Published in final edited form as:

Sleep Med. 2020 March ; 67: 217–224. doi:10.1016/j.sleep.2019.11.1259.

Sleep heart rate variability assists the automatic prediction of long-term cardiovascular outcomes

Lulu Zhang^a, Huili Wu^b, Xiangyu Zhang^c, Xinfu Wei^d, Fengzhen Hou^{a,*}, Yan Ma^e

^aKey Laboratory of Biomedical Functional Materials, School of Science, China Pharmaceutical University, Nanjing 210009, China

^bENT Sleep Monitoring Center, Coal General Hospital, Beijing 100028, China

^cSEU-lenovo S-H-E Wearable Intelligent Monitoring Lab, State Key Laboratory of Bioelectronics, The School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China

^dDepartment of Otolaryngology, Coal General Hospital, Beijing 100028, China

^eDivision of Interdisciplinary Medicine and Biotechnology, Department of Medicine, Beth Israel Deaconess Medical Center/Harvard Medical School, Boston, United States

Abstract

Objective: We aimed to investigate the association between sleep HRV and long-term cardiovascular disease (CVD) outcomes, and further explore whether HRV features can assist the automatic CVD prediction.

Methods: We retrospectively analyzed polysomnography (PSG) data obtained from 2111 participants in the Sleep Heart Health Study, who were followed up for a median of 11.8 years after PSG acquisition. During follow-up, 1252 participants suffered CVD events (CVD group) and 859 participants remained CVD-free (non-CVD group). HRV measures, derived from time-domain and frequency-domain, were calculated. Regression models were created to determine the independent predictor for long-term CVD outcomes, and to explore the association between HRV and CVD latency. Furthermore, based on HRV and other clinical features, a model was trained to automatically predict CVD outcomes using the eXtreme Gradient Boosting algorithm.

Results: Compared with the non-CVD group, decreased HRV during sleep was found in the CVD group. HRV, particularly its component of high frequency (HF), was demonstrated to be independent predictor of CVD outcomes. Moreover, normalized HF was positively correlated with

*Corresponding author.: houfz@cpu.edu.cn (F. Hou).

Author contribution statement

Lulu Zhang: Investigation, Formal analysis, Validation, Roles/Writing - original draft.

Huili Wu: Conceptualization, Project administration.

Xiangyu Zhang: Data curation, Software.

Xinfu Wei: Project administration.

Fengzhen Hou: Data curation, Project administration, Supervision, Writing - review & editing.

Yan Ma: Methodology.

Conflict of interest

None.

The ICMJE Uniform Disclosure Form for Potential Conflicts of Interest associated with this article can be viewed by clicking on the following link: <https://doi.org/10.1016/j.sleep.2019.11.1259>.

CVD latency. The proposed prediction model achieved a total accuracy of 75.3%, in which sleep HRV features served as a supplement to the well-recognized CVD risk factors, such as aging, adiposity and sleep disorders.

Conclusions: Association between sleep HRV and long-term CVD outcomes was demonstrated here, suggesting that altered HRV during sleep might occur many years prior to the onset of CVD. Machine learning models, combining sleep HRV and other clinical characteristics, should be promising in the early prediction of CVD outcomes.

Keywords

Heart rate variability; Cardiovascular diseases; Sleep; Machine learning

1. Introduction

Cardiovascular disease (CVD) is a major cause of mortality, claiming 33% of all deaths worldwide [1]. Early detection and control of CVD risk factors are therefore greatly encouraged for health management. Heart rate variability (HRV) is a useful term which is widely applied to describe the variation of intervals between two successive heart beats [2], such intervals are often called RR intervals. Since first was proposed, HRV has been rapidly adopted as a non-invasive method to study the cardiac autonomic modulation [3]. Evidences of an association between HRV and CVD, such as myocardial infarction [4,5], stroke [6], angina [7,8], coronary heart disease [9], coronary artery disease [10–12] and sudden cardiac death [13], have been reported. Furthermore, studies have put forward that HRV has predictive value for CVD outcomes [14–16]. For the general population, reduced HRV has a high correlation with incident coronary heart disease and death.

Among studies focused on the association between HRV and CVD, HRV signals used were usually acquired during daytime on those awake participants. Sleep, a totally different physiological condition from daytime awareness, constitutes a fundamental behavioral mechanism for all living organisms. For humans in particular, numerous evidences show that sleep is vital on maintaining physical health [17,18], cognitive function [19,20], recovery [21], memory [22], mood [23] and daytime functioning [24,25]. Furthermore, sleep, or sleep-related mechanisms, impose regulatory control over the cardiovascular system, since modulation of autonomic nervous system (ANS) are profoundly influenced by the sleep-wake cycle [26,27]. Eguchi et al., demonstrated that HRV during sleep was independently associated with an increased risk of CVD in patients with type 2 diabetes [28]. Vanoli et al., found that sleep HRV was highly relevant to the identification of autonomic derangements which may account for a higher risk of lethal events after myocardial infarction [29]. Recently, reduced parasympathetic modulation during sleep been reflected by the high frequency component of HRV was reported to be one potential mechanism underlying the increased prevalence of CVD among veterans with posttraumatic stress disorder [30]. Although the association between HRV and CVD has been well recognized, it is still largely unknown that whether HRV features, especially during sleep, can assist the prediction of the occurrence of CVD events after years of latency.

Several assessment systems for CVD risk have been proposed to predict individual CVD events, such as Framingham risk score [31], Reynolds risk score [32,33], QRISK2 risk score [34] and the prediction algorithm which is recommended by the American Heart Association/American College of Cardiology (ACC/AHA) [35]. Typical risk factors in these systems include age, systolic blood pressure, total and high-density lipoprotein cholesterol, smoking, hypertension and diabetes status. By means of such assessment systems, a large number of individuals at risk of CVD fail to be detected while some not at risk are given preventive treatment unnecessarily [36], new approaches are therefore still in demand to improve the accuracy of CVD prediction. Machine learning (ML) is a subset of artificial intelligence in the field of computer science that allows computers to use data to learn [37]. ML approaches have been widely applied in disease diagnosis and prognosis which achieving satisfied accuracy [38]. Recently, by adopting routine clinical data [36], ML algorithms were employed in a large-scale study to predict the first CVD event after 10 years. In comparison with the AHA/ACC risk prediction algorithm, results shows that the accuracy of CVD risk prediction is significantly improved by the application of ML [36]. However, to our best knowledge, few study employs HRV indices in ML models for CVD event prediction. Therefore, more explorations are required in predicting CVD outcomes automatically by using ML algorithms and HRV features.

It's worth noting that sleep is not a stable, but rather a complex process. For general population, a nocturnal sleep comprises four to six sleep cycles, which normally begin with light sleep, continue to deep sleep and end in rapid eye movement (REM) sleep [39]. According to the rules introduced by Rechtschaffen and Kales (R & K rules) [40], non-REM (NREM) sleep can be further classified into four stages (Stage 1, 2, 3, 4), and the current American Academy of Sleep Medicine rules combined Stages 3 and 4 and termed it N3 [41]. During sleep, ANS function is influenced by sleep state [26], resulting in an alteration of HRV across different sleep stages [42–44]. For healthy adults, HRV was observed to be decreased during NREM sleep with augmented parasympathetic modulation, and increased during REM sleep with a reduction in parasympathetic modulation [42]. Therefore, it is a rational way to comprehensively evaluate HRV in different sleep stages when using sleep HRV to predict CVD outcomes.

In the present study, we retrospectively analyze HRV data derived from an open-access database. On one hand, we target to investigate whether there is an association between sleep HRV and long-term CVD outcomes. On the other hand, we aim to find out whether ML model based on sleep HRV data and clinical characteristics can predict long-term CVD outcomes.

2. Participants and methods

2.1. Participants

The HRV data used in this study were obtained from the Sleep Heart Health Study (SHHS) database [45]. The SHHS is a multi-center cohort study which aims to investigate whether sleep-disordered breathing is associated with an increased risk of cardiovascular events. In all, 6441 men and women aged 40 years and older were enrolled between November 1, 1995 and January 31, 1998 to take part in SHHS for a baseline polysomnography (PSG)

monitoring using the Compumedics PS polysomnograph at home. Before PSG monitoring, the medical history, recent medications, and life habits of the participants were recorded in detail. Furthermore, a follow-up visit was conducted to monitor and adjudicate the CVD outcomes (eg, stroke, heart attack) for each participant between baseline and 2011. Key outcomes for SHHS include incident or recurrent CVD events or diagnoses occurring subsequent to the baseline PSG, such as hospitalized acute myocardial infarction, coronary surgical intervention, congestive heart failure, any coronary heart disease, any cardiovascular disease and angina pectoris. The CVD latency, defined as the duration between the baseline PSG recording and to the first CVD event during follow-up, was also recorded. In result, PSG recordings and related outcome data obtained from 5802 participants were available online.

As the primary purpose of the present study was to investigate the association between sleep HRV and CVD outcomes, we focused on the participants who were free of CVD at baseline and had a normal sleep during the study night. Thus, as shown in Fig. 1, participants were excluded if they had a self-reported or MD-reported CVD before baseline PSG monitoring ($n = 1795$), or had extremely low sleep efficiency during the study night (defined as a value less than the mean minus double of standard deviation of all the 5802 subjects, $n = 146$). Moreover, benzodiazepines, tricyclic, and non-tricyclic antidepressants were reported to play a potential role in ANS and HRV [46], participants who used such medicines within two weeks prior to PSG monitoring ($n = 408$) were also excluded. Furthermore, as HRV features from different sleep stages (REM, Stage 2 and N3 sleep) and CVD risk factors, such as age, waist/hip ratio, body mass index (BMI), height, apnea hypopnea index (AHI), respiratory disturbance index (RDI), smoking status, lifetime cigarette smoke, alcohol intake, diabetes or hypertension were required in the present study, participants lack of such information were also excluded. Eventually, 2111 participants were included in the present study and further classified into two groups depending on whether the participant had at least one CVD event recorded during follow-up. As illustrated in Fig. 1, 1252 participants were included in the CVD group and 859 subjects were included in the non-CVD group. These included participants were followed for a median of 11.8 years (Q1–Q3, 11.1–12.4 years) until death or last contact.

The study protocol of the SHHS was approved by the institutional review board of each participating center. Each participant provided signed informed consent before the study. All methods were carried out in accordance with relevant guidelines and regulations. The current study analyzed de-identified data from the SHHS database, and did not involve a research protocol requiring approval by the relevant institutional review board or ethics committee.

2.2. Signal preprocessing

The PSG recordings included one channel electrocardiogram (ECG) data from a bipolar lead with a sampling rate of 125 Hz. A Butterworth band-pass filter (0.5–45 Hz) was first applied to the ECG recordings. Then the PanTompkins's method [47] was used to detect R-waves of the ECG. A time series of RR intervals, ie, the HRV signals, can therefore be obtained by calculating the time intervals between each pair of successive R-wave peaks. Finally, for

each participant, we picked up all the HRV segments with successive 5 min which had the same stage (REM, Stage 1, Stage 2 and N3 sleep).

For each HRV segment, artifacts or ectopic beats were directly eliminated with interpolation, resulting in segments consisting of normal-to-normal heartbeat (NN) intervals. The segments with a percent of artifacts or ectopic beats more than 10% were excluded from further study [48]. Since only 304 out of 2111 participants had at least one 5-min segment during Stage 1 sleep, in the present study, HRV segments during REM, Stage 2 and N3 sleep were taken into consideration. For all the included participants, the average number of segments was 10 during REM sleep, 22 during Stage 2 sleep and 9 during N3 sleep, respectively. Meanwhile, the maximum number of segments was 27 during REM sleep, 55 during Stage 2 sleep and 41 during N3 sleep, respectively. All the participants had at least one segment of 5-min HRV during REM, Stage 2 or N3 sleep.

2.3. HRV analysis

In the present study, traditional short-term HRV metrics derived from the analysis of time-domain and frequency-domain were adopted for each 5-min HRV segment [49]. Those metrics included the standard deviation of the NN intervals (SDNN), the square root of the sum of the squares of the differences between NN intervals (RMSSD), power in the low frequency range (0.04–0.15 Hz, LF), power in the high frequency range (0.15–0.4 Hz, HF), HF power in the normalized units (HFnorm, $HF/(HF + LF) \times 100$), and total power (TP).

2.4. Statistical analysis

Statistical analyses were performed using MATLAB (Mathworks Inc., Natick, MA) and SPSS version 22 (IBM SPSS Statistics, NY, United States). First, the difference in clinical characteristics between the CVD and non-CVD groups was assessed. Chi-square test was applied to categorical variables, such as gender, smoking status, diabetes, and hypertension. Meanwhile, the non-parametric Whitney test was used for continuous variables, such as age, waist/hip ratio, BMI, height, lifetime cigarette smoking, alcohol intake, AHI, and RDI. Between-group difference in HRV metrics was evaluated in a similar way. Second, logistic regression analysis was applied to identify the independent HRV metrics to CVD outcomes. Third, multivariable linear regression analysis was utilized to detect the HRV indices which were significantly correlated with CVD latency.

2.5. The predication of CVD outcomes based on ML

In the current study, a binary classifier based on the eXtreme Gradient Boosting (XGboost) algorithm was used to predict CVD outcomes during follow-up using HRV features and clinical characteristics. The XGBoost algorithm is an implementation of gradient boosting machines, first proposed by Chen and Guestrin [50]. Since its proposition, the XGBoost algorithm has been used widely by data scientists to achieve state-of-the-art results in many challenges [50]. XGBoost is a decision-tree based algorithm and a XGBoost model usually consists of a number of classification and regression trees (CARTs).

By using a given training dataset with n examples which consist of m features and one target label, the construction of a CART is aimed to map each example to a continuous score on a

leaf of the tree, called a prediction of the example. The main idea underlying the XGBoost algorithm is to combine a high number of ‘weak learners’ with low accuracy into a ‘strong learner’ [51] by establishing a series of CARTs iteratively. The first CART in a XGBoost model is trained to fit the whole space of the training dataset [52]. For each example in the training dataset, the difference between the prediction of the first CART and the target, called the residual, is then computed. To overcome the shortcomings of the first learner, the second CART is trained with a refreshed training dataset which utilizes the residual as a new target, as shown in Fig. 2. Such refreshing of targets with residuals, ie, the residual of the i -th CART is served as the target of the $(i+1)$ -th CART, and training of new learners is repeated until some stopping criterion is satisfied [52], resulting in a final model with a number of CARTs. By summing up the score in the corresponding leaves of all CARTs, we can get the final prediction of the XGBoost model.

For each participant, feature vector of the XGBoost model consisted of 18 HRV metrics and 11 clinical characteristics, including age, waist/hip ratio, BMI, height, AHI, RDI, smoking status, lifetime cigarette smoke, alcohol intake, diabetes and hypertension. The former was obtained according to the calculation of HRV indices in REM, Stage 2 and N3 sleep, respectively. To improve the generalization ability of the model, five-fold cross validation was employed by randomly distributing the subjects into five equal subsets. Then the model was trained on four subsets and tested on the rest subset five times until all the subsets were tested. Five indices, accuracy, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of the classifier were used to evaluate the model’s performance and the performances in the five-fold cross validations were averaged. The model was trained on python 3.7.0 with the XGBoost 0.7 package (<https://pypi.org/project/xgboost/>).

3. Results

3.1. Baseline clinical characteristics of the included participants

Table 1 presents the baseline clinical characteristics of all included participants. CVD group had a significantly higher age, waist/hip ratio, BMI, height and RDI than the non-CVD group. Besides, compared with non-CVD group, CVD group had higher prevalence of diabetes and hypertension.

3.2. Between-group comparison of HRV metrics

The results of cross-sectional comparisons of HRV metrics are illustrated in Table 2. Compared with non-CVD group, a significantly ($p < 0.05$) decreased HF was found in CVD group regardless of sleep stages. Additionally, lower SDNN and LF during REM sleep, as well as decreased RMSSD and HFnorm during N3 sleep were found in CVD group.

3.3. Independent HRV predictors of CVD outcomes

Logistic regression models were conducted to determine the independent HRV measures for long-term CVD outcomes. As shown in Table 3, HF and HFnorm exhibit statistically significant ($p < 0.05$) effect on CVD prediction in all three stages, after adjusting clinical characteristics, ie, age, gender, waist/hip ratio, RDI, lifetime cigarette smoke, alcohol intake,

diabetes and hypertension. Furthermore, RMSSD, LF and TP during REM sleep, as well as SDNN, LF and TP during N3 sleep were also statistically significant ($p < 0.05$) predictors to CVD outcomes.

3.4. Correlation between HRV metrics and CVD latency

A sub-group analysis was conducted including participants ($n = 384$) in the CVD group for whose CVD latency (median: 5.8 years, Q1–Q3: 3.1–9.1 years) was available. Clinical characteristics (age, gender, waist/hip ratio, RDI, lifetime cigarette smoke, alcohol intake, diabetes and hypertension) and HRV metrics were included in the multivariable linear regression to explore the sensitive HRV indices correlated with CVD latency. As shown in Table 4, HFnorm was positively related with CVD latency after adjusting clinical characteristics during all the sleep states. However, no other HRV metrics correlated with CVD latency.

3.5. The results of XGBoost classifier

Binary classifiers based on the XGBoost algorithm and five-fold cross validation were trained and tested on all the included 2111 participants to determine whether there would be at least one CVD event recorded during the follow-up visit or not. The feature vector for each participant was constructed using 29 features, including clinical characteristics (age, waist/hip ratio, BMI, height, AHI, RDI, smoking status, lifetime cigarette smoke, alcohol intake, diabetes, and hypertension) and HRV metrics during REM, Stage 2 and N3 sleep. Such a feature vector was then fed into the models as their input. Among the trained five models, the best model predicted 231 CVD cases from 250 CVD cases, and 97 non-CVD cases from 171 non-CVD cases. As shown in Table 5, the model achieved an average accuracy of 75.3%, average sensitivity of 87.9%, and average specificity of 57%.

The importance of a feature (ie, score) in the XGBoost model, is measured based on the total times it was used to split the data across all CARTs [53]. Features with high score are considered to be more important in the model than those with low score. As a five-fold cross validation was used when training the model, for each feature, we used an average score of the five trained XGBoost models to measure its importance. Fig. 3 illustrated that clinical characteristics such as age, waist/hip ratio, height, AHI, BMI and RDI contributed considerably to the prediction of CVD outcomes. HRV metrics, in particular, HFnorm during all three sleep stages, HF during REM sleep and LF during N3 sleep also showed non-trivial effects on the automatic identification of potential CVD outcomes. However, hypertension, lifetime cigarette smoke, diabetes or smoking status had relatively small contributions to the model.

4. Discussion

In the present study, we investigated the association between sleep HRV and long-term CVD outcomes, and further adopted ML method to predict the outcomes based on HRV and other clinical features. Compared with non-CVD group, decreased HRV was observed during sleep in CVD group, in which participants have at least one CVD event during follow-up. Sleep HRV was further found to be independent predictor of CVD outcomes and positively

correlated with CVD latency. Moreover, our findings demonstrated that ML should be a promising tool for the prediction of CVD outcomes. Demographics (age, waist/hip ratio, height, and BMI), measurements of sleep-disordered breathing and sleep HRV characteristics (such as HFnorm during all three sleep stages, HF during REM sleep and LF during N3 sleep) were important features in the suggested ML model.

Aging, adiposity and sleep disorders are common risk factors for CVD outcomes [54–56]. Both the incidence of CVD and the levels of cardiovascular risk factors increase with age [57,58]. Adiposity, generally measured by waist/hip ratio or BMI, positively correlates with CVD [59–62]. Previous studies demonstrate that sleep-disordered breathing correlates with CVD morbidity and mortality [63]. In line with previous studies, indices (ie, age, waist/hip ratio, BMI, AHI and RDI) were vital features in the proposed CVD prediction model based on XGBoost algorithm. Moreover, we find that height is another important feature in the model, which is in consistence with previous observation that height is inversely correlated with coronary heart disease [64,65]. Although gender is a potential risk in coronary heart disease [58,66] and myocardial infarction [67], also medical history of diabetes [68,69] or hypertension [70,71] is associated with a high risk of CVD outcomes, those features contributed little to the prediction of CVD outcomes in the present study. Similarly, we find less importance of usual alcohol intake per day or smoking status to the proposed CVD prediction model.

The current study reveals an association between HRV (in particular, its HF component) and long-term CVD outcomes. A significant decline of HF was found in CVD group when comparing with non-CVD group. HF and HFnorm were demonstrated to be independent predictors of CVD outcomes in the logistic regression models. HFnorm was also demonstrated to be important in the proposed ML model. Physiologically, HF component of HRV is generally attributed to vagal activity [72,73]. Therefore, our study might suggest that a withdrawal of vagal activity occurs even several years before the onset of CVD event. Furthermore, the results of multivariable linear regression analysis show that HFnorm is positively correlated with CVD latency, which might indicate an association between increased vagal activity and risk levels of CVD outcomes. We believe that such a finding will bring achievement of early intervention in CVD.

The proposed XGBoost model, which utilized sleep HRV and clinical characteristics to construct its feature vector, provided a total accuracy of 75.3% in the prediagnosis of CVD events after years (median: 5.8 years) of latency, suggesting a promising application of ML in automatic prediction of long-term CVD outcomes. Although the common CVD risk factors, such as age, adiposity and sleep-disorder breathing, were most important in the XGBoost model, it is the Author's belief that HRV metrics cannot not overlooked. By using the clinical characteristics or the HRV metrics as their features respectively, we further constructed another two XGBoost models to predict long-term CVD outcomes. The total accuracy was decreased to 73.7% under the situation without HRV features, while the accuracy remained 58.8% when utilizing HRV features only. Our results indicate that the presence of sleep HRV cannot be overlooked by serving as a supplement to common CVD risk factor in long-term CVD prediction. While such presence seems trivial due to the

excessive CVD latency, future work can be tried on explore the sleep HRV in short-term prediction of CVD outcomes.

Although the current study elucidated the association between sleep HRV and long-term CVD outcomes in a large cohort, one limitation is that only conventional HRV metrics were considered. As nonlinear features of HRV have been widely recognized, it is worth including indices derived from HRV nonlinear dynamics analysis in the prediction of CVD outcomes. Besides, a comprehensive investigation shall be of significance by employing more kinds of ML algorithm and more clinical characteristics, such as systolic blood pressure and the level of serum total and high-density lipoprotein cholesterol.

5. Conclusions

For subjects with CVD risks, ANS alterations during sleep may present a long time prior to the onset of a CVD event. Such alterations can be captured by the changes in multiple HRV metrics, specially, decreased HF. A combination of sleep HRV measuring and ML techniques can assist the early prediction of CVD outcomes. Since ambulatory ECG monitoring is readily available and accessible in a clinical setting, large-scale screening to detect HRV alterations may be assistant in early diagnosis and interventions of adverse cardiovascular events.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 61401518, 31671006 and 61771251), Jiangsu Provincial Key R & D Program (Social Development) (Grant No. BE2015700 and BE2016773), and Natural Science Research Major Program in Universities of Jiangsu Province (Grant No. 16KJA310002). The authors would like to acknowledge the support team of the forum in the Sleep Heart Health Study for their detailed explanations and assistance in our use of the dataset.

References

- [1]. Roth GA, Johnson C, Abajobir A, et al. Global, regional, and national burden of cardiovascular diseases for 10 causes, 1990 to 2015. *J Am Coll Cardiol* 2017;70:1–25. [PubMed: 28527533]
- [2]. Pumpura J, Howorka K, Groves D, et al. Functional assessment of heart rate variability: physiological basis and practical applications. *Int J Cardiol* 2002;84:1–14. [PubMed: 12104056]
- [3]. Sztajzel J Heart rate variability: a noninvasive electrocardiographic method to measure the autonomic nervous system. *Swiss Med Wkly* 2004;134:514–22. [PubMed: 15517504]
- [4]. Milovanovic B, Furundzic D, Gligorijevic T, et al. Nonlinear parameters of heart rate variability and detection of high risk patients after myocardial infarction based on artificial intelligence analysis. *J Electrocardiol* 2018;53:e1.
- [5]. Oliveira NL, Ribeiro F, Alves AJ, et al. Heart rate variability in myocardial infarction patients: effects of exercise training 2013;32:687–700.
- [6]. Lees T, Shad-Kaneez F, Simpson AM, et al. Heart rate variability as a biomarker for predicting stroke, post-stroke complications and functionality. *Biomark Insights* 2018;13.
- [7]. Lanza GA, Cianflone D, Rebuzzi AG, et al. Prognostic value of ventricular arrhythmias and heart rate variability in patients with unstable angina 2006;92: 1055–63.
- [8]. Huang J, Sopher SM, Leatham E, et al. Heart rate variability depression in patients with unstable angina. *Am Heart J* 1995;130:772–9. [PubMed: 7572585]
- [9]. Dekker JM, Crow RS, Folsom AR, et al. Low heart rate variability in a 2-minutes rhythm strip predicts risk of coronary heart disease and mortality from several causes: the ARIC Study. *Circulation* 2000;102:1239–44. [PubMed: 10982537]

- [10]. Yu LC, Lin I-M, Fan S-Y, et al. One-Year cardiovascular prognosis of the randomized, controlled, short-term heart rate variability biofeedback among patients with coronary artery disease. *Int J Behav Med* 2018;25:271–82. [PubMed: 29297147]
- [11]. Li H-R, Lu T-M, Cheng H-M, et al. Additive value of heart rate variability in predicting obstructive coronary artery disease beyond Framingham risk. *Circ J* 2015;80:494–501. [PubMed: 26701182]
- [12]. Bhattacharyya MR, Whitehead DL, Rakhit R, et al. Depressed mood, positive affect, and heart rate variability in patients with suspected coronary artery disease. *Psychosom Med* 2008;70:1020–7. [PubMed: 18941130]
- [13]. Dougherty CM, Burr RL. Comparison of heart rate variability in survivors and nonsurvivors of sudden cardiac arrest. *Am J Cardiol* 1992;70:441–8. [PubMed: 1642181]
- [14]. Dekker JM, Schouten EG, Klootwijk P, et al. Heart rate variability from short electrocardiographic recordings predicts mortality from all causes in middle-aged and elderly men. The Zutphen Study. *Am J Epidemiol* 1997;145: 899–908. [PubMed: 9149661]
- [15]. Mäkikallio TH, Tapanainen JM, Tulppo MP, et al. Clinical applicability of heart rate variability analysis by methods based on nonlinear dynamics. *Card Electrophysiol Rev* 2002;6:250–5. [PubMed: 12114847]
- [16]. Huikuri HV, Stein PK. Heart rate variability in risk stratification of cardiac patients. *Prog Cardiovasc Dis* 2013;56:153–9. [PubMed: 24215747]
- [17]. Spiegel K, Tasali E, Leproult R, et al. Effects of poor and short sleep on glucose metabolism and obesity risk. *Nat Rev Endocrinol* 2009;5:253–61. [PubMed: 19444258]
- [18]. Patel SR, Hu FB. Short sleep duration and weight gain: a systematic review. *Obesity* 2012;16:643–53.
- [19]. Yeonsu S, Terri B, Kristine Y, et al. Relationships between sleep stages and changes in cognitive function in older men: the MrOS sleep study. *Sleep* 2015;38:411–21. [PubMed: 25325465]
- [20]. Blackwell T, Yaffe K, Ancoliisrael S, et al. Poor sleep is associated with impaired cognitive function in older women: the study of osteoporotic fractures. *J Gerontol* 2006;61:405–10.
- [21]. Faraut B, Boudjeltia KZ, Vanhamme L, et al. Immune, inflammatory and cardiovascular consequences of sleep restriction and recovery. *Sleep Med Rev* 2012;16:137–49. [PubMed: 21835655]
- [22]. Born J System consolidation of memory during sleep. *Psychol Res* 2012;76: 192–203. [PubMed: 21541757]
- [23]. Ford DE, Cooper-Patrick L. Sleep disturbances and mood disorders: an epidemiologic perspective. *Depress Anxiety* 2010;14:3–6.
- [24]. Ma Y, Hou F, Yang AC, et al. Symbolic dynamics of electroencephalography is associated with the sleep depth and overall sleep quality in healthy adults. *Phys A Stat Mech Appl* 2018;513:22–31.
- [25]. McCrae CS, Rowe MA, Tierney CG, et al. Sleep complaints, subjective and objective sleep patterns, health, psychological adjustment, and daytime functioning in community-dwelling older adults. *J Gerontol* 2005;60:182–9.
- [26]. Zambotti MD, Trinder J, Silvani A, et al. Dynamic coupling between the central and autonomic nervous systems during sleep: a review. *Neurosci Biobehav Rev* 2018;90.
- [27]. De ZM, Trinder J, Silvani A, et al. Dynamic coupling between the central and autonomic nervous systems during sleep: a review 2018;90:84–103.
- [28]. Kazuo E, Schwartz JE, Pickering TG, et al. Increased heart rate variability during sleep is a predictor for future cardiovascular events in patients with type 2 diabetes. *Hypertens Res Offic J Jpn Soc Hypertens* 2010;33:737–42.
- [29]. Vanoli E, Adamson PB, Pinna GD, et al. Heart rate variability during specific sleep stages. A comparison of healthy subjects with patients after myocardial infarction. *Circulation* 1995;91:1918–22. [PubMed: 7895347]
- [30]. Ulmer CS, Hall MH, Dennis PA, et al. Posttraumatic stress disorder diagnosis is associated with reduced parasympathetic activity during sleep in US veterans and military service members of the Iraq and Afghanistan wars. *Sleep* 2018;41:174.

- [31]. Marma AK, Lloyd-Jones DM. Systematic examination of the updated Framingham heart study general cardiovascular risk profile. *Circulation* 2009;120:384–90. [PubMed: 19620502]
- [32]. Chua D, Shalansky SJ, Legal MG, et al. Incorrect wording in: development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds risk score. *J Am Med Assoc* 2007;297:1433.
- [33]. Ridker PM, Paynter NP, Nader R, et al. C-reactive protein and parental history improve global cardiovascular risk prediction: the Reynolds Risk Score for men. *Circulation* 2008;118:2243–51. [PubMed: 18997194]
- [34]. Julia HC, Carol C, Yana V, et al. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. *BMJ* 2008;336: 1475–82. [PubMed: 18573856]
- [35]. Goff DC Jr, Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA guideline on the assessment of Cardiovascular risk: a report of the American College of Cardiology/American heart association task force on practice guidelines. *J Am Coll Cardiol* 2014;63:2886.
- [36]. Weng SF, Reps J, Kai J, et al. Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS One* 2017;12:e0174944. [PubMed: 28376093]
- [37]. Serra A, Galdi P, Tagliaferri R. Machine learning for bioinformatics and neuroimaging. *Wiley Interdisciplinary Reviews Data Mining & Knowledge Discovery*; 2018 p. e1248.
- [38]. Safdar S, Zafar S, Zafar N, et al. Machine learning based decision support systems (DSS) for heart disease diagnosis: a review. *Artif Intell Rev* 2017: 1–27.
- [39]. Hou FZ, Li FW, Wang J, et al. Visibility graph analysis of very short-term heart rate variability during sleep. *Phys A Stat Mech Appl* 2016;458:140–5.
- [40]. Wolpert EA. A manual of standardized terminology and scoring system for sleep stages of human subjects. *Arch Gen Psychiatr* 1969;20:246.
- [41]. Danker-Hopfe H, Anderer P, Zeitlhofer J, et al. Interrater reliability for sleep scoring according to the Rechtschaffen & Kales and the new AASM standard. *J Sleep Res* 2010;18:74–84.
- [42]. Zemaityte D, Varoneckas G, Sokolov E. Heart rhythm control during sleep. *Psychophysiology* 2010;21:279–89.
- [43]. Bonnet MH, Arand DL. Heart rate variability: sleep stage, time of night, and arousal influences. *Electroencephalogr Clin Neurophysiol* 1997;102:390–6. [PubMed: 9191582]
- [44]. Berlad II, Shlitrer A, Ben-Haim S, et al. Power spectrum analysis and heart rate variability in Stage 4 and REM sleep: evidence for state-specific changes in autonomic dominance. *J Sleep Res* 2010;2:88–90.
- [45]. Quan SF, Howard BV, Iber C, et al. The sleep heart health study: design, rationale, and methods. *Sleep* 1997;20:1077–85. [PubMed: 9493915]
- [46]. Licht CMM, de Geus EJC, van Dyck R, et al. Longitudinal evidence for unfavorable effects of antidepressants on heart rate variability. *Biol Psychiatry* 2010;68:861–8. [PubMed: 20843507]
- [47]. Pan J, Tompkins WJ. A real-time QRS detection algorithm. *IEEE Trans Biomed Eng* 2007;BME-32:230–6.
- [48]. Mäkikallio TH, Ristimäe T, Airaksinen KE, et al. Heart rate dynamics in patients with stable angina pectoris and utility of fractal and complexity measures. *Am J Cardiol* 1998;81:27–31. [PubMed: 9462601]
- [49]. Malik M, Bigger JT, Camm AJ, et al. Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Task force of the European society of Cardiology and the North American society of pacing and electrophysiology. *Circulation* 1996;93:1043–65. [PubMed: 8598068]
- [50]. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. *Acm Sigkdd international conference on knowledge discovery & data mining*. 2016.
- [51]. Ren X, Guo H, Li S, et al. A novel image classification method with CNNXGBoost model. *Int Workshop Digit Watermark* 2017;10431:378–90.
- [52]. Urraca R, Martinez-De-Pison E, Sanz-Garcia A, et al. Estimation methods for global solar radiation: case study evaluation of five different approaches in central Spain. *Renew Sustain Energy Rev* 2017;77:1098–113.

- [53]. https://xgboost.readthedocs.io/en/latest/python/python_api.html#. [Accessed 10 November 2019].
- [54]. Newman AB, Nieto FJ, Guidry U, et al. Relation of sleep-disordered breathing to cardiovascular disease risk factors: the Sleep Heart Health Study. *Am J Epidemiol* 2001;154:50–9. [PubMed: 11434366]
- [55]. Valderrama AL, Loustalot F, Gillespie C, et al. Million hearts: strategies to reduce the prevalence of leading cardiovascular disease risk factors—United States, 2011. *MMWR Morb Mortal Wkly Rep* 2011;60:1248–51. [PubMed: 21918495]
- [56]. Erhardt L Cigarette smoking: an undertreated risk factor for cardiovascular disease. *Atherosclerosis* 2009;205:23–32. [PubMed: 19217623]
- [57]. Nandi B, Kundu K, Banerjee N, et al. Myocardial infarction and coronary deaths in the World Health Organization MONICA Project. Registration procedures, event rates, and case-fatality rates in 38 populations from 21 countries in four continents. *Circulation* 1994;90:583–612. [PubMed: 8026046]
- [58]. Castelli WP. Epidemiology of coronary heart disease: the Framingham study. *Am J Med* 1984;76:4–12. [PubMed: 6702862]
- [59]. Hubert HB, Feinleib M, McNamara PM, et al. Obesity as an independent risk factor for cardiovascular disease: a 26-year follow-up of participants in the Framingham Heart Study. *Circulation* 1983;67:968–77. [PubMed: 6219830]
- [60]. Wilson PWF, D'Agostino RB, Lisa S, et al. Overweight and obesity as determinants of cardiovascular risk: the Framingham experience. *Arch Intern Med* 2002;162:1867–72. [PubMed: 12196085]
- [61]. Sowers JR. Obesity as a cardiovascular risk factor. *Am J Med* 2003;115:37–41.
- [62]. Welborn TA, Dhaliwal SS, Bennett SA. Waist:hip ratio is the dominant risk factor predicting cardiovascular death in Australia. *Med J Aust* 2003;179.
- [63]. Stopford E, Ravi K, Nayar V. The association of sleep disordered breathing with heart failure and other cardiovascular conditions. *Cardiol Res Pract* 2013;2013.
- [64]. Yarnell JW, Limb ES, Layzell JM, et al. Height: a risk marker for ischaemic heart disease: prospective results from the Caerphilly and Speedwell Heart Disease Studies. *Eur Heart J* 1992;13:1602–5. [PubMed: 1289087]
- [65]. Smith G, Davey Hart C, et al. Height and risk of death among men and women: aetiological implications of associations with cardiorespiratory disease and cancer mortality. *J Epidemiol Community Health* 2000;54: 97–103. [PubMed: 10715741]
- [66]. Kuhn FE, Rackley CE. Coronary artery disease in women. Risk factors, evaluation, treatment, and prevention. *Arch Intern Med* 1993;153:2626–36. [PubMed: 8250659]
- [67]. Njølstad I, Arnesen E, Lund-Larsen PG. Smoking, serum lipids, blood pressure, and sex differences in myocardial infarction. A 12-year follow-up of the Finnmark Study. *Circulation* 1996;93:450–6. [PubMed: 8565161]
- [68]. Kannel WB, McGee DL. Diabetes and cardiovascular disease: the Framingham study. *JAMA* 1979;241:2035–8. [PubMed: 430798]
- [69]. Stamler J, Vaccaro O, Neaton JD, et al. Diabetes, other risk factors, and 12-yr cardiovascular mortality for men screened in the Multiple Risk Factor Intervention Trial. *Diabetes Care* 1993;16:434–44. [PubMed: 8432214]
- [70]. Levy D, Larson MG, Vasan RS, et al. The progression from hypertension to congestive heart failure. *JAMA* 1996;275:1557–62. [PubMed: 8622246]
- [71]. Kannel WB. Hypertension as a risk factor for cardiac events: epidemiologic results of long-term studies. *J Cardiovasc Pharmacol* 1993;21:S27–37. [PubMed: 7692148]
- [72]. Akselrod S, Gordon D, Ubel FA, et al. Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control. *Science* 1981;213:220–2. [PubMed: 6166045]
- [73]. Pomeranz B, Macaulay RJ, Caudill MA, et al. Assessment of autonomic function in humans by heart rate spectral analysis. *Am J Physiol* 1985;248:151–3.

Fig. 1.
Flowchart of the inclusion of participants.

Fig. 2.

Schematic diagram of a XGBoost model which is comprised of n classification and regression trees (CARTs) used for in a prediction task. The dataset is firstly divided into training set (80%) and testing set (20%). Each example in the dataset consists of m features and a targeting label. The first CART (ie, CART 1) is trained to fit the whole space of the training set. In CART 1, the difference between the i -th example's prediction (denoted as $P1_i$) and target, called the residual (denoted as $R1_i$), is computed. Then, CART 2 is trained with a refreshed training set which utilizes $R1$ as a new target. Such refreshing of targets and training of more CARTs is repeated until some stopping criterion is satisfied, resulting in a final model with n CARTs. When testing the model, the final prediction (denoted as y-predict) was the sum of the individual predictions of n CARTs.

Fig. 3.

The importance of features used in the XGBoost classifier. Features with suffix _REM, _Stage2, or _N3, represents for HRV features during REM sleep, Stage 2 sleep or N3 sleep, respectively.

Table 1

Baseline clinical characteristics of subjects involved in this study.

	CVD	non-CVD	<i>p</i>
No. subjects(n)	1252	859	
Age (years)	62 [57,69]	60 [50,73]	<0.001 *
Gender (male/female)	499/753	316/543	0.155
Waist/hip ratio	94.7 [88.8,99.1]	89.4 [81.96,2]	<0.001 *
BMI (kg/m ²)	27.9 [25.2,31.3]	27.2 [24.4,30.5]	0.002 *
Height (centimeters)	166 [160,173]	165 [158,173]	0.014 *
AHI (events/hour)	9.06 [3.83, 18]	8.12 [3.35, 16.7]	0.261
RDI (events/hour)	28.5 [18.3,41.8]	26.1 [17.1,38.9]	0.008 *
Smoking status (never/current/former)	614/113/525	447/62/350	0.213
Lifetime cigarette smoke (packs/year)	0 [0,18.8]	0 [0,14.4]	0.054
Alcohol intake (drinks/day)	0 [0,3]	0 [0,3]	0.187
Diabetes (yes/no)	81/1171	25/834	<0.001 *
Hypertension (yes/no)	486/766	292/567	0.024 *

Note: Values are expressed as median [lower quartile, upper quartile] or the ratio as indicated. Waist/hip ratio, waist/hip ratio; BMI, body mass index; AHI, apnea hypopnea index, RDI, respiratory disturbance index.

* represents a significant difference, $p < 0.05$ (Chi-square test or non-parametric Whitney test).

Table 2
Results of HRV analysis for non-CVD and CVD groups during REM, Stage 2 and N3 sleep.

	SDNN (ms)	RMSSD (ms)	LF (ms ²)	HF (ms ²)	HFnorm (%)	TP (ms ²)
REM						
CVD	50.2 [29.5,62.1]	30.1 [21.3,52.5]	421 [240,687]	177 [80,455]	30.8 [19.3,47.4]	1958 [1216,2892]
Non-CVD	52.1 [39.6,65.9]	32.3 [22.1,57.1]	456 [229,864]	214 [91,570]	30.4 [19.1,49]	2059 [1221,3279]
<i>p</i>	0.041 *	0.053	0.014 *	0.009 *	0.487	0.097
Stage 2						
CVD	45.8 [35.6,57.3]	36.9 [25.7,56.5]	504 [282,871]	308 [142,614]	37.8 [25.5,51.8]	1699 [1008,2516]
non-CVD	47.5 [35.8,60]	39 [25.9,60.8]	529 [275,963]	348 [152,716]	39.4 [26.3,53.6]	1772 [1015,2729]
<i>p</i>	0.09	0.098	0.221	0.013 *	0.05	0.192
N3						
CVD	33.4 [23.7,47.4]	31.3 [20.9,53.9]	231 [123,441]	214 [94,546]	48.9 [33,65.3]	821 [431,1470]
non-CVD	34.8 [24.4,50.2]	33.8 [22.5,55.5]	243 [123,504]	262 [110,645]	49.9 [35.8,65.6]	887 [454,1611]
<i>p</i>	0.058	0.033 *	0.183	0.005 *	0.036 *	0.055

Note: Values are expressed as median [lower quartile, upper quartile].

* represents a significant difference, $p < 0.05$ (non-parametric Whitney test).

Table 3

Independent HRV predictors of long-term CVD outcomes.

Sleep stage	Indices (per 100 units)	OR	95%CI	<i>p</i>
REM sleep	RMSSD	0.681	0.479, 0.969	0.033
	LF	0.981	0.964, 0.998	0.029
	HF	0.972	0.952, 0.993	0.009
	HFnorm	0.599	0.363, 0.989	0.045
Stage 2 sleep	HF	0.980	0.962, 0.999	0.034
	HFnorm	0.529	0.311, 0.900	0.019
N3 sleep	SDNN	0.591	0.352, 0.994	0.047
	RMSSD	0.706	0.510, 0.978	0.036
	HF	0.983	0.967, 0.999	0.045
	HFnorm	0.604	0.376, 0.970	0.037
	TP	0.991	0.983, 0.999	0.034

Note: HRV metric and clinical characteristics were included in the multivariable analysis. OR = odds ratio; CI = confidence interval.

Table 4

Results of Multivariable linear regression analysis for the prediction of CVD latency.

Indices	β	95% CI	<i>p</i>
REM sleep:			
SDNN	-0.01	-7.702, 6.373	0.853
RMSSD	0.062	-1.576, 7.930	0.190
LF	-0.078	-0.387, 0.051	0.133
HF	0.081	-0.050, 0.476	0.113
HFnorm	0.105	0.142, 13.636	0.045*
TP	-0.072	-0.120, 0.020	0.164
Stage 2 sleep:			
SDNN	-0.021	-8.819, 5.822	0.688
RMSSD	0.055	-2.206, 7.541	0.283
LF	-0.055	-0.264, 0.081	0.296
HF	0.081	-0.052, 0.464	0.117
HFnorm	0.111	0.391, 14.861	0.039*
TP	-0.046	-0.123, 0.047	0.384
N3 sleep:			
SDNN	-0.002	-6.802, 6.544	0.970
RMSSD	0.052	-2.114, 6.663	0.309
LF	-0.072	-0.357, 0.063	0.169
HF	0.054	-0.096, 0.320	0.290
HFnorm	0.119	0.814, 13.717	0.027*
TP	-0.029	-0.129, 0.072	0.575

Note: HRV metric and clinical characteristics were included in multivariable analysis. CI = confidence interval.

Table 5

The performance of XGBoost prediction models.

	Accuracy (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
1-fold	72.4	87.2	50.9	72.2	73.1
2-fold	75.8	87.6	58.5	75.5	76.3
3-fold	75.8	84.8	62.6	76.8	73.8
4-fold	77.9	92.4	56.7	75.7	83.6
5-fold	74.7	87.3	56.6	74.3	75.6
average	75.3	87.9	57.0	74.9	76.5

Note: PPV = positive predictive value; NPV = negative predictive value.