0067

COMPUTATIONAL PHENOTYPING IN POLYSOMNOGRAPHY: USING INTERPRETABLE PHYSIOLOGY-BASED MACHINE LEARNING MODELS TO PREDICT HEALTH OUTCOMES

Fernandez C[1,2], Rusk S[1,2], Glattard N[1,2], Shokoueinejad M[3]
[1]Department of Population Health Sciences, University of Wisconsin-Madison, Madison, WI, [2]EnsoData Research, EnsoData, Inc., Madison, WI, [3]Department of Biomedical Engineering, University of Wisconsin-Madison, Madison, WI

Introduction: Machine learning models have grown in popularity for the analysis of Polysomnographic (PSG) data, but many are disadvantaged by their significant lack of interpretability. From a clinical standpoint, it can be challenging to understand what determinant health factors are considered by predictive models to estimate the likelihood of health outcomes. In contrast, we utilize a Computational Phenotyping approach to predict adverse health outcomes based on common clinical variables and interpretable physiological features, providing a clear explanation as to why each estimation is made.
Methods: We used cross-sectional analyses of adults (N = 5,803), ages 39–90 (M ± SD = 63.2 ± 11.2 years), who completed an at-home PSG while participating in the Sleep Heart Health Study. In total, 1,541 interpretable physiological and clinical features were computationally derived from the dataset and used to predict 8 outcome variables including all-cause mortality, stroke, CHD, or CVD. Machine learning techniques including Random Forest, SVM, and Neural Networks were trained, optimized, and evaluated to model the relationship between the interpretable features and health outcomes.
Results: The Random Forest achieved the best predictive performance using a subset of 30 physiological and clinical features. The overall accuracy was 75.3%, with the best single variable performance on all-cause mortality (86% precision, 76% recall). These top 30 features included age, cigarette packs per year, blood pressure, cholesterol, and other variables that are well understood to contribute to the outcomes analyzed. Interestingly however, two thirds of these features represented PSG derived physiological measures. On a quantitative basis, measures of hypoxia, sleep fragmentation, sleep time, and HRV during arousal were observed to have comparable, and in some cases greater, importance than the better understood factors for predicting specific health outcomes.
Conclusion: Computational Phenotyping allows for the generation of accurate and interpretable predictive models for adverse health outcomes that rely on an intuitive subset of physiological and clinical variables. This work represents one of the largest studies analyzing the relationship between health outcomes and PSG based variables using novel machine learning algorithms, and highlights the critical role that sleep physiological measures play in contributing to health outcomes.
Support (If Any):

0068

ESTIMATION OF SLEEP STAGES USING CARDIAC AND ACCELEROMETER DATA FROM A WRIST-WORN DEVICE

Beattie Z, Pantelopoulos A, Ghoreyshi A, Oyang Y, Statan A, Heneghan C
Fitbit, San Francisco, CA

Introduction: We investigated the ability of a wrist-worn tracker to estimate sleep stages in normal adult sleepers. Such a device could be useful in simplifying sleep research and in increasing public knowledge of sleep.
Methods: Movement and cardiac data was collected from 60 adult subjects (36 M: 24 F, ages 34 ± 10 yrs) wearing two wrist worn devices (left and right hand) containing a 3D-accelerometer and an optical photoplethysmogram (PPG), while undergoing a sleep stage study using a Type III home sleep testing device. The accelerometer was used to generate various features of movement; the PPG records cardiac peaks generated by each heartbeat, and can be used to determine heart rate and heart rate variability metrics. The sleep study was scored independently by two registered PSG technicians, using consensus AASM scoring rules. Using these labels, an automated classifier and post-processing rule was developed to label 30-second epochs as one of Wake/Light/Deep/REM (note that Stages N1 and N2 were combined into a single "Light" classification). The estimated performance of this automated classifier system was calculated using a leave-one out validation method. The performance metrics were Cohen's kappa (measures the level of agreement greater than chance) and per-epoch accuracy (percent of epochs correctly labeled).
Results: The estimated Cohen's kappa was 0.52 ± 0.14 for left hand wear, and 0.53 ± 0.14 (right hand). The per-epoch accuracy was 69%. Across the population, there was no statistically significant bias in the estimated durations of the wake, light, deep and REM stages versus the gold standard measurements.
Conclusion: These results suggest that a wrist worn device with movement and cardiac sensors can be used to determine sleep stages with a reasonable degree of accuracy in normal adult sleepers, but without the cost and artificial sleep environment of a sleep laboratory. The reported performance figures are similar or better than previously reported results from non-EEG based sleep staging using combinations of cardiac, respiratory and movement information.
Support (If Any): This work was supported by Fitbit.

0069

DEVELOPMENT AND VALIDATION OF AN ALGORITHM FOR THE STUDY OF SLEEP USING A BIOMETRIC SHIRT IN YOUNG HEALTHY ADULTS

Pion-Massicotte J[1,2], Chicoine M[1,3], Chevrier É[1], Roy J[2], Savard P[4], Godbout R[1,5,3]
[1]Sleep Laboratory and Clinic, Hôpital Rivière-des-Prairies, Montréal, QC, CANADA, [2]Carré Technologies, Montréal, QC, CANADA, [3]Research Center, CIUSSS du Nord-de-l'Île-de-Montréal, Montréal, QC, CANADA, [4]Institut de génie biomédical, Polytechnique Montréal, Montréal, QC, CANADA, [5]Department of Psychiatry, Université de Montréal, Montréal, QC, CANADA

Introduction: Portable polysomnography systems are often too complex and encumbering for home sleep recordings. We assessed the feasibility of measuring sleep with a biometric shirt.
Methods: Twenty healthy young adults (12 women, 8 men; 21.9 ± 2.0 years) were recorded in a sleep laboratory for two consecutive nights using standard polysomnography and a biometric shirt, simultaneously. Polysomnographic recordings were scored using standard methods. The biometric shirt had embedded electrocardiogram sensors, two respiratory inductance plethysmography bands, a 3-axis accelerometer and a detachable microcontroller performing signal acquisition, data processing and communication protocols. The shirt size was selected for each subject so that the signal was optimal. An algorithm was developed to classify the biometric shirt recordings into three vigilance states: wake, nonREM sleep and REM sleep. The algorithm was based on breathing rate and heart rate variability, body movement and included a correction for sleep onset and offset. The results from the two types of recordings were compared with percentages of agreement and kappa coefficients.
Results: Five nights from four subjects were rejected due to recurrent signal artefacts caused by an ill-fitting or misplaced shirt. The overall mean percentage of agreement for 35 recording pairs was 77.55%. When NREM and REM sleep epochs were grouped together, the