Field-Based Assessments of Behavioral Patterns During Shiftwork in Police Academy Trainees Using Wearable Technology

Melissa L. Erickson†,1,2, Will Wang§,1, Julie Counts*, Leanne M. Redman†, Daniel Parker†,10, Janet L. Huebner†, Jessilyn Dunn§,11, and William E. Kraus†,§

†Duke Molecular Physiology Institute, Duke University, Durham, North Carolina, †Translational Research Institute, AdventHealth, Orlando, Florida, §Pennington Biomedical Research Center, Louisiana State University, Baton Rouge, Louisiana, ‡Department of Biomedical Engineering, Duke University, Durham, North Carolina, and ||Department of Biostatistics & Bioinformatics, Duke University, Durham, North Carolina

Abstract  Circadian misalignment, as occurs in shiftwork, is associated with numerous negative health outcomes. Here, we sought to improve data labeling accuracy from wearable technology using a novel data pre-processing algorithm in 27 police trainees during shiftwork. Secondarily, we explored changes in four metabolic salivary biomarkers of circadian rhythm during shiftwork. Using a two-group observational study design, participants completed in-class training during dayshift for 6 weeks followed by either dayshift or nightshift field-training for 6 weeks. Using our novel algorithm, we imputed labels of circadian misaligned sleep episodes that occurred during daytime, which were previously were mislabeled as non-sleep by Garmin, supported by algorithm performance analysis. We next assessed changes to resting heart rate and sleep regularity index during dayshift versus nightshift field-training. We also examined changes in field-based assessments of salivary cortisol, uric acid, testosterone, and melatonin during dayshift versus nightshift. Compared to dayshift, nightshift workers experienced larger changes to resting heart rate, sleep regularity index (indicating reduced sleep regularity), and alterations in sleep/wake activity patterns accompanied by blunted salivary cortisol. Salivary uric acid and testosterone did not change. These findings show wearable technology combined with specialized data pre-processing can be used to monitor changes in behavioral patterns during shiftwork.

Keywords  circadian rhythm, wearables, field-based study, circadian misalignment, shiftwork, cortisol

1. Indicates co-first authorship.
2. To whom all correspondence should be addressed: Melissa L. Erickson, Translational Research Institute, Advent Health, 301 East Princeton Street, Orlando, Fl. 32804, USA; e-mail: Melissa.L.Erickson@AdventHealth.com.
Shiftwork, characterized by behavioral circadian disruption, is associated with numerous negative health consequences including increased chronic disease risk (Antunes et al., 2010; Pan et al., 2011; Lieu et al., 2012; Barbadoro et al., 2013; Depner et al., 2014; Vetter et al., 2016; Manohar et al., 2017; Shan et al., 2018; Gao et al., 2020; Dutheil et al., 2020; Rivera et al., 2020; Schilperoort et al., 2020; Maidstone et al., 2021), obesity, type 2 diabetes (Antunes et al., 2010; Shan et al., 2018; Gao et al., 2020), hypertension (Manohar et al., 2017), dyslipidemia (Dutheil et al., 2020), asthma (Maidstone et al., 2021), and increased breast cancer risk and stroke (Rivera et al., 2020). To advance our understanding of this relationship, a feasible methodology for assessing field-based behavioral circadian disruption on a large-scale is needed. This requisite was recently highlighted in a white paper summarizing discussions at the 2018 Sleep Research Society’s sponsored workshop, “International Biomarkers Workshop and Wearables in Sleep and Circadian Science” (Depner et al., 2020).

The widespread development of commercially available activity trackers affords researchers new opportunities to survey novel behavioral patterns in real-world settings that can be linked to key health indicators (Shcherbina et al., 2017). Wrist-worn smart watches provide information on behavioral regularity of when an individual sleeps and exercises. Wearable technology allows access to a wide range of physiological variables such as activity and heart rate data, which can be used to infer individual circadian phase (Bowman et al., 2021). When combined with mathematical modeling, wearable technology may guide chronotherapy, a strategy that takes individual circadian rhythms into account to understand health outcomes (Stone et al., 2020).

Consumer-based wearable devices are an accessible and cost-effective tool for field-based assessments of behavior on a large scale. While wearable technology holds promise, the accuracy of these devices in challenging situations may be compromised. Sleep detection performance is reduced during disrupted sleep (Chinoy et al., 2021), and whether this extends to shiftwork is unknown. Current technology is optimized for use in “typical” sleep-wake behaviors, while shiftwork requires an individual to be active during the nighttime hours and sleep during daytime hours. This convention biases data labeling performance, in that devices are more likely to label activity data as “wake” during daytime hours and “sleep” during nighttime hours. These misaligned behaviors during shiftwork are often mislabeled, particularly as it relates to sleep episodes during daytime. However, this shortcoming may be overcome by developing a novel data pre-processing step that emphasizes physiological measures and reduces the algorithm’s dependence on external clock time, thereby increasing the accuracy of data labeling when used in the shiftwork setting.

We anticipated that proprietary sleep algorithms originally developed for use by consumers with regular sleep patterns might perform poorly during night shiftwork: daytime sleep episodes would be mislabeled as “awake” (or non-sleep). Therefore, the first aim of this study was to develop a novel algorithm for sleep label imputation that is not biased by external clock time, in a sample of shift working police trainees. The second aim was to assess the feasibility of concurrent field-based salivary sampling to detect changes in known biomarkers of circadian patterns. We hypothesized that our novel algorithm for sleep imputation would provide reasonable labels for daytime sleep episodes that are mislabeled as “awake” (or non-sleep) by commercial technology; and, second, that nightshift work would be reflected by aberrations in biological samples (cortisol, uric acid, testosterone, and melatonin).

MATERIALS AND METHODS

Study Design

This was a two-group observational, repeated measures study design, leveraging the established schedule followed by 36 police recruits. Schedules of police recruits involved 24 weeks of in-class training followed by 14 weeks of field-training. This pilot study lasted approximately 12 weeks and occurred during the last 6 weeks of in-class training (baseline phase) and the first 6 weeks of field-training. During in-class training, classes were held Monday through Friday during daytime (0730-1700 h) hours; this represented normal circadian alignment. This baseline phase was subsequently followed by 6 weeks of field-training. During the field-training phase, 13 participants maintained a normal daytime schedule, representing circadian alignment and 14 participants switched to night shift work, representing circadian misalignment (Figure 1, Panel a). During the second phase (circadian misalignment), trainees were assigned to one of the following four shift work schedules:

- Schedule A: 0600-1700 h (circadian alignment);
- Schedule B: 1000-2100 h (circadian alignment);
- Schedule C: 1600-0300 h (circadian misalignment);
- Schedule D: 2000-0700 h (circadian misalignment).

Two of these four field-training schedules (A and B) align with the 24-h day/night cycle and represented a...
Erickson et al. / ASSESSING CIRCADIAN DISRUPTION IN POLICE TRAINEES 3

[Image: Diagram showing study design with timelines for in-class training, field-based training, and regular duty]

maintenance of behavioral circadian alignment. One participant engaged in office work continued to follow a 0800-1700 schedule. The other two schedules (C and D) were misaligned with the day/night cycle and represented acute circadian misalignment. Work schedules were maintained for four consecutive days, followed by four consecutive days off. Activity monitors were worn continuously and thus captured behavior during both the in-class training phase and the field-training phase. Three salivary samples were collected during the in-class training phase and six samples were collected during the field-training phase, totaling nine salivary samples for each participant. This study design was advantageous because it controls for the job transition from in-class training to field-training due to nightshift and dayshift transition comparisons.

Ethical Approval

This study was approved by the Duke University Health System Institutional Review Board for Clinical Investigations (IRB# Pro00077319). All participants provided written informed consent prior to study participation.

Participants

Study inclusion criteria were as follows: (1) enrolled in a local public safety training program and (2) owned a smartphone. We conducted on-site recruitment events to raise general study awareness by partnering with a local police department. We presented the study to a total of 77 trainees, or four academy classes, and enrolled 36 participants. Participants provided informed consent electronically using a secure web application (REDcap).

Study Protocol

As a field-based study, all assessments were collected outside of the laboratory. After providing informed consent, participants were instructed on
use of the activity tracker (Garmin vívosmart® HR, Olathe, KS) and supplied with six self-administered saliva collection kits, using either drool sampling (SalivaBio Passive Drool, Salimetrics®, Carlsbad, CA) or oral swab method (SalivaBio Oral Swab, Salimetrics®), and then instructed on their use and the collection protocol.

**Activity Tracker Assessments**

Activity monitors were worn on the wrist 24/7 (except for when the watch was being charged). The Garmin vívosmart® HR-recorded observations of activity level, heart rate, and algorithmically generated sleep/wake labels every 15 min, totaling 96 measurements per person during a 24-h period. Wear time was required to be at least 80% over a given 24-h period and individuals meeting this criterion for at least 50% of the days were considered complete and included in the analysis. For days with more than 80% valid observations, missing points were imputed based on the closest value to the epoch. From these final data, changes in resting heart rate and in sleep regularity using the methods described subsequently were evaluated.

**Novel Sleep Labeling Method Development.** Garmin vívosmart® HR relies on user input of anticipated regular bedtime—a key input to their proprietary sleep detection algorithm. However, shift workers followed irregular sleep/wake patterns and this may potentially contribute to inaccuracies in data labeling. We posit that a novel data labeling method that does not require user input information, such as anticipated bedtime, will increase sleep labeling accuracy, particularly as it relates to sleep episodes during daytime, as is typical in the shiftwork setting. Thus, we developed a novel logistic regression-based sleep imputation algorithm that relies on heart rate and activity data—rather than anticipated bedtime—as input information to classify data as either “sleep” or “wake.” Specifically, we developed a model that labels 15-min epochs as sleep or wake based on new input information—heart rate and activity data—collected by Garmin vívosmart® HR and application.

To accomplish this, we first defined our ground truth data set using the following rationale. We acquired reliable sleep periods detected by the Garmin vívosmart® HR and application. Given that Garmin technology is optimized for a typical circadian aligned schedule, we assumed that Garmin vívosmart® HR sleep labels (positive labels) collected during in-class training (which follows a daytime schedule) were reliable. Specifically, reliable awake labels were defined as periods 4-8 h before the sleep period start and 4-8 h after the sleep period end. We used this data set to train and test models to be used for label imputation. First, we performed nested cross-validation (CV) by trainees, in which the outer layer is 5-fold CV by trainee and the inner layer is hyperparameter tuning with parameters selected according to each model. In each outer layer, we used data from 20 trainees for model training, with the inner CV layer for hyperparameter tuning. The best model selected from the inner CV layer was trained on the whole training data set from the 20 trainees, and evaluated on the last five trainee’s data. We have on average 7400 episodes per trainee, which means that for every outer layer of CV, we have approximately 148,000 training samples and 37,000 testing samples. The results are reported in Table 1. From the four algorithms evaluated (Logistic Regression, Random Forest, Support Vector Machine with Radial Basis Function Kernel, and AdaBoost), each of these algorithms achieved great performance in accuracy (>0.94), F1-score (>0.94) and area under curve of receiver operating characteristic plot (ROC-AUC > 0.95). Random forest achieved the best performance based on the nested CV.

**Resting Heart Rate.** Daily resting heart rate was calculated as the mean heart rate at rest, or when the maximum Motion Intensity < 3. Motion Intensity was derived from minute-level accelerometry data and is an aggregate measure of overall activity level for each 15-min epoch. Motion Intensity takes integer values between zero and seven inclusively, with zero corresponding to stillness and higher scores corresponding to more activity.
Sleep Regularity Index. We calculated a sleep regularity index, using sleep/wake labels obtained after the sleep imputation algorithm, to quantify day-to-day sleep regularity over the course of five consecutive days. This is a previously established index that ranges from zero to 100, in which a greater value indicated increased sleep regularity (Lunsford-Avery et al., 2018). The equation for calculation of the sleep regularity index has been described previously. It was used initially on ActiGraph’s sleep/wake label data streams; and therefore, could be easily applied to sleep/wake labels derived from Garmin vívosmart® HR activity and heart rate data for use in this study.

Saliva Assessments

Saliva samples were self-collected using either the cheek cotton swab method or the passive drool method, in which saliva is passed into a collection container via straw. Participants stored saliva samples in their home −20 °C freezer until collected by study staff at the following protocol visit. Samples were then stored at −80 °C until batched analyses. During in-class training, which represents baseline, participants collected three samples: before bed, upon waking, and 30 min after waking (sample must be collected within 60 min after waking to be included in final analysis) on a workday (totaling three samples). During field-training, participants collected three samples at the same behavioral events on a workday and non-workday (totaling six samples). The workday and non-workday samples were averaged, to represent the behavioral time points for field-training.

The differences between in-class training and field-training (average of workday and non-workday) were calculated for each behavioral time point:

1. Before bed<sub>in class-training</sub> − before bed<sub>field-training; average of workday and non-workday</sub>
2. Upon waking<sub>in class-training</sub> − upon waking<sub>field-training; average of workday and non-workday</sub>
3. Wake + 30 min<sub>in class-training</sub> − wake + 30 min<sub>field-training; average of workday and non-workday</sub>

Assessment timepoints are shown in Figure 1, Panel b. These calculations were performed on salivary biomarkers: cortisol, uric acid, testosterone, and melatonin. Next, we compared to deltas between in-class dayshift versus those during in-class nightshift.

Salivary Circadian Biomarker Assays. To assess salivary biomarkers, manufacturer’s instructions were followed using commercially available immunoassay kits; salivary cortisol (Salimetrics #1-3002), salivary uric acid (Salimetrics #1-3802), salivary testosterone (Salimetrics #1-2402), and salivary melatonin (Salimetrics #1-3402). To minimize batch effects, all three behavioral timepoints from a participant were analyzed on the same plate (e.g., saliva sample collected at baseline, upon waking, and wake + 30 min during both in-class training and field-based training). Samples were run in singlicate. Manufacturer-provided controls were run in duplicate on each plate in order to assess intra- and inter-assay variability and to establish an acceptable control range. Laboratory personnel were blinded to study condition.

Statistical Analysis

Data are presented as mean ± standard deviation unless otherwise noted. Python 3.6 (packages statsmodels 0.11.0 and pingouin 0.3.11) was used for statistical analyses.

Aim 1. Novel algorithm performance: we tested agreement between our novel algorithm and reliable Garmin vívosmart® HR sleep labels by evaluating the following: testing and training accuracy, testing F1-score, and testing ROC-AUC. To accomplish this, we used four performance evaluation models: logistic regression, random forest, AdaBoost, and support vector machine (radial basis function). We considered 0.90 testing F1-score and testing ROC-AUC as acceptable performance.

Aim 2. Within each group, differences between in-class training versus field-training were determined using the Wilcoxon signed-rank test for activity monitor measures (resting heart rate and sleep regularity index) and salivary measures (cortisol, uric acid, testosterone, and melatonin). Between-group differences (circadian misalignment vs. circadian alignment) during the transition from in-class to field-training were determined using the non-parametric Kruskal-Wallis test for both activity monitor measures (resting heart rate and sleep regularity index) and salivary measures (cortisol, uric acid, testosterone, and melatonin). The significance threshold was \( p < 0.05 \). Salivary biomarkers were adjusted for multiple comparisons using Bonferroni corrections. Outliers were identified by Grubb’s test.

RESULTS

Participant Characteristics

The study cohort was predominately male (67%; 18 M/9 F). The mean age was 28 years (6.2) ranging...
from 21 to 47 years. The mean body mass index (BMI) was 27 kg/m² (±3.4) ranging from 21 to 33 kg/m². A consort figure is shown in Figure 2. Nine of 36 enrolled participants did not have data due to various reasons (e.g., lost to follow-up, did not follow sample collection instructions).

**Activity Tracker Data Labeling Performance During Shift Work**

Of the 27 participants that completed both phases, we had complete data for 25 participants. We excluded activity tracker data from two participants as they did not meet our wear time criteria.

To determine whether we can rely on sleep labels provided by Garmin vívosmart® HR, we evaluated the performance of Garmin vívosmart® HR sleep labeling during both dayshift and nightshift. We expect trainees to have at least one main sleep event every 24 h, which we defined as the largest block of time spent asleep or in bed exceeding 4 h, regardless of circadian alignment. However, it is possible that trainees stayed awake for more than 24 h at a time, as has been previously reported in shift working nurses (Gamble et al., 2011). We compared the number of Garmin vívosmart® HR-generated sleep periods to this expectation and observed that during the day shift field-training (circadian alignment), 89.7% of the defined main sleep events were accurately labeled as sleep. In comparison, during the night shift field-training (circadian misalignment) only 49.7% of the defined main sleep events were accurately labeled as sleep. We interpreted this to mean activity tracker proprietary algorithms have high sleep detection accuracy used during typical circadian aligned schedule, but poor sleep detection performance during circadian misalignment. These findings reaffirm the need to improve data labeling performance, particularly as it relates to labeling sleep episodes during daytime, as is typically for shiftwork.

**Novel Sleep Imputation Method Development**

Our algorithm demonstrated high epoch-by-epoch prediction accuracy on the test data set, with logistic regression achieving a testing accuracy, or level of agreement with the Garmin vívosmart® HR algorithm, of 94% (Table 1). Our test data set consisted of continuous wearable data captured during dayshift and nightshift. While all four models demonstrated high level of agreement, we ultimately chose logistic regression because of model simplicity and less risk of overfitting. We then used the logistic regression model to determine sleep versus wake labels for each epoch during nightshift work (circadian misalignment) and
imputed sleep labels during daytime that were previously missed by the Garmin’s vívosmart® HR proprietary algorithm. In addition, we compared the performance of this algorithm versus one that only takes activity into account, as well as one that only takes heart rate into account. We observed best sleep episode labeling accuracy using the algorithm that takes both heart rate and activity into account.

Figure 3 compares sleep labeling by Garmin vívosmart® HR propriety software (gray shading) versus sleep labeling by our novel method (overlaid orange shading) over the course of 4 days, plotted as heart rate (Panel a), Mean Motion Intensity (Panel b), and Max Motion Intensity (Panel c). These data show that Garmin vívosmart® HR labeled a main sleep event during the first 24-h period, but mislabeled sleep events as “awake” (or non-sleep) events during the subsequent three nights. In contrast, our sleep labeling method was able to impute a main sleep event for each 24-h period using heart rate and activity raw data. These four (sequential) days were chosen arbitrarily.

For further demonstration, Figure 4 compares Garmin vívosmart® HR sleep labeling (blue shading) versus our novel sleep imputation method (overlaid orange shading) over the course of day-shift in-class training and the transition to night-shift field-based training. The nightshift-start date is indicated by the dashed horizontal line. These data show that the novel sleep imputation method increases data labeling performance of sleep episodes occurring during daytime, which overcomes data labeling accuracy limitations specifically in the context of nightshift work.

Activity Tracker Assessments

To determine whether physical activity and algorithm-derived sleep patterns across phases of the day were indicative of the occurrence of a circadian misalignment, we developed a polar plot to visualize activity and sleep behavior fluctuations during both daytime training versus nightshift field-training, totaling 42 days. As shown (Figure 5), the polar plots
depict behavioral pattern shifts relative to the external clock time over long durations (e.g., several weeks) and demonstrate the dramatic shift in the sleep/wake routine relative to the external clock time that is absent during dayshift work (Figure 5a) but present during nightshift (Figure 4b). Hence, we concluded that following a nightshift schedule resulted in behavioral circadian misalignment.

**Resting Heart Rate.** For the trainees remaining in circadian alignment, resting heart rate was 63.5 ± 6.4 bpm (beats per minute) during in-class training and increased to 65.4 ± 6.7 bpm during field-training; whereas for the trainees who underwent circadian misalignment, resting heart rate increased from 66.1 ± 4.5 bpm during in-class training to 72.5 ± 6.0 bpm during field-training. Circadian misalignment resulted in a significantly larger change in resting heart rate ($p = 0.009$, Table 2).

**Sleep Regularity Index.** For the trainees remaining in circadian alignment, sleep regularity index was 65.5 ± 13.4 during in-class training and changed to 67.0 ± 10.2 during field-training; whereas for the trainees who underwent circadian misalignment, sleep regularity index was 64.5 ± 8.2 during in-class training and decreased to 55.0 ± 9.8 during field-training. Circadian misalignment resulted in a significantly larger decrease in sleep regularity index ($p = 0.050$, Table 2).

### Table 2. Behavioral and biological changes during shiftwork.

<table>
<thead>
<tr>
<th>Outcome Measures</th>
<th>Delta From In-Class Training to Field-Training</th>
<th>Dayshift: Circadian Alignment</th>
<th>Nightshift: Circadian Misalignment</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity tracker assessments</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resting heart rate, bpm</td>
<td>1.84 ± 2.33</td>
<td>6.48 ± 4.28</td>
<td>0.00898*</td>
<td></td>
</tr>
<tr>
<td>Sleep regularity index</td>
<td>1.44 ± 10.8</td>
<td>9.49 ± 10.9</td>
<td>0.0500</td>
<td></td>
</tr>
<tr>
<td>Salivary biomarker</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cortisol</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before bed</td>
<td>-0.0018 ± 0.13</td>
<td>-0.046 ± 0.11</td>
<td>0.142</td>
<td></td>
</tr>
<tr>
<td>Upon waking</td>
<td>-0.00056 ± 0.015</td>
<td>-0.11 ± 0.21</td>
<td>0.310</td>
<td></td>
</tr>
<tr>
<td>Wake + 30 min</td>
<td>0.04 ± 0.18</td>
<td>-0.32 ± 0.23</td>
<td>0.00156*</td>
<td></td>
</tr>
<tr>
<td>Uric acid</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before bed</td>
<td>0.012 ± 1.32</td>
<td>-0.29 ± 3.1</td>
<td>0.280</td>
<td></td>
</tr>
<tr>
<td>Upon waking</td>
<td>-1.01 ± 3.79</td>
<td>2.81 ± 3.47</td>
<td>0.128</td>
<td></td>
</tr>
<tr>
<td>Wake + 30 min</td>
<td>-0.12 ± 1.85</td>
<td>-0.52 ± 3.4</td>
<td>0.939</td>
<td></td>
</tr>
<tr>
<td>Testosterone</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before bed</td>
<td>-15.4 ± 70.0</td>
<td>-27.2 ± 145</td>
<td>0.440</td>
<td></td>
</tr>
<tr>
<td>Upon waking</td>
<td>232 ± 721</td>
<td>40.3 ± 83.5</td>
<td>0.866</td>
<td></td>
</tr>
<tr>
<td>Wake + 30 min</td>
<td>-40.9 ± 126</td>
<td>-3.32 ± 61.3</td>
<td>0.537</td>
<td></td>
</tr>
<tr>
<td>Melatonin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before bed</td>
<td>47.5 ± 162</td>
<td>3.91 ± 9.29</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Upon waking</td>
<td>-5.35 ± 15.3</td>
<td>0.93 ± 11.7</td>
<td>0.482</td>
<td></td>
</tr>
<tr>
<td>Wake + 30 min</td>
<td>2.11 ± 16.9</td>
<td>-4.86 ± 13.83</td>
<td>0.482</td>
<td></td>
</tr>
</tbody>
</table>

Values presented as mean ± SD. $p \leq 0.05$.

Salivary Assessments

Salivary data were analyzed from 19 participants. For the trainees remaining in circadian alignment, cortisol measured 30 min after waking was 0.31 ± 0.15 µg/dL during in-class training and changed to 0.35 ± 0.16 µg/dL during field-training; whereas for the trainees who underwent circadian misalignment, cortisol measured 30 min after waking was 0.57 ± 0.25 µg/dL during in-class training and decreased to 0.25 ± 0.14 µg/dL during field-training. Circadian misalignment resulted in a significantly larger decrease in cortisol measured 30 min after waking ($p = 0.0002$, Table 2). Cortisol measures before sleep and upon waking did not significantly change during circadian alignment versus circadian misalignment (Table 2).

Uric acid measured before bed, upon waking, and 30 min after waking, testosterone measured before bed, upon waking, and 30 min after waking, and melatonin measured before bed, upon waking, and 30 min after waking did not significantly change during circadian alignment versus circadian misalignment (Table 2).

### DISCUSSION

Shiftwork is associated with increased disease risk (Antunes et al., 2010; Pan et al., 2011; Lieu et al., 2012; Barbadoro et al., 2013; Depner et al., 2014; Vetter et al.,...
To further our understanding of the health risks associated with this highly prevalent occupational demand, we are in need of field-based methodologies that quantify behavioral circadian disruption at large scale (Depner et al., 2020). Here, we used a commercially available wrist-worn activity trackers (Garmin vívosmart® HR) to assess alterations in activity and sleep patterns occurring as a result of changing from a dayshift to a nightshift work schedule.

Consumer activity trackers have long been used in research settings to assess behavior in various patient populations (Adams et al., 2021; Bayoumy et al., 2021; Ometov et al., 2021). Previous approaches to studying circadian rhythm use heart rate and activity data, assessed by wearables, to estimate individual circadian phase response curves in field-based settings (Bowman et al., 2021). Other approaches apply systems modeling and machine learning to estimate circadian phase (Stone et al., 2020) to guide chronotherapy (Tyler et al., 2020; Kim et al., 2020). The current work differs from these in that we seek to specifically address the shortcoming of data labeling inaccuracies during shiftwork. Using a novel data pre-processing step, we improve upon the limitations of Garmin vívosmart® HR devices, to better quantify behavioral patterns in shift workers.

Our new approach specifically overcomes the issue that Garmin vívosmart® HR devices rely on external clock time and self-reported sleep time of the user to detect sleep and wake episodes. Reliance on external clock time increases the risk of bias toward mislabeling periods of low activity during nighttime hours as “sleep,” which may not necessarily be accurate during

---

Figure 5. Polar plot display of activity patterns assessed using Garmin vívosmart® HR (shown in blue) and sleep patterns derived from the novel sleep labeling method (shown in red). Days are plotted on the radial axis for two consecutive observational periods (in-class training followed by field-training). Each activity and sleep data pair represents one participant. Panel a shows paired activity and sleep data from four participants assessed during circadian alignment. Panel b shows paired activity and sleep data from four participants during circadian misalignment. Black dashed circle indicates approximate timing of the transition from in-class training to field-training. Intensity of the activity data are represented by increased pixel color intensity as indicated in the figure legend on the top level of both panels.
nighttime shiftwork. Reciprocally, there is a risk of bias toward mislabeling actual sleep episodes as “wake” when sleep occurs during daytime hours. Hence, we adapted this novel sleep labeling method to overcome these limitations. Our method relies on heart rate and activity data as input information, which are physiological indicators of activity, rather than anticipated bedtime, to label sleep/wake episodes. We had approximately 7340 event epochs per person and 76.5 days of continuous data collected for each person. This high frequency of data points yields a more accurate prediction compared to input variables collected with low frequencies (Dunn et al., 2021). This approach is particularly advantageous for smaller samples sizes and was effective in our 27-person current sample size.

Behavioral regularity contributes to internal circadian timing, while behavioral irregularities contribute to mistiming, or internal circadian dyssynchrony (Bass and Lazar, 2016). Sleep and wake patterns, in addition to eating and exercise, are relevant behaviors impacting circadian timing (Bass and Lazar, 2016; Zhang et al., 2021). The sleep regularity index was established as a tool to quantify the degree of sleep regularity in a group of older adults (mean age = 68.7 ± 9.2 years) (Lunsford-Avery et al., 2018). The initial validation study reports that greater sleep irregularity was associated with 10-year cardiovascular disease risk, as well as greater obesity, hypertension, fasting glucose, hemoglobin A1c, and diabetes status (Lunsford-Avery et al., 2018). In our study, we compared the sleep regularity index assessed during circadian misaligned and circadian aligned behavioral conditions. As expected, we observed a decline in sleep regularity during night shiftwork. This decline in the sleep regularity index occurred concurrently with changes in activity and sleep patterns assessed using the novel sleep labeling method. These complementary findings support the use of our sleep labeling method as a novel tool to assess changes in activity and sleep behavioral patterns imposed by a night shift schedule.

Circadian rhythms are intrinsic, self-sustaining patterns generated by internal molecular clocks residing in virtually all cells of the body (Takahashi et al., 2008). The gold-standard for assessing circadian rhythm in human is the constant routine protocol (Duffy and Dijk, 2002) and dim-light melatonin onset (Lewy et al., 1999). Several hormones also display 24-h oscillating rhythms. Alternatively, the secretion patterns of these hormones—including cortisol (Hofstra and de Weerd, 2008)—can be used to infer circadian phase. In this study, we assessed salivary cortisol using self-administered saliva detection kits. We observed that salivary cortisol decreased during circadian misalignment. This occurred in parallel with the decline in sleep regularity index as well as changes alternations in Garmin vívosmart® HR-reported activity and sleep. In addition to cortisol, we examined changes in salivary testosterone, uric acid, and melatonin. Testosterone and uric acid were unchanged during circadian misalignment. Testosterone was highly variable, in part, because 33% of the cohort was female; thus, we did not detect significant changes. While we observed changes in salivary cortisol, further work with increased sampling frequency and larger sample size will be needed to draw more definitive conclusions.

Assessments of individual behavioral patterns in real-life settings will require large-scale field-tested methods. To do so, we partnered with local police trainees and leveraged their established training schedule. We controlled for the stress of transitioning from in-class training to field-training through comparisons of both nightshift and dayshift schedules. Our field-based detections of aberrations in physiological parameters such as heart rate and cortisol may point toward potential ways in which circadian dyssynchrony impacts health parameters, yet further work in this area is needed. Overall, there was high compliance to the study protocol. This may be, in part, due to the fact that research staff largely conducted recruitment, consenting, and data collection at the work site (local police academy) and electronically rather than requiring inpatient laboratory visits. We seek to overcome a critical methodological barrier by quantifying circadian rhythm disruption in field-based settings. And as with any field-based study, there were some challenges. The melatonin assay requires a relatively high sample volume (100 µL), whereas other biomarkers require lower volumes (salivary cortisol: 25 µL; salivary testosterone: 25 µL; salivary uric acid: 10 µL). We initially used the oral swab method of sample collection; however, this did not capture adequate volume resulting in missing values for melatonin. After 12 participants, we switched to the passive drool method in efforts to address this limitation and collect larger volumes; yet, this method still resulted in inadequate volume. Therefore, we are unable to draw definite conclusions on changes in melatonin in this study.

The use of Garmin vívosmart® HR devices for assessing sleep/wake patterns in a field-based setting is not without limitations. Due to the proprietary nature of the algorithm, we are not fully aware of data processing methods; this includes Motion Intensity calculations, which are relied on for the current sleep label imputation method. Others have previously reported that Garmin devices have low specificity in detecting wake episodes after sleep onset (Chinoy et al., 2021), highlighting a need to address shortcomings. Despite these limitations,
however, Garmin vívosmart® HR devices were employed in this current pilot study due user-friendliness, cost-effectiveness (relative to actigraphy), device durability that is necessary for field-based studies with extended wear times, as well as accessibility to raw data (relative to other commercially available wearable devices). Thus, the sleep imputation method described herein provides novel value to the field by improving upon Garmin vívosmart® HR device limitations.

Wearable technology has several advantages that make it a suitable tool for furthering the study of circadian rhythms in field-based settings. There are existing approaches using actigraphy data for determination of sleep-wake patterns (Cole et al., 1992). In contrast, we focus on consumer wearable devices as they are better suited for large-scale use due to accessibility, cost-effectiveness, and ease of use relative to traditional actigraphy. Consumer wearables have limited accuracy in challenging scenarios, such as shiftwork. Here, we overcome this limitation by improving the accuracy of data labeling, particularly as it relates to labeling sleep episodes during daytime during shiftwork. This was accomplished by using heart rate and activity data as input, rather than external clock time, during a data pre-processing step. In line with wearable-detected alterations in activity and sleep patterns, we observed declines in the sleep regularity index and lower salivary cortisol an endogenous marker of circadian phase. This methodological advancement is a first step toward our goal of assessing sleep-wake behavioral patterns on large scale by long-term assessments in large and diverse populations. Long-term, we expect this work will lead to improved methodologies for studying the negative health consequences of living in circadian misalignment.

ACKNOWLEDGMENTS

This study was supported by NIH IP2CHD086851 to the Rehabilitation Research Resource to Enhance Clinical Trials Center at the University of Alabama, Birmingham. MLE was supported, in part, by NIH T32DK064584. This project has been made possible, in part, by grant number 2020-218599 from the Chan Zuckerberg Initiative DAF, an advised fund of Silicon Valley Community Foundation.

CONFLICT OF INTEREST STATEMENT

The author(s) have no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ORCID iDs

Melissa L. Erickson https://orcid.org/0000-0003-2953-3073
Daniel Parker https://orcid.org/0000-0001-8512-1044

REFERENCES


