250

AI-SUPPORTED SLEEP STAGING FROM ACTIVITY AND HEART RATE

Samadrita Chowdhury,[1] TzuAn Song,[2] Richa Saxena,[3] Shaun Purcell,[4] Joyita Dutta[5]

[1]University of Massachusetts Lowell, [2]University of Massachusetts, [3]Massachusetts General Hospital, [4]Brigham and Women's Hospital, [5]University of Massachusetts Lowell / Massachusetts General Hospital / Harvard Medical School

Introduction: Polysomnography (PSG) is considered the gold standard for sleep staging but is labor-intensive and expensive. Wrist wearables are an alternative to PSG because of their small form factor and continuous monitoring capability. In this work, we present a scheme to perform such automated sleep staging via deep learning in the MESA cohort validated against PSG. This scheme makes use of actigraphic activity counts and two coarse heart rate measures (only mean and standard deviation for 30-s sleep epochs) to perform multi-class sleep staging. Our method outperforms existing techniques in three-stage classification (i.e., wake, NREM, and REM) and is feasible for four-stage classification (i.e., wake, light, deep, and REM).
Methods: Our technique uses a combined convolutional neural network coupled and sequence-to-sequence network architecture to appropriate the temporal correlations in sleep toward classification. Supervised training with PSG stage labels for each sleep epoch as the target was performed. We used data from MESA participants randomly assigned to non-overlapping training (N=608) and validation (N=200) cohorts. The under-representation of deep sleep in the data leads to class imbalance which diminishes deep sleep prediction accuracy. To specifically address the class imbalance, we use a novel loss function that is minimized in the network training phase.
Results: Our network leads to accuracies of 78.66% and 72.46% for three-class and four-class sleep staging respectively. Our three-stage classifier is especially accurate at measuring NREM sleep time (predicted: 4.98 ± 1.26 hrs. vs. actual: 5.08 ± 0.98 hrs. from PSG). Similarly, our four-stage classifier leads to highly accurate estimates of light sleep time (predicted: 4.33 ± 1.20 hrs. vs. actual: 4.46 ± 1.04 hrs. from PSG) and deep sleep time (predicted: 0.62 ± 0.65 hrs. vs. actual: 0.63 ± 0.59 hrs. from PSG). Lastly, we demonstrate the feasibility of our method for sleep staging from Apple Watch-derived measurements.
Conclusion: This work demonstrates the viability of high-accuracy, automated multi-class sleep staging from actigraphy and coarse heart rate measures that are device-agnostic and therefore well suited for extraction from smartwatches and other consumer wrist wearables.
Support (if any): This work was supported in part by the NIH grant 1R21AG068890-01 and the American Association for University Women.

251

AGREEMENT AND RELIABILITY OF A NEW POLYSOMNOGRAPHY SLEEP STAGING ALGORITHM AGAINST MULTIPLE HUMAN SCORERS

Ulysses Magalang,[1] Brendan Keenan,[2] Bethany Staley,[2] Peter Anderer,[3] Marco Ross,[3] Andreas Cerny,[3] Raymond Vasko,[3] Samuel Kuna,[4] Jessie Bakker[5]

[1]The Ohio State University Wexner Medical Center, [2]University of Pennsylvania, [3]Philips Sleep and Respiratory Care, [4]Philadelphia VA Medical Center, [5]Philips Sleep & Respiratory Care

Introduction: Scoring algorithms have the potential to increase polysomnography (PSG) scoring efficiency while also ensuring consistency and reproducibility. We sought to validate an updated sleep staging algorithm (Somnolyzer; Philips, Monroeville PA USA) against manual sleep staging, by analyzing a dataset we have previously used to report sleep staging variability across nine center-members of the Sleep Apnea Global Interdisciplinary Consortium (SAGIC).
Methods: Fifteen PSGs collected at a single sleep clinic were scored independently by technologists at nine SAGIC centers located in six countries, and auto-scored with the algorithm. Each 30-second epoch was staged manually according to American Academy of Sleep Medicine criteria. We calculated the intraclass correlation coefficient (ICC) and performed a Bland-Altman analysis comparing the average manual- and auto-scored total sleep time (TST) and time in each sleep stage (N1, N2, N3, rapid eye movement [REM]). We hypothesized that the values from auto-scoring would show good agreement and reliability when compared to the average across manual scorers.
Results: The participants contributing to the original dataset had a mean (SD) age of 47 (12) years and 80% were male. Auto-scoring showed substantial (ICC=0.60-0.80) or almost perfect (ICC=0.80-1.00) reliability compared to manual-scoring average, with ICCs (95% confidence interval) of 0.976 (0.931, 0.992) for TST, 0.681 (0.291, 0.879) for time in N1, 0.685 (0.299, 0.881) for time in N2, 0.922 (0.791, 0.973) for time in N3, and 0.930 (0.811, 0.976) for time in REM. Similarly, Bland-Altman analyses showed good agreement between methods, with a mean difference (limits of agreement) of only 1.2 (-19.7, 22.0) minutes for TST, 13.0 (-18.2, 44.1) minutes for N1, -13.8 (-65.7, 38.1) minutes for N2, -0.33 (-26.1, 25.5) minutes for N3, and -1.2 (-25.9, 23.5) minutes for REM.
Conclusion: Results support high reliability and good agreement between the auto-scoring algorithm and average human scoring for measurements of sleep durations. Auto-scoring slightly overestimated N1 and underestimated N2, but results for TST, N3 and REM were nearly identical on average. Thus, the auto-scoring algorithm is acceptable for sleep staging when compared against human scorers.
Support (if any): Philips.

252

NON-INVASIVE QUANTIFICATION OF HUMAN BRAIN LACTATE CONCENTRATIONS ACROSS SLEEP-WAKE CYCLES

Selda Yildiz,[1] Miranda Lim,[2] Manoj Sammi,[1] Katherine Powers,[1] Charles Murchison,[3] Jeffrey Iliff,[4] William Rooney[1]

[1]Oregon Health & Science University, [2]VA Portland Health Care System, [3]University of Alabama at Birmingham, [4]VA Puget Sound Health Care System

Introduction: Cellular mechanisms underlying changes in small animal brain lactate concentrations have been investigated for more than 70 years and report sharp reductions in lactate (12-35%) during sleep or anesthesia relative to wakefulness. The goal of this study was to investigate alterations in human cerebral lactate concentrations across sleep-wake cycles. Toward this goal, we developed a novel non-invasive methodology, quantified changes in human cerebral lactate during sleep stages, and investigated potential mechanisms associated with changes in lactate.
Methods: Nine subjects (four females, five males; 21-27 y-o, mean age 24.2 ±2) were sleep deprived overnight, and underwent (5:45~11:00 am) experiments combining simultaneous MR-spectroscopy (MRS) and polysomnography (PSG) in a 3 T MR instrument using a 64-channel head/neck coil. A single voxel MRS (1H-MRS) acquired signals from a volume of interest (12~24 cm3) for every 7.5-s for 88~180-min. Lactate signal intensity was determined from each 7.5-s spectrum, normalized to corresponding water signal, and averaged over 30-s for each PSG epochs. Artifact corrected PSG data were scored for each