromyogram (EMG) data [6].

Early night is typically dominated by deep sleep called also as slow-wave sleep, SWS, or N3. Deep sleep is associated with slow brain waves, increased protein synthesis, increased growth hormone secretion, and decreased cortisol secretion. The latter part of the night is typically dominated by rapid eye movement sleep, i.e. REM sleep. REM sleep is widely linked to dreaming and information processing [7-8]. A sleep is initiated by relaxed wakefulness (N1) and most of the night is spent in basic light sleep stage (N2). A typical night's sleep includes 4-5 cycles from non-REM (incl. N1-N3) to REM sleep. Awake periods typically account for less than 5% of the night. [6]
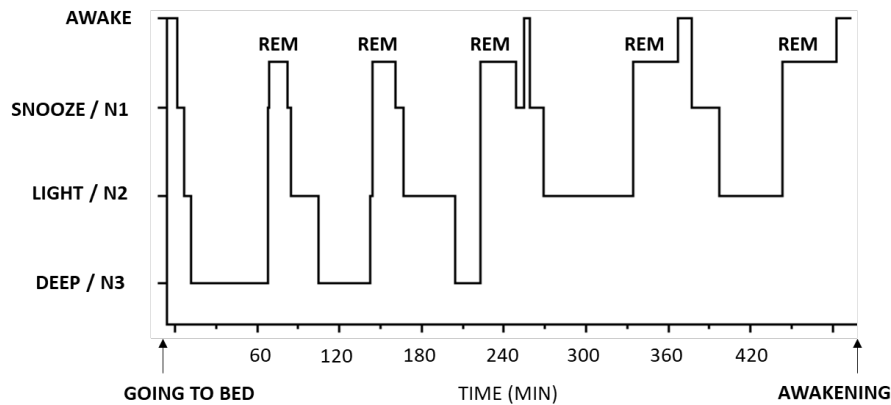
Figure 1. Sleep hypnogram from a typical sleep session showing five non-REM to REM cycles with typically deep sleep dominant during the early part of night and REM sleep dominance later in the night.

Sleep-wake cyclicity is delicately controlled by homeostatic processes and the body's circadian rhythms, where changes in daylight conditions play a major role. The suprachiasmatic nucleus in the hypothalamus seems to act as the master clock of the human body [2]. Elongated awakening periods and/or high load on the body (e.g., resulting from large amounts of vigorous physical activity) increases the homeostatic drive for sleep and increases the recuperative deep sleep the next night [7].

Sleep and Health Status

Sleep has various substantial health effects, and a lack of sleep and sleep deprivation may cause serious negative health consequences. For example, sleeping less than 7 hours per night has been associated with weight gain and obesity, diabetes, hypertension, heart disease, stroke, depression, impaired immune function, increased pain, impaired performance, increased error rates during tasks, greater risk of accidents, and increased risk of death [9]. Sleep duration and morbidity seems to have a U-shaped association, where those who sleep the least and the most are at higher risk for all-cause mortality [10].

Sleep affects learning and memory, cognitive performance, and alertness. Inadequate sleep is associated with difficulties in concentrating, focusing and with memory recall [11]. Sleep can be considered good when falling asleep is easy, when sleep is continuous, when the person does not wake up too early, when waking is accompanied by feelings of being refreshed, and when daytime performance levels are not decreased [12]. These elements, however, can be challenging to evaluate objectively as a whole, as some are definitively subjective in nature [13].

Sleep Recommendations

Sleep's restorative effect is influenced by aspects of sleep quality and quantity. Sleep experts have identified and agreed upon clear recommendations for healthy sleep durations. Recognizing that sleep needs and patterns vary according to age, these expert recommendations are age groups specific. For example, according to the expert panel of the National Sleep Foundation adults (26-64) and young adults (18-25) should sleep between 7 to 9 hours each night [14].

There is less scientific consensus, however, regarding how sleep quality should be defined and measured. Nevertheless, there are first statements available that suggest sleep quality metrics. These suggestions include sleep latency, number of awakenings lasting more than 5 minutes, the amount of wake after sleep onset, and sleep efficiency [15].

The National Sleep Foundation expert panel has also made preliminary suggestions regarding normal, healthy distribution of sleep stages across various age groups. Among adults, for example, according to the panel, deep sleep (N3) of 16-20% and REM sleep of 21-30 % indicates good sleep quality; whereas, transition times from wake to sleep (N1) of over 20% or light sleep (N2) of over 80% from the sleep period does not indicate good sleep quality [15].

Moreover, the autonomic nervous system (ANS) is a key regulatory system for the body, and sleep is reflected in ANS activity. Generally, the parasympathetic branch of the ANS should be primarily dominant during sleep. This reflects a relaxed state. A high sympathetic drive can be a sign of suboptimal recovery and physiological stress. The ANS state is also affected by different stages of sleep. Cardiovascular activity, for example, is very stable in the N3 sleep, whereas, it can be highly variable during REM sleep, often reaching levels seen during wakefulness. [16-18].

Factors Affecting Sleep

Sleep can be affected by various internal and external stressors. High levels of perceived stress, worries, and anxiety can make it difficult to fall asleep and to stay asleep [19-20]. Stress can reduce amounts of deep and REM

sleep and increase awakenings [21]. Too much daylight can disrupt and confuse the body's circadian regulation [22]. Shift work is a typical factor which decreases the amount and quality of sleep by disturbing the body's regulative processes via changing sleeping times [23]. Travel across time zones can similarly disturb the body's internal clock and disrupt sleep. [22]. Alcohol, other stimulating substances, and medications can markedly disturb sleep by affecting brain function, sleep structure, and autonomic nervous system activity [24-25]. Regular physical activity may promote sleep, [26] but disrupting homeostasis with strenuous physical activity can cause sympathetic overdrive in the ANS and negatively affect sleep, especially if performed too close to bedtime [27]. Environmental factors (i.e., light, noise, and room temperature) can also affect sleep [28-29].

While many factors can negatively affect sleep, it is also possible to improve sleep through voluntary behaviors. We can decide, for example, when and how hard we exercise, how we occupy our time in the evenings, and when we go to bed. Sleep can be supported through a commitment to good sleep hygiene and with lifestyle choices [30]. Thus, it would be beneficial to assess and understand the restorative effect of our sleep. By improving our understanding of how lifestyle and personal decisions affect our sleep, we can make better, more informed decisions in support of our health and well-being. To serve that purpose, Firstbeat has developed a comprehensive method for credibly estimating sleep quantity, structure, and the restorative effect of sleep using wearable data. This white paper describes the factors behind the method and presents its validation results.

DESCRIPTION OF THE FIRSTBEAT SLEEP DETECTION AND STAGE CATEGORIZATION METHOD

Firstbeat has created a unique sleep algorithm which uses a neural network-based model to provide real-time sleep analysis. Figure 2 shows a schematic illustration of the method. The algorithm uses beat-by-beat heart rate data (heart rate variability), HRV-derived respiration rate, and wrist/body movement, and time of day data. The user's age, height, weight, and gender are required as background information.

Making the analysis special is the fact that it is capable of providing feedback with little to no delay. This offers a significant advantage over other competing neural network analytic models that require the inclusion of data from both before and after the sleep period, a requirement that results in a substantial delay in feedback availability. Additionally, the calculation process of the Firstbeat algorithm is light and highly efficient, making it possible to implement the algorithm in various wearable devices such as smart- and fitness watches and trackers having limited data processing resources. It also means that it's possible to analyze sleep and offer insight without needing to transfer data and rely on the processing power of a paired smartphone or cloud-based computational solution.

The Firstbeat method capably evaluates when the person is in bed and when they are out of bed awake. For in-bed sleep periods, data is classified into light (N1+N2), deep (N3), and REM sleep stages. Awakening segments that occur after and within the onset of the night's sleep period are also detected. Moreover, the method
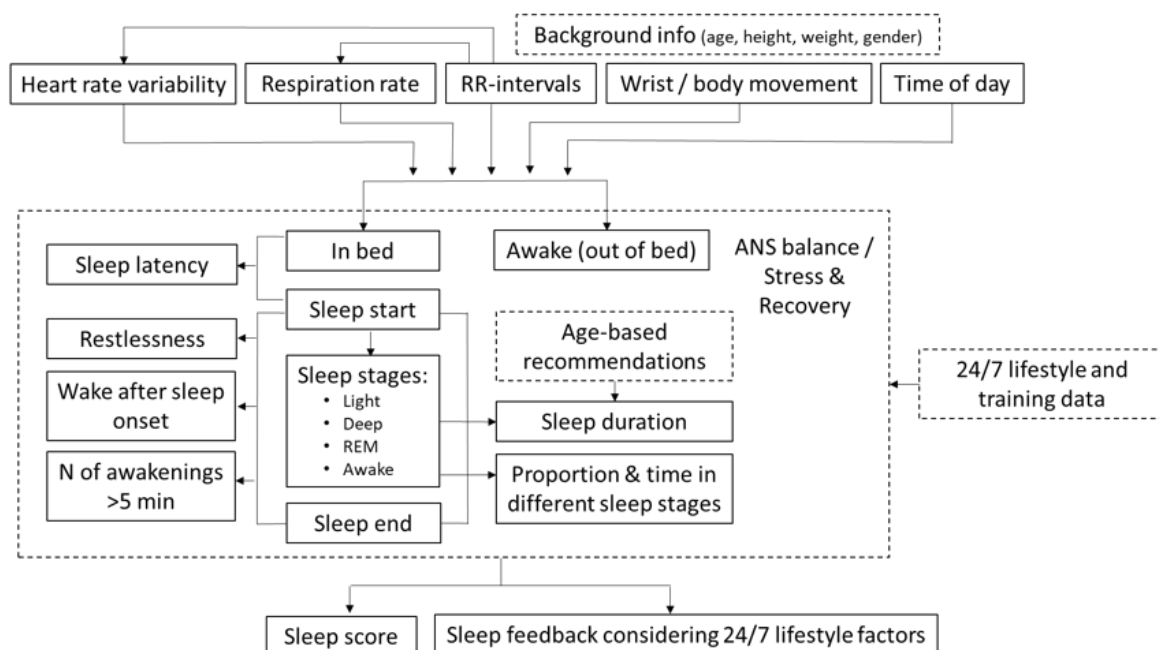

Figure 2. A schematic illustration of the procedural steps involved in the Firstbeat sleep analysis.

evaluates autonomic nervous system (ANS) function using HRV data. This allows for assessment of body stress and recovery states during the sleep. Sleep data is further complemented using a restlessness metric based on acceleration data.

The ANS balance is assessed using beat-to-beat heart rate data. When parasympathetic modulation is dominant, heart rate (HR) is individually low and heart rate variability (HRV) high. This is detected by the Firstbeat analysis and described as a recovery/relaxation state. When sympathetic modulation predominates, HR elevates and HRV generally declines from the individual's baseline levels. This is detected as a stress state by the analysis. The stronger the parasympathetic or sympathetic modulation, the stronger the relaxation or stress intensity, respectively. It is known that during wakefulness parasympathetic activity decreases and/or sympathetic activity increases. During sleep, parasympathetic modulation should predominate to ensure body restitution. The deeper the sleep is, the stronger the parasympathetic modulation is (referring especially to light or deep sleep). However, bursts of sympathetic activity occur during restless periods or awakenings during sleep.

The final feedback allows comprehensive information on:

1. Sleep duration by evaluating the start and end of sleep,
2. Sleep latency (i.e. the time from going to bed to the onset of actual sleep),
3. Sleep stage identification (proportion of the sleep period and the overall accumulated time), and
4. Overall sleep score based on sleep duration, HRV-based stress and recovery information, sleep stages, and movement-based restlessness factors.

In addition, the user of the method can be provided actionable insights in the form of feedback sentences. These feedback sentences can address sleep duration, recovery, restlessness and sleep structure for a given night's sleep. The method also considers lifestyle factors such as repeated behavior patterns and habits in 24/7 use and gives insights on their impacts on the restorative effect of the person's sleep. In this way, the method can become used to help optimize sleep and recovery.

VALIDATION RESULTS AND EXAMPLES

Validation of sleep and sleep stage identification methodologies poses some unique challenges that extend from the current state of the art. The contemporary gold standard for sleep identification and sleep stage classification is human expert interpretation of polysomnography data, including EEG recordings of brainwave activity. Well-trained sleep experts examining

---

ACCURACY METRIC TERMS

There are a multitude of accuracy metrics for measuring the performance of a classification estimator. One that is commonly used in scientific literature related to sleep stage classification is the agreement rate, also called recall, sensitivity, or true positive rate. Defined for any specific sleep stage, the metric stands for the percentage of epochs correctly estimated as a given sleep stage out of the epochs that are classified as the same sleep stage by a reference (e.g. PSG-based) classifier.

In a percentage confusion matrix, such as in Table 1, the agreement rate is represented by the diagonal values. In other words, the rows of a confusion matrix tell how the epochs with a given reference sleep stage are distributed between different sleep stages classified by the estimator.

Complementary to the agreement rate is another metric called specificity, also called true negative rate. The metric is mainly relevant in the case of binary classification, such as sleep vs. awake, where the specificity of sleep is defined as the percentage of epochs correctly estimated as "negative sleep", i.e. awake, out of the epochs that are classified as awake by a reference classifier.

---

the same data, however, do not always perfectly agree on the same set of sleep stage classifications.

Agreement rates between individual experts classifying sleep stages from the same data seems to be slightly over 80% [31], but also values in the range of 67-82 % have been earlier reported [32]. A study with over 2,500 skilled scorers showed an average agreement rate of 82.6 %. The level of agreement was highest for REM sleep, closely followed by Light sleep (N2) and Awake. Scoring agreement for Deep sleep (N3) was 67.4 % and the lowest for stage N1 at 63.0 %. Scorers had difficulty defining sleep onset, the first epoch of stage of N2 after N1 and the first epoch of REM after N2. Discrimination between stages Light and Deep sleep was particularly difficult for scorers [31]. This confirms that, in practice, sleep stage classification is a rather challenging task to perform even for skilled sleep experts with direct access to highly detailed brain activity data.

The dataset used to develop and validate the Firstbeat sleep detection and sleep stage classification methodology included expertly stage categorized polysomnography sleep data from 110 adult subjects (780 hours of sleep data). It also included heart rate data in the form of RR-intervals obtained from ECG measurement and was accompanied by movement data obtained using accelerometers.

Only a small portion of the total dataset was used at a time to create the analytic model during the development stage compared to the validation stage. The entire dataset was utilized to validate the method. This was done to reduce the risk of overfitting and therefore over-estimating the accuracy of the results.

When assessing the validity of sleep stage classifications made by the Firstbeat method compared to polysomnography-derived reference classifications, the reference sleep stages were resampled from 30-second bins to 5-second bins. This resulted in a total of 557,776 datapoints, and each 5-second sample was considered as a single datapoint for agreement comparison.

The resulting Firstbeat method differentiated sleep state from awake state with 94 % sensitivity (true-positive rate) and 63 % specificity (true-negative rate). Epoch-to-epoch agreement weighted by the manifestation of sleep stages was 66%. If an offline implementation of the same algorithm (no optimization for limited resources available on embedded systems) was used, the epoch-to-epoch classification accuracy improved by approximately 3 percentage points to around 69%. The accuracy of the method in classifying sleep stages can be considered reasonably good.

Table 1 offers a confusion matrix for estimated sleep stages based on the classification of all 557,776 data points. The table shows the agreement rate achieved between sleep stage classifications arrived at using the Firstbeat method and those identified through expert interpretation of polysomnography recordings. For cases where the two methods disagreed, the rates at which other possible states were assigned according to each method are also shown.

		Firstbeat estimate			
		Deep	Light	REM	Awake
Reference	Deep	**67%** 73900	31% 34647	0% 423	1% 1433
	Light	12% 35775	**67%** 198636	11% 33385	9% 27278
	REM	1% 504	34% 32304	**64%** 61326	2% 1716
	Awake	4% 2059	28% 15628	7% 3789	**62%** 34973

Table 1. The confusion matrix of sleep stage classification with the Firstbeat Method. Diagonal values in the table indicate the classification agreement rate expressed as a percentage and the number of agreed epochs for each sleep stage.

Figures 3 and 4 offer two examples of sleep documented with HR, HRV, and accelerometer-based movement data. ANS activity derived from HRV-data is interpreted and shown in the form of stress and recovery. Sleep stages classified using the Firstbeat method are shown along with those arrived at by expert interpretation of polysomnography data. Examples of feedback sentences based on the Firstbeat method are also included.
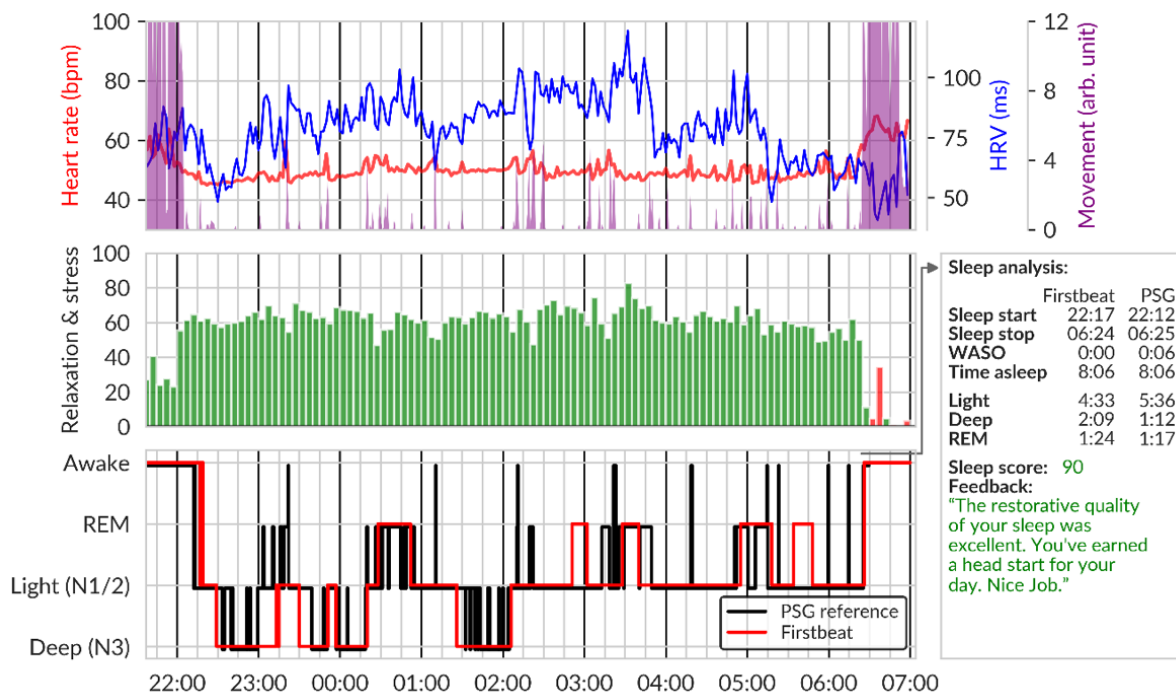

Sleep analysis:

	Firstbeat	PSG
Sleep start	22:17	22:12
Sleep stop	06:24	06:25
WASO	0:00	0:06
Time asleep	8:06	8:06
Light	4:33	5:36
Deep	2:09	1:12
REM	1:24	1:17

Sleep score: 90
Feedback:
"The restorative quality of your sleep was excellent. You've earned a head start for your day. Nice Job."

Figure 3. An example from individual sleep period. The figure represents heart rate, heart rate variability and acceleration data together with the HRV-based stress and recovery analysis results, and the sleep and awake detection with sleep stages compared to the polysomnographic reference result
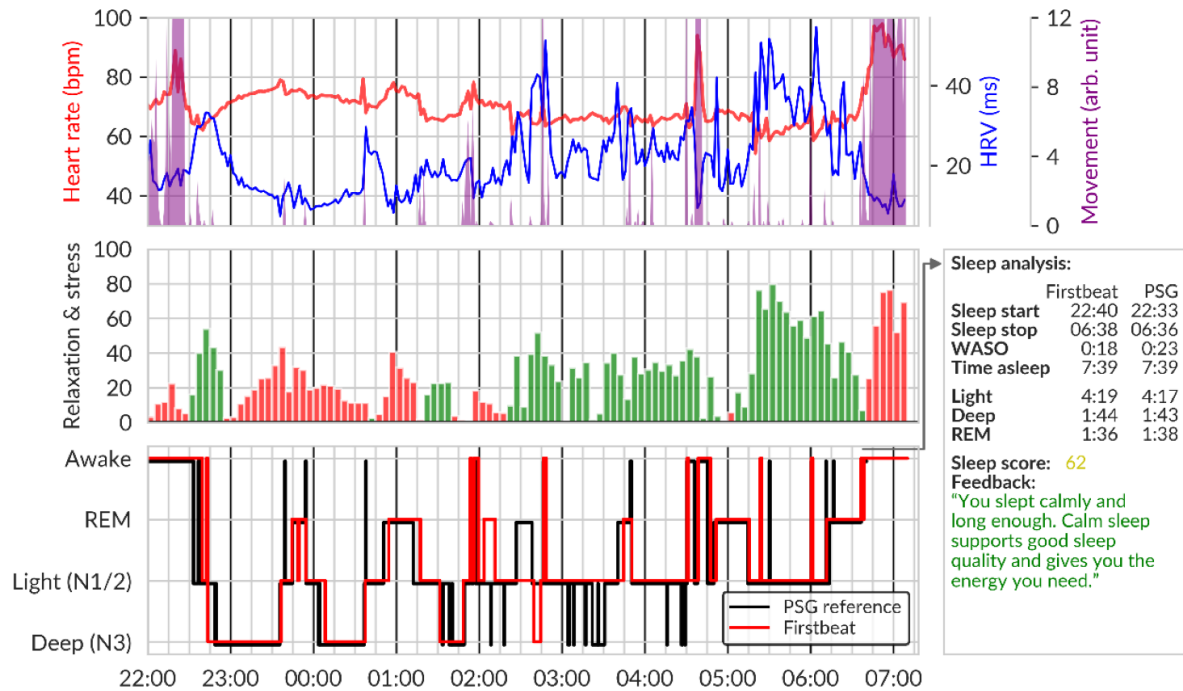
Figure 4. Another example from individual sleep period.

USE OF SLEEP FEEDBACK

The Firstbeat sleep analysis can be used for various health and performance related purposes. It is possible to comprehensively track whether a person gets enough sleep over time, precluding the need for keeping a sleep diary. By documenting going-to-bed and sleep times, the method facilitates effortless tracking of regularity of sleep-wake cyclicity. Sleep quality metrics can be used to help understand the restorative effects of sleep and how various lifestyle factors such as stress, daily activity, and exercise impact sleep. This way, the method can help to identify and promote activities, behaviors, and habits which support good sleep for a given individual.

From a fitness and performance perspective, sleep metrics can be used to optimize training decisions, providing a basis for training prescription and guidance. For example, poor sleep can increase the time needed to fully recover from a strenuous workout, so a lighter training session or even rest might be recommended instead of a more demanding workout By making sleep more visible, the method removes the mystery from the biologically important time we spend sleeping, unconscious of our situation and surroundings. As a result, vital insight becomes available regarding how well a person has rested, and this insight can be used to make informed daily decisions for better health and performance.

REFERENCES AND FURTHER READING

[1] **Schmidt MH (2014).** The energy allocation function of sleep: A unifying theory of sleep, torpor, and continuous wakefulness. Neuroscience & Biobehavioral Reviews, 47: 122-153

[2] **Barone DA & Krieger AC (2015).** The Function of Sleep, review. AIMS Neuroscience, 2 (2): 71-90. DOI: 10.3934/ Neuroscience. 2015.2.71

[3] **Committee on Sleep Medicine and Research (2006).** Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem. Harvey R. Colten HR & Altevogt BM, (Eds.). The National Academies Press, Washington DC.

[4] **Åkerstedt T et al. (2009).** "Sleep and recovery", In Sonnentag et al. (Ed.) Current Perspectives on Job-Stress Recovery, Emerald Group Publishing Limited, Bingley, 205-247.https://doi.org/ 10.1108/S1479-3555(2009)0000007009

[5] **Åkerstedt T & Nilsson PM (2003).** Sleep as restitution: an introduction. Journal of Internal Medicine, 254: 6–12

[6] **Carskadon MA & Dement WC (2011).** Monitoring and staging human sleep. In Kryger T et al. (Eds.), Principles and practice of sleep medicine, 5th edition, pp:16-26. St. Louis: Elsevier Saunders.

[7] **Dijk D-J (2009)**. Regulation and Functional Correlates of Slow Wave Sleep. Journal of Clinical Sleep Medicine. Supplement to Vol.5 (2): S6-S15

[8] Hobson JA & Pace-Schott EF (2002). The cognitive neuroscience of sleep: neuronal systems, consciousness and learning. Nature Reviews Neuroscience, 3: 679–693

[9] Watson et al. (2015) Recommended Amount of Sleep for a Healthy Adult: A Joint Consensus Statement of the American Academy of Sleep Medicine and Sleep Research Society. Sleep, 38(6): 843-844. doi: 10.5665/sleep.4716.

[10] Callicchio L & Kalesan B (2009) Sleep duration and mortality: a systematic review and meta-analysis. Journal of Sleep Research, 18: 148-158. doi: 10.1111/j.1365-2869.2008.00732.x

[11] Rasch B & Born J (2013) About sleep's role in memory. Physiological Reviews, 93 (2): 681-766. doi: 10.1152/physrev. 00032.2012.

[12] Harvey et al. (2008). The Subjective Meaning of Sleep Quality: A Comparison of Individuals with and without Insomnia. Sleep, 31(3): 383–393.

[13] Krystal AD & Edinger JD (2008). Measuring Sleep Quality. Sleep Medicine, 9 (Suppl 1): S10-7 DOI: 10.1016/S1389-9457(08)70011-X

[14] Hirskowitz et al. (2015). National Sleep Foundation's sleep time duration recommendations: methodology and results summary. Sleep Health, 1 (1): 40-43.

[15] Ohayon et al. (2017). National Sleep Foundation's sleep quality recommendations: first report. Sleep Health, 3 (1): 6-19.

[16] Boudreau P et al. (2013). Circadian variation of heart rate variability across sleep stages. Sleep, 36 (12): 1919-1928.

[17] Trinder et al. (2012). Sleep and cardiovascular regulation. European Journal of Physiology, 463:161–168. DOI 10.1007/ s00424-011-1041-3

[18] Tobaldini et al. (2013). Heart rate variability in normal and pathological sleep, Frontiers in Physiology, 4: 4-11.

[19] Papadimitriou GN & Linkowski P (2005). Sleep disturbance in anxiety disorders. International Review of Psychiatry. 17(4): 229-36.

[20] Sanford LD et al. (2015). Stress, arousal, and sleep. Current Topics in Behavioral Neuroscience, 25: 379-410. doi: 10.1007/7854_2014_314.

[21] Eui-Joong & Dimsdale (2007). The Effect of Psychosocial Stress on Sleep: A Review of Polysomnographic Evidence, Behavioral Sleep Medicine, 5:4, 256-278, DOI: 10.1080/ 15402000701557383

[22] Albrecht U & Ripperger JA (2018). Circadian Clocks and Sleep: Impact of Rhythmic Metabolism and Waste Clearance on the Brain. Trends in Neuroscience 41(10):677-688. doi: 10.1016/j.tins.2018.07.007.

[23] Kecklund G & Axelsson J (2016). Health consequences of shift work and insufficient sleep. BMJ 2016;355:i5210 doi: 10.1136/bmj.i5210.

[24] He S et al. (2019). Alcohol and sleep-related problems. Current Opinions in Psychology, 30: 117-122. doi: 10.1016/ j.copsyc.2019.03.007.

[25] Pietilä et al. (2018). Acute Effect of Alcohol Intake on Cardiovascular Autonomic Regulation During the First Hours of Sleep in a Large Real-World Sample of Finnish Employees: Observational Study. JMIR Mental Health, 5 (1), e23. doi:10.2196/mental.9519

[26] Kredlow MA (2015). The effects of physical activity on sleep: a meta-analytic review. Journal of Behavioral Medicine, 38 (3): 427–449

[27] Myllymäki T et al. (2011). Effects of vigorous late-night exercise on sleep quality and cardiac autonomic activity. Journal of Sleep Research, 2011 20(1 Pt 2):146-53. doi: 10.1111/j.1365-2869.2010.00874.x.

[28] Halperin D (2014). Environmental noise and sleep disturbances: A threat to health? Sleep Science, 7 (4): 209-212. https://doi.org/10.1016/j.slsci.2014.11.003

[29] Okamoto-Mizuno & Koh Mizuno K (2012). Effects of thermal environment on sleep and circadian rhythm. Journal of Physiological Anthropology, 31(1): 14. **doi: 10.1186/1880-6805-31-14**

[30] Irish LA et al. (2015). The role of sleep hygiene in promoting public health: A review of empirical evidence. Sleep Medicine Reviews, 22:23-36. doi: 10.1016/j.smrv.2014.10.001.

[31] Rosenberg & Van Hout (2013). The American Academy of Sleep Medicine inter-scorer reliability program: sleep stage scoring. Journal of Clinical Sleep Medicine, 9 (1): 81-87. doi: 10.5664/jcsm.2350.

[32] Norman et al. (2000). Interobserver agreement among sleep scorers from different centers in a large dataset. Sleep, 23(7): 901-908.

For more information:
Firstbeat Technologies Oy
Yliopistonkatu 28 A
FI-40100 Jyväskylä, FINLAND
info@firstbeat.com
www.firstbeat.com
www.firstbeat.com/physiology/research-and-publications