Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 155 (2019) 242–249

Procedia Computer Science

www.elsevier.com/locate/procedia

The 16th International Conference on Mobile Systems and Pervasive Computing (MobiSPC)
August 19-21, 2019, Halifax, Canada

How Physical Exercise Level Affects Sleep Quality? Analyzing Big Data Collected from Wearables

Xiaoli Liu[a,*], Satu Tamminen[a], Topi Korhonen[b], Juha Röning[a]

[a] *University of Oulu, Pentti Kaiteran katu 1, 90014 Oulu, Finland*
[b] *Polar Electro Oy, Professorintie 5, 90440 Kempele, Finland*

Abstract

Physical exercise and sleep have independent, yet synergistic, impacts on the health. However, the effects of acute exercise level on sleep quality have not been well investigated. We utilize statistical methods to investigate the differences of exercise level between the good and bad sleep nights. Our results present a complex interrelation between physical exercise and sleep quality with analyzing large personal data sets collected from wearables. As far as we know, this is the first study to investigate insights of interrelation of physical exercise and sleep quality based on a big volume of data collected from wearable devices of real users.

Keywords: Data analytics; wearables; sleep quality; statistical methods

1. Introduction

Sleep is incredibly important for the health and has a deep impact on quality of life. Sleep deprivation can cause catastrophic results, especially for those in professions requiring high accuracy and safety levels. Epidemiological studies have reported an association between daily physical activities and sleep quality. Exercise has many benefits, including improving the sleep quality, increasing the sleep amount, reducing stress, and helping with disorder. However, how to do exercise, such as how much exercise to do per day (exercise duration), when to do exercise (exercise time), and what kind of exercise to do (exercise intensity) can lead to a good sleep night are not well studied yet. Therefore, it is critical to perform systematic research to address this challenge.

Pervasive wearable actigraphy devices monitor the people during daily life and support data analytics technologies to enable discovery of valuable insights into the behavioral patterns related to sleep. Because wearables spread sleep awareness beyond clinicians and researchers to include million users who rely on health applications to track their

* Corresponding author. Tel.: +358452766168.
E-mail address: xiaoli.liu@oulu.fi

10.1016/j.procs.2019.08.035

sleep quality and physical activities, they have paved the way for scalable sleep health study. Actigraphy devices utilize internal sensors to collect data of physical activities and sleep. These devices generate big volumes of data in real time and at a high velocity. Analyzing wearable device data to build descriptive statistical models and provide knowledge for practitioners and users is a challenge for pervasive systems. Depending on the analytic goals and model types, data can be presented in a raw or processed form and new features need to be formulated.

Most existing research focuses on how to effectively measure sleep efficiency with wearables [1] and how to predict sleep quality based on small exercise data sets collected from volunteers [2]. In this research, we explore an innovative idea of utilizing statistical methods for studying the effect of exercise level on sleep quality and investigating how to build a generalized model to predict the sleep quality based on the exercise information of users. To achieve our goal, we start with investigating users with large exercise and sleep variations and focus on those users with good and bad sleep nights. We perform a comprehensive analysis on the difference of exercise level for good and bad sleep nights with big real data sets collected from wearables. Our results of analyzing selected 271 users show that the interrelation between physical exercise and sleep is complex and weak. It is hard to build a good generalized predictive model to predict the sleep efficiency using the exercise information. Therefore, building the predictive models within similar users group is more practical. The main contribution of this article is a comprehensive analytics of interrelation of physical exercise and sleep quality based on voluminous real data collected from wearable devices. We identify the possibility of building a generalized predictive model and important predictors can be used to build personalized predictive models. As far as we know, this is the first effort at the domain based on a large volume of data from real users. Our approach serves as the foundation for tools which provide knowledge for practitioners and potentially predict sleep disorders and even identify the lifestyle factors affecting a particular disorder.

2. Background and related work

In this section, we discuss the background and extensively review the related work on physical exercise and sleep efficiency with focusing on the utilization of wearables. Sleep has a significant impact on quality of life. For centuries, researchers have been exploring characteristics of sleep and its clinical aspects, but most research is limited only into laboratory and hospital experiments. Traditionally, the analytics of physical exercise and sleep requires an extensive and time-consuming clinical interpretation by medical experts. For example, Passos et al. [3] performed an evaluation of the effect of long-term moderate aerobic exercise on sleep, quality of life and mood of individuals with chronic primary insomnia. Myllymäki et al. [4] studied the effort of vigorous late-night exercise on sleep quality and cardiac autonomic activity.

Wearables are significant technology advancements with huge potential for physical exercise and sleep analytics. Only during recent years, the pervasiveness of wearable devices and health and wellness applications by the public have ignited the quantified self-revolution. Wearables, such as actigraph accelerometers, generate a continuous time series of a person's daily physical exertion and rest. Wearables enable us to monitor sleep and physical activity behaviors for longer periods outside the lab or clinic settings. Sathyanarayana et al. [5] explored the usage of wearables for sleep, including algorithms for robust human activity recognition and predictive methodologies for sleep quality estimation. Hoque et al. [6] presented a body position and movements monitoring system based on the WISP platform. In addition, a few works demonstrate the relationship between sleep quality of people and contexts [2].

Data analytics has been extensively performed to study sleep patterns and characteristics from actigraphy data [7][8][9][10]. In addition, interest in sleep data analytics is not restricted to researchers and clinicians, but are shared by health-conscious citizens nowadays as well. Data are collected from millions of people using affordable pervasive wearable devices to track their daily physical activity and sleep. This data are integrated into repositories, allowing integration with other medical sensors and supporting holistic data analytics. In addition to the data collected from millions of fitness trackers and smart-watches that monitor sleep, there exist data from context sensors. Multiple data sources require the development of new data fusion algorithms to discover novel insights. For example, Polar Electro fitness data are collected from multiple sources and integrated into their fitness platform. The system enables the end users to share the data with applications they prefer. Polar Electro API Open AccessLink provides a direct information sharing link between the Polar ecosystem and data systems of other organizations. By correlating human activity sensed by wearables, we pave the way to provide insights into clinicians and individuals for the early diagnosis of sleep problems, which directly influence the quality of sleep of individuals.

3. Datasets and Data preparation

3.1. Datasets

We utilize anonymous datasets from Polar Flow database continuously collected from 2943 users for eight months with Polar wearables. This dataset integrates the data from physical activities and data from Polar Sleep Plus®. Polar Sleep Plus has been proved to be an accurate method for measuring sleep quality [11] [12]. Exercise data include exercise starting time, exercise ending time, calories consumption, exercise duration, duration in different heart zones, maximum heart rate, minimum heart rate, resting heart rate, etc. The sleep data contain sleep starting time, waking time, sleep span, sleep continuity, actual sleep (sleep efficiency), sleep feedback, etc. Sleep efficiency is generally used in sleep science and it refers to the actual percentage of sleep time. More specifically, it is the ratio of the total asleep time compared to the total amount of time spent in bed in a night. Sleep continuity is a parameter which is only calculated by Polar devices and it describes how continuous the sleep is [13]. Sleep Plus utilizes intelligent algorithms based on wrist movements and positions to track the sleep timing, duration, and quality [14]. To study sleep quality, we focus on sleep efficiency as the sleep quality index in this paper.

3.2. Data preparation

Data preparation is performed to guarantee the data quality. Before discussing data preparation, we introduce Basal Metabolic Rate (BMR), Metabolic Equivalent of Task (MET) and trimp in order to better understand the methods of performing new feature construction. BMR represents the amount of energy expended while at rest. MET is a physiological measure expressing the energy cost of physical activity and is defined as the ratio of the energy consumption to the resting metabolic rate. Trimp is a method of quantifying training load, which takes the intensity of exercise into consideration. In this research, we calculate the new feature trimp based on exercise maximum heart rate, average heart rate, resting heart rate, exercise duration, and gender.

We select exercises with which calories per hour consumption ranges from 50 cal to 2000 cal and filter the records with exercise duration less than 10 minutes. Records with calories marked with N/A are removed. Each exercise duration is calculated based on the exercise starting and stopping time. According to expert knowledge, exercise intensity may have impact on sleep quality, and relative calories take personal information into account. Therefore, three new features intensity1, intensity2, and relative calories are constructed. Intensity1 and intensity2 are constructed to represent the exercise intensity. Intensity1 is defined as the average MET value calculated based on calories, duration and BMR. Intensity2 is constructed using trimp and duration. Relative calories is calories/BMR, which eliminates the impact of variation in calories consumption between individual.

As a motivating example, we consider a scenario that one person perform one or more exercises in one day. For each day, the total exercise information, the last exercise information, sleep efficiency and sleep continuity are collected. The total trimp in one day for each person is calculated as the sum of each exercise trimp value; the total duration is the sum of all exercise duration; the total relative calories is calculated as the sum of all exercise relative calories; and the total intensity1 and total intensity2 are calculated based on the total calories consumption, total duration, total trimp, and BMR value. The last exercise is the one with the latest exercise starting time in one day. Variables total duration (TD), total trimp (TR), total relative calories (TRC), total intensity1 (TI1), total intensity2 (TI2), last exercise trimp (R), last exercise duration (D), last exercise relative calories (RC), last exercise intensity1 (I1), last exercise intensity2 (I2), span time between last exercising time to sleep start time (TOS) in a day are selected for observing the difference of exercises between the good and the bad sleep days. The total trimp in one day for each person is calculated as the sum of each exercise trimp value; the total duration is the sum of all exercise durations; the total relative calories is calculated as the sum of all exercise relative calories; and the total intensity1 and total intensity2 are calculated based on the total calories consumption, total duration, total trimp, and BMR value.

4. User selection and ratio identification

In this section, we describe how to select the users with larger sleep and exercise variations and how the ratios are defined and calculated for each variable. The following four steps are carried out in order to achieve our goals.

Step 1: Selection of users with the larger sleep variation. Sleep records with sleep efficiency ranges from 75% to 100% are selected for the analysis. Sleep efficiency under 75% are usually regarded as non reasonable measurements. The sleep data are grouped based on user IDs. For each user i, the mean value of sleep efficiency $\mu_i(e)$, the standard deviation of sleep efficiency $sd_i(e)$, the mean value of the sleep continuity $\mu_i(c)$, and the standard deviation of sleep continuity $sd_i(c)$ are calculated for $i = 1, 2, ..., k$, where k is the number of users. The character e is used to represent efficiency. $\mu(sd, e) = \frac{\sum_{i=1}^{k} sd_i(e)}{k}$ and $sd(sd, e) = \sqrt{\frac{\sum_{i=1}^{k}(sd_i(e)-\mu(sd,e))^2}{k-1}}$ are calculated respectively, which are the mean value and the standard deviation of the $sd_i(e)$. The threshold can be defined as $T(\rho, e) = \mu(sd, e) + \rho sd(sd, e)$. The larger the value of ρ is, the bigger the differences of sleep efficiency between the good and bad days are. In this research, we take $\rho = 1.281552$ and choose the users that satisfy $sd_i(e) > T(\rho, e)$. By doing this, we can have the sleep efficiency in good days and bad days with less overlapping, while still have enough sample users.

Step 2: Selection of users with larger exercise variation based on variable trimp. The exercise data are grouped based on user IDs and the standard deviation of trimp for each person $sd_i(trimp)$ is calculated. The threshold $T(trimp)$ is defined as the 80% of $sd_i(trimp)$. By inner joining the two selected users groups, the users having larger sleep and exercise variation are selected.

Step 3: Identify very good and bad nights for the selected users. Let S_i^j denote j^{th} sleep records of user i. The good and bad sleeping nights are defined based on mean and standard deviation value of sleep efficiency. The night for user i is considered as good, if $S_i^j > \mu_i(e) + sd_i(e)$; and considered as bad, if $S_i^j > \mu_i(e) + sd_i(e)$. $\mu_i(e)$ and $sd_i(e)$ are calculated in step 1 and based on all sleep history records of user i. When good and bad sleep nights are selected for each user, the exercise information of corresponding days will be left joined to the nights sleep information. During the process of left join, there are nights having corresponding days exercise information and nights without the corresponding days exercise information. Values with zeros are filled for those days without exercise information in the left join process. Therefore, two cases are considered here. Case 1 considers those nights that the corresponding days are exercise days and case 2 consider all those nights no matter the corresponding days are exercise or non-exercise days. In case 2, variables TI1, TI2, I1, I2, and TOS are not meaningful. Therefore, only seven variables are considered. Let $n_i(g)$ and $n_i(b)$ represent the number of exercise days in good and bad sleep nights for user i, respectively, where $i = 1, 2, ..., q$ and q is the number of the selected users. The selected users need to have both good and bad nights with exercise, which means $n_i(g) \geq 1$ and $n_i(b) \geq 1$.

Step 4: Define and compute the ratios of each variable for each user. The goal is to observe if there is any difference of exercise between the good and bad sleep nights for users. To achieve the goal, a ratio to measure the difference is proposed. The ratio for each variable and for each person is defined as Variable(r)=mean(variable, in good nights)/mean(variable, in bad nights). For example in case 1, user i has n good nights records and corresponding days exercise information $(TD_i^j, TR_i^j, ..., I2_i^j, TOS_i^j)$, where $j = 1, 2, ..., k_1$ and $k_1 \leq n$, user i also has m bad nights records and corresponding exercise information $(TD_i^l, TR_i^l, ..., I2_i^l, TOS_i^l)$, where $l = 1, 2, ..., k_2$ and $k_2 \leq m$. Then the ratio of variable TD for user i is $TD_i(r) = \frac{\sum_{j=1}^{k1} TD_i^j/k_1}{\sum_{l=1}^{k2} TD_i^l/k_2}$. One user's other ratios $(TR_i(r), ..., TOS_i(r))$ and all users' ratios are calculated by using this approach. If the ratio is greater than 1, it indicates that for good nights, users averagely exercise more or consume more or do exercise earlier compared to their bad sleep nights. For some users, good sleep nights might be associated with more exercise compared with bad sleep nights; for some users, good sleep might be associated with less exercise compared with bad sleep nights; and for some users, exercise time might have an important affect on the good and bad sleep nights.

Finally, 271 out of 2943 users with larger sleep and exercise variation are selected to compare the differences of exercises between good sleep days and bad sleep days. In total, 8945 sleep nights were tracked where 2414 nights with exercise days and 6531 nights without exercise days. Ratios are calculated for each variable for each of the selected users.

5. Effects of exercise on sleep efficiency

5.1. Results of selected users

Figure 1 presents the sleep efficiency density plot for all selected 271 users, which shows an obvious sleep efficiency difference between the good and bad nights. The density plot of bad sleep nights has a little intersection with the

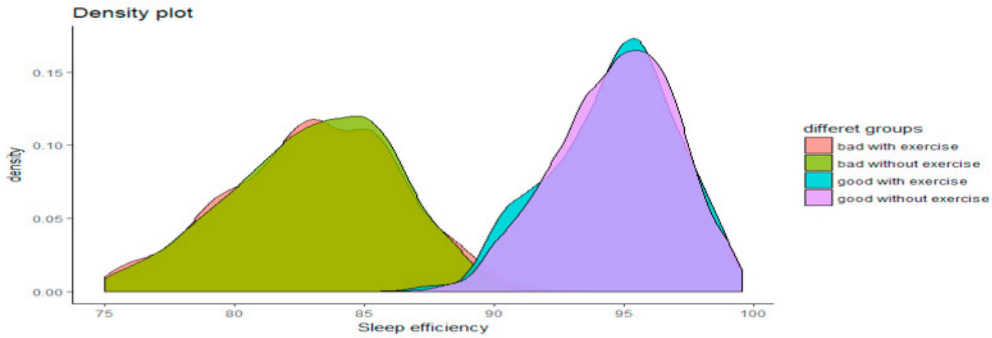

Fig. 1. Sleep efficiency density for users in different groups.

density plot of good sleep nights. Within the bad nights, the efficiency distribution of bad nights with exercise is heavily overlapping with bad nights without exercise. The same situation happens within the good nights.

Table 1. 271 users' exercise statistics and test results

Variables	m±sd (Bad Sleep Days)	m±sd (Good Sleep Days)	p-value of wilcox-test	m(Ratios)	m(log(Ratios))	p-value of t-test
TR	180.30±112.96	182.54±112.59	0.4464	1.26	0.00005	0.999
TD(m)	94.39±58.26	94.63±56.76	0.4554	1.28	0.0035	0.9291
TRC	690.58±456.61	692.39±448.76	0.7403	1.39	-0.007	0.8746
TI1	7.51±2.72	7.49±2.75	0.8971	1.04	-0.0115	0.5643
TI2	2.01±0.71	2.02±0.71	0.7679	1.04	-0.0038	0.8351
T	151.97±105.35	154.82±107.94	0.6316	1.32	0.0019	0.6629
D(m)	79.44±53.86	80.61±54.09	0.4877	1.33	0.02	0.6232
RC	590.25±434.86	599.49±452.93	0.9243	1.54	0.0019	0.6962
I1	7.43±2.77	7.41±2.81	0.8229	1.06	-0.0018	0.5643
I2	2.01±0.72	2.02±0.73	0.7714	1.05	-0.0012	0.9488
TOS(h)	8.19±4.48	8.87±4.59	0.00825	1.17	-0.005	0.8871

When all the selected 271 users are taken into consideration for the nights with exercise days, the unpaired two-samples wilcoxon test and one sample t-test for log (Ratios) are performed. The statistics information of the exercise data in good and bad sleep nights and results of the two tests are presented in Table 1. Table 1 indicates that there is not a significant exercise difference between the good and bad nights except the exercise time with the p-value of the wilcox-test smaller than 0.05. The p-values of the one sample t-test for log(ratios) are all larger than 0.05. The ratio eliminates the personal impact and is good for the analysis of more or less exercise.

Four subgroups from the selected 271 users are chosen to show the complexity of the association between exercise and sleep. Figure 2 presents the exercise information of group A and Group B users in bad and good sleep days, where 0 represents bad sleep days and 1 represents good sleep days. Pictures in the left two columns in figure 2 present exercise information of group A users in good and bad sleep days. Group A having good sleep days averagely exercise more compared to their bad sleep days. The first, second, and third quartiles for total duration in sleep group 0 is (28.64, 43.67, 68.43), and in sleep group 1 is (60.10, 97.33, 192.11). Two-samples wilcoxon test has been performed to compare the means of the total duration in those two groups with p-value=1.529e-07, which means the exercise shift in bad sleep days and good sleep days' means is not equal to 0. We put the variables in the following order ($TD, TR, TRC, TI1, TI2, D, R, I1, I2, TOS$), and the p-values of the two-samples wilcoxon test for the variables are (1.529e-07, 2.858e-09, 6.937e-07, 0.3676, 0.1091, 2.5898e-09, 8.297e-10, 2.892e-07, 0.394, 0.277, 0.535). The p-values of exercise duration, relative calories consumption and trimp are smaller than 0.05. The p-values for the two intensities and exercise time are larger than 0.05. Pictures in the right two columns in figure 2 present exercise information of group B users in good and bad sleep days. Those users having good sleep days averagely exercise less compared with their bad sleep days. The p-values of the two-samples wilcoxon test for the variables are (1.934e-

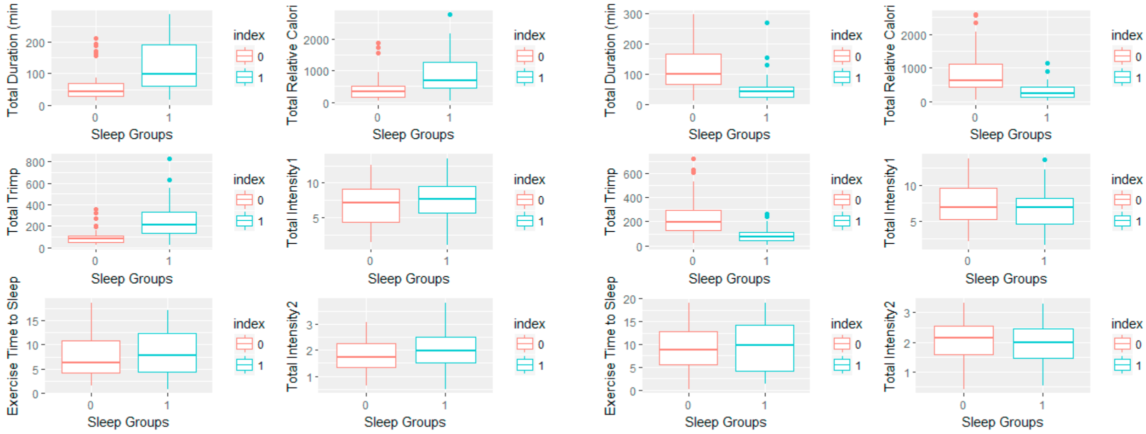

Fig. 2. Exercise variables in good and bad days for users group A (first two columns) and users group B (last two columns).

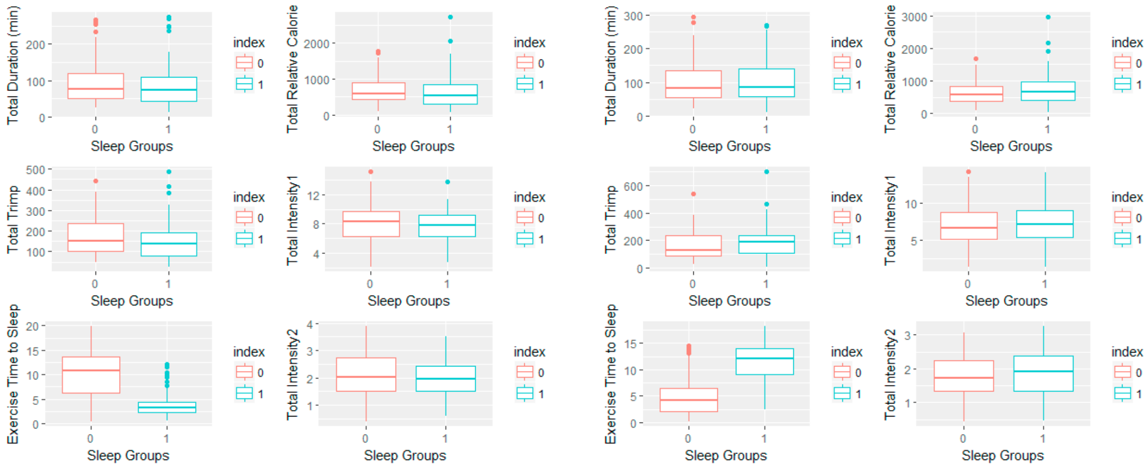

Fig. 3. Exercise variables in good and bad days for users group C (left two columns) and users group D (right two columns).

07, 6.45e-12, 3.458e-11, 0.4119, 0.3186, 3.012e-09, 2.661e-09, 2.838e-08, 0.5289, 0.4027, 0.5778). The p-values of the two-samples wilcoxon test for Group B users are similar to the Group A users. Thus, for those 56 users, the exercise intensity and exercise time do not make a significant difference in the good and bad sleep groups while the exercise duration, relative consumption and trimp have significant difference between the two sleep groups. The left two columns picture in figure 3 are boxplots for 27 users (Group C) and the right two columns pictures are boxplots for another 27 users (Group D) exercise information in good and bad sleep days. For those users, exercise time plays an important role in their good and bad sleep nights. The p-values of the two-samples wilcoxon test for group C users are (0.4664, 0.1037, 0.1819, 0.2579, 0.3559, 0.4114, 0.0645, 0.1196, 0.2142, 0.2442, 1.392e-13) and for group D users are (0.5982, 0.1442, 0.1856, 0.483, 0.1225, 0.05167, 0.004726, 0.01347, 0.6205, 0.3733, 1.662e-15). For those 54 users, the difference in exercise intensity are not significant between those two groups, but exercise time has significant difference between those two sleep groups. The boundary p-value used in the test are set to 0.05. p-value<0.05 is regarded as the shift is not equal to 0.

From the above analysis, we observe that the association between exercise and sleep is very complex. For example, with exercise duration 100 minutes, some users with high probability of getting a good sleep night; some users with high probability of having a bad night; and for others, it might have a good or a bad night as the main factor might be the exercise time (when they do the exercise). Table 1 shows that building the good sleep generalized predictive model that can be used for every person based on the exercise is not the best choice. As we can see from Figure 2 and

Figure 3, grouping similar users and developing the predictive model within groups are more practical compared with building a generalized predictive model.

5.2. Results of a defined background group

We consider two situations of a defined background group and describe the corresponding results. For the selected 271 users, difference of exercise between the good and bad groups is weak. There exists exercise difference within user groups A-D. However, the backgrounds of group users are hard to capture. For example, the age of group A users ranges from 26 years old to 71 years old and bmi ranges from 19.82 to 34. For some users, the good sleep nights are associated with more exercise, but for some users the association is opposite. Because the background information may have a significant impact on the sleep as well, a subset users from 271 users with a defined background (male, age between 35 and 45, bmi between 18.5 and 25, and regular, frequent, and heavy training background) are selected to analyze the exercise difference between good and bad sleep nights for efficiency. Two situations are considered: efficiency in case 1 and efficiency in case 2.

Table 2 presents the mean value and the standard deviation of each exercise variable in good and bad sleep days, the mean value of each ratio, the mean values of the logarithm of each ratio, and the p-value of the t-test with alternative greater for each exercise variable for sleep efficiency. Log values of the ratios are used in order to satisfy underlying assumptions of t-tests, which require the normal distribution. The p-values for efficiency in case 1 for the variables TR, TD, TRC, T, D, RC and TOS are less than 0.05, which means the difference has statistical significance. The p-values for efficiency in case 2 for the variables TR, TD, TRC, T, D, and RC are also less than 0.05. Therefore, compared with bad sleep efficiency days, sleep good efficiency days are associated with a slightly longer duration, more trimp, higher relative calories consumption, earlier exercise time. Table 2 shows which exercise variables can be used when building the personal model.

Table 2. Defined background group users' statistics and test results for sleep efficiency

Efficiency (case1)	Mean±SD (Bad Days)	Mean±SD (Good Days)	Mean (Ratios)	Mean (log(Ratios))	p-value of t-test greater
TR	163.92±81.22	216.21±137.97	1.58	0.29	0.02507
TD(m)	88.41±46.36	105.95±48.82	1.57	0.25	0.04379
TRC	741.15±399.47	923.82±553.96	1.82	0.30	0.04483
TI1	8.42±2.43	8.86±2.29	1.13	0.06	0.2571
TI2	1.96±0.46	2.10±0.54	1.11	0.07	0.1763
T	132.73±58.73	204.80±134.95	1.77	0.40	0.00565
D(m)	71.97±34.25	98.16±47.81	1.67	0.33	0.01336
RC	597.72±282.71	878.09±554.93	1.99	0.39	0.01995
I1	8.21±2.36	8.85±2.30	1.16	0.08	0.1898
I2	1.94±0.48	2.14±0.52	1.15	0.10	0.08466
TOS(h)	6.91±2.58	9.20±3.07	1.63	0.30	0.02915
Efficiency (case2)	Mean±SD (Bad Days)	Mean±SD (Good Days)	Mean (Ratios)	Mean (log(Ratios))	p-value of t-test greater
TR	80.80±85.49	97.11 ±91.42	2.81	0.47	0.03162
TD(m)	41.92±45.60	44.70±38.51	2.60	0.43	0.01072
TRC	359.45±375.69	386.19±315.52	3.91	0.48	0.0474
T	60.06±55.17	88.20±74.94	3.02	0.58	0.04019
D(m)	30.93±28.57	39.78±28.70	2.67	0.51	0.01573
RC	267.18±237.16	351.96±257.26	4.07	0.57	0.02395

6. Conclusions and future directions

Sleep is often the longest behavior we perform every day, which allows our bodies to have a proper rest. Hence, it is critical to have efficient sleep. Though pervasive wearables are widely accepted for fitness tracking, which potentially improves the sleep quality, very limited research has been done about how wearables can be utilized for improving sleep quality. Our goals are addressing the challenges of does exercise level affect sleep quality and what is the best way of performing exercise in order to improve sleep quality. We present a comprehensive analysis utilizing statistical methods to investigate the difference of exercise level for the bad and good sleep nights with big real data sets collected from wearables.

Groups of users with larger sleep variation and exercise variation are selected for finding out if there is any exercise difference between the good and bad sleep days. The analysis shows that exercise level and sleep quality are quite personal. With a same exercise plan, some users can have a good sleep night, but some users can have a bad sleep night. The exercise difference between the good and bad sleep efficiency night is weak from the population level. The analysis of the defined background group users indicates that exercise has stronger affects on sleep efficiency than sleep continuity. Variables exercise duration, relative calories consumption, trimp, exercise time are statistically important. Compared with bad sleep efficiency days, the exercise duration is longer, relative calorie consumption is larger, and the exercise time is earlier in good sleep efficiency days.

This research is one step towards fully integrated sleep related data analytic. The more interesting questions concerning the exercise and sleep data include: how to predict our sleep quality using our daily exercise data (exercise time, exercise duration, exercise intensity)? How shall we plan the daytime exercise in advance, in order to have a good sleep night? Our ultimate goal is to build a prediction model using exercise information as part of predictors. The analysis results propose the hints and future directions as follows. 1) It is impractical to build a generalized predictive model to predict the sleep efficiency using the exercise information. 2) Exercise duration, relative calories consumption, trimp and exercise time can be used as predictors in building personal predictive model. Exercise intensity does not have a significant impact on sleep quality. 3) Development of personalized predictive models for user groups with a similar background is an optimized solution. 4) Building personalized models with tracked exercise and sleep data using statistical methods or deep learning such as LSTM would be a most promising direction. Instead of building a generalized predictive model on the Cloud, it would be wise for building personalized models on own devices.

References

[1] A. Sathyanarayana, J. Srivastava, and L. Fernandez-Luque, "The science of sweet dreams: predicting sleep efficiency from wearable device data," Computer, vol. 50, no. 3, pp. 30–38, March 2017.

[2] Y. Bai, B. Xu, Y. C. Ma, G. D. Sun, and Y. Zhao, "Will you have a good sleep tonight?: sleep quality prediction with mobile phone," Proceedings of the 7th International Conference on Body Area Networks. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), ICST, Brussels, Belgium, Belgium, pp. 124–130, 2012.

[3] G. S. Passos, D. Poyares, M. G. Santana, C. V. ĎAurea, S. D. Youngstedt, S. Tufik, and M. T. de Mello, "Effects of moderate aerobic exercise training on chronic primary insomnia," Sleep Medicine, Vol. 12, Iss. 10, pp. 1018–1027, 2011.

[4] T. Myllymäki, H. Kyrölinen, K. Savolainen, L. Hokka, R. Jakonen, T. Juuti, K. Martinmäki, J. Kaartinen, M. L. Kinnunen, and H. Rusko," Effects of vigorous late-night exercise on sleep quality and cardiac autonomic activity," J. Sleep Res., pp. 146–153, March 2011.

[5] A. Sathyanarayana, F. Ofli, L. Fernandez-Luque, J. Srivastava, A. Elmagarmid, T. Arora, and S. Taheri. Robust Automated Human Activity Recognition and Its Application to Sleep Research . International Conference on Data Mining Workshops. Barcelona, Spain: IEEE. 2016.

[6] E. Hoque, R. F. Dickerson, and J.A. Stankovic. Monitoring body positions and movements during sleep using WISPs. In Wireless Health 2010 (WH '10). ACM, New York, NY, USA, pp. 44-53. 2010.

[7] R. J. Cole, D. F. Kripke, W. Gruen, D. J. Mullaney, and J. C. Gillin, "Automatic sleep/wake identification from wrist activity," Sleep, vol. 15, pp. 461–469, 1992.

[8] M. Enomoto, T. Endo, K. Suenaga, N. Miura, Y. Nakano, S. Kohtoh, Y. Taguchi, S. Aritake, S. Higuchi, M. Matsuura, K. Takahashi, and K. Mishima, "Newly developed waist actigraphy and its sleep/wake scoring algorithm," Sleep and Biological Rhythms, vol. 7, pp. 17–22, 2009.

[9] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, "Activity recognition from accelerometer data," Proceedings of the 17th conference on Innovative applications of artificial intelligence (IAAI'05), vol. 5, pp. 1541–1546, July 2005.

[10] J. Baek, G. Lee, W. Park, and B. J. Yun, "Accelerometer signal processing for user activity detection, " International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, pp. 610–617, Springer Berlin Heidelberg, September 2004.

[11] A. A. Parent, V. Guadagni, and M.J. Poulin, "Validation of novel algorithm for recognition of sleep or awake states," Proceedings of the Canadian Society for Exercise Physiology 51st Annual General Meeting, pp 87, 2018.

[12] AK. Pesonen and L. Kuula, "The Validity of a New Consumer-Targeted Wrist Device in Sleep Measurement: An Overnight Comparison Against Polysomnography in Children and Adolescents," J. Clin Sleep Med., vol. 14, pp. 585–591, Apr, 2018.

[13] P. CelkaEmail, A. Martinmaki, T. Korhonen, N. Santaniemi, and J. Virkkala, "Sleep-wake detection and computation of sleep continuity from a wrist unit in children, adolescents and adults," In: Eskola H., Visnen O., Viik J., Hyttinen J. (eds) EMBEC NBC 2017, EMBEC 2017, NBC 2017. IFMBE Proceedings, vol. 65, Springer, Singapore.

[14] K. Martinmki, T. Korhonen, K. Synjkangas, N. Santaniemi, T. Posio, P. Puolakanaho, and E. Pekonen, "Measuring and estimating sleep quality," Justia Patents, August 2018.