#Correspondence should be addressed to: Melissa A. St. Hilaire, PhD, 510 Turnpike Street, Room 202-H, Department of Computer and Data Sciences, Merrimack College, North Andover, MA 01845, United States, sthilairem@merrimack.edu.

[1] Present address: Turner Institute for Brain and Mental Health, School of Psychological Sciences, Monash University, Clayton, VIC, Australia

[2] Present address: Department of Computer and Data Sciences, School of Engineering and Computational Sciences, Merrimack College, North Andover, MA, USA

[3] Present address: Center for Innovation in Digital Health Care, Massachusetts General Hospital, Boston, MA, USA (present address)

[4] Present address: Department of Neurology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

[*] These authors contributed equally to this work.

[†] These authors contributed equally to this work.

Conflicts of Interest

Dr. Phillips has received research funding from Versalux and Delos, and he is co-founder and co-director of Circadian Health Innovations PTY LTD. He also holds a patent (US20150080756A1) for estimating arousal states from ambulatory recordings by using sleep and wake models.

Dr. St. Hilaire has received grant/research support from NIH, NASA, and Merrimack College.

Dr. Barger reports honorariums from AAA Foundation, University of Arizona and University of British Columbia.

Mr. O'Brien has no COI to report from the past 3 years.

Dr. Rahman holds patents for (1) Method and device for preventing alterations in circadian rhythm (U.S. patent application Ser. No. 10/525,958), and (2) Methods and devices for improving sleep performance in subject exposed to light at night (U.S. Application No. 61/810,985); owns equity in Melcort Inc.; has provided paid consulting services to Sultan & Knight Limited, Bambu Vault LLC, Lucidity Lighting Inc.; and has received honoraria as an invited speaker and travel funds from Starry Skies Lake Superior, University of Minnesota Medical School, PennWell Corp., and Seoul Semiconductor Co. Ltd., FALK FOUNDATION E.V.; and has received grant/research support from Seoul Semiconductor Co. Ltd., Biological Innovation and Optimization Systems, LLC, Merck & Co., Inc., Pfizer Inc., Vanda Pharmaceuticals Inc., Lighting Science Group, NIH, and NASA. These interests were reviewed and managed by Brigham and Women's Hospital and Mass General Brigham in accordance with their conflict-of-interest policies.

Dr. Landrigan holds equity in and has consulted with the I-PASS Patient Safety Institute, which seeks to train institutions in best handoff practices and aid in their implementation. In addition, Dr. Landrigan has received monetary awards, honoraria, and travel reimbursement from multiple academic and professional organizations for teaching and consulting on sleep deprivation, physician performance, handoffs, and safety, and has served as an expert witness in cases regarding patient safety and sleep deprivation.

Dr. Lockley reports consulting fees from the Hintsa Performance AG, Stantec and View Inc, and has consulting contracts with Absolute Rest, Akili Interactive, Apex 2100 Ltd, Ashurst Risk Advisory, Consumer Sleep Solutions, KBR Wyle Services, Light Cognitive; Lighting Science Group Corporation/HealthE; and Mental Workout/Timeshifter. He has received honoraria and travel or accommodation expenses from Bloxhub, Clifton College, Danish Centre for Lighting, and University of Toronto; and travel or accommodation expenses (no honoraria) from Wiley; and royalties from Monash University and Oxford University Press. He holds equity in iSleep pty. He has received an unrestricted equipment gift and investigator-initiated grant from F. Lux Software LLC, and a Clinical Research Support Agreement and Clinical Trial Agreement with Vanda Pharmaceuticals Inc. He is an unpaid Board Member of the Midwest Lighting Institute (nonprofit). He was a Program Leader for the CRC for Alertness, Safety and Productivity, Australia, through an adjunct position at Monash University (2015–2019). He is part-time adjunct professor at the University of Surrey, UK. He holds several pending patents (US2019366032A1; USD943612S1; US2021162164A1, US20220151552A1). He has served as a paid expert in legal proceedings related to light, sleep, and health. His interests are reviewed and managed by Mass General Brigham in accordance with their conflict-of-interest policies.

Dr. Czeisler reports that he serves as the incumbent of an endowed professorship provided to Harvard Medical School by Cephalon, Inc. and reports institutional support for a Quality Improvement Initiative from Delta Airlines and Puget Sound Pilots; education support to Harvard Medical School Division of Sleep Medicine and support to Brigham and Women's Hospital from: Jazz Pharmaceuticals PLC, Inc, Philips Respironics, Inc., Optum, and ResMed, Inc.; research support to Brigham and Women's Hospital from Axsome Therapeutics, Inc., Dayzz Ltd., Peter Brown and Margaret Hamburg, Regeneron Pharmaceuticals, Sanofi SA, Casey Feldman Foundation, Summus, Inc., Takeda Pharmaceutical Co., LTD, Abbaszadeh Foundation, CDC Foundation; educational funding to the Sleep and Health Education Program of the Harvard Medical School Division of Sleep Medicine from ResMed, Inc., Teva Pharmaceuticals Industries, Ltd., and Vanda Pharmaceuticals; personal royalty payments on sales of the Actiwatch-2 and Actiwatch-Spectrum devices from Philips Respironics, Inc; personal consulting fees from Axsome Therapeutics, Bryte Foundation, With Deep, Inc. and Vanda Pharmaceuticals; honoraria from the Associated Professional Sleep Societies, LLC for the Thomas Roth Lecture of Excellence at SLEEP 2022, from the Massachusetts Medical Society for a New England Journal of Medicine Perspective article, from the National Council for Mental Wellbeing, from the National Sleep Foundation for serving as chair of the Sleep Timing and Variability Consensus Panel, for lecture fees from Teva Pharma Australia PTY Ltd. and Emory University, and for serving as an advisory board member for the Institute of Digital Media and Child Development, the Klarman Family Foundation, and the UK Biotechnology and Biological Sciences Research Council. CAC has received personal fees for serving as an expert witness on a number of civil matters, criminal matters, and arbitration cases, including those involving the following commercial and government entities: Amtrak; Bombardier, Inc.; C&J Energy Services; Dallas Police Association; Delta Airlines/Comair; Enterprise Rent-A-Car; FedEx; Greyhound Lines, Inc./Motor Coach Industries/FirstGroup America; PAR Electrical Contractors, Inc.; Puget Sound Pilots; the San Francisco Sheriff's Department; Schlumberger Technology Corp.; Union Pacific Railroad; United Parcel Service; and Vanda Pharmaceuticals. CAC has received travel support from Merck, Sharpe and Dohme for travel to Japan and Australia; equity interest in Vanda Pharmaceuticals, With Deep, Inc, WHOOP, Inc., and Signos, Inc.; and institutional educational gifts to Brigham and

Predicting neurobehavioral performance of resident physicians in a Randomized Order Safety Trial Evaluating Resident-physician Schedules (ROSTERS)

Andrew J. K. Phillips, Ph.D.[a,b,1,*], **Melissa A. St. Hilaire, Ph.D.**[a,b,2,*,#], **Laura K. Barger, Ph.D.**[a,b], **Conor S. O'Brien**[a,b,3], **Shadab A. Rahman, Ph.D. M.P.H.**[a,b], **Christopher P. Landrigan, M.D. M.P.H.**[a,b,c], **Steven W. Lockley, Ph.D.**[a,b], **Charles A. Czeisler, Ph.D., M.D.**[a,b,†], **Elizabeth B. Klerman, M.D., Ph.D**[a,b,4,†]

[a]Division of Sleep and Circadian Disorders, Departments of Medicine and Neurology, Brigham and Women's Hospital, Boston, MA, USA

[b]Division of Sleep Medicine, Department of Medicine, Harvard Medical School, Boston, MA, USA at the time this work was completed.

[c]Division of General Pediatrics, Department of Pediatrics, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA

Abstract

Objectives: Mathematical models of human neurobehavioral performance that include the effects of acute and chronic sleep restriction can be key tools in assessment and comparison of work schedules, allowing quantitative predictions of performance when empirical assessment is impractical.

Methods: Using such a model, we tested the hypothesis that resident physicians working an extended duration work roster (EDWR), including 24–28 hours of continuous duty and up to 88 hours per week averaged over 4 weeks, would have worse predicted performance than resident physicians working a rapidly cycling work roster (RCWR) intervention designed to reduce the duration of extended shifts. The performance metric used was attentional failures (i.e., PVT lapses). Model input was 169 actual work and sleep schedules. Outcomes were predicted hours per

week during work hours spent at moderate (equivalent to 16–20 hours of continuous wakefulness) or high (equivalent to 20 hours of continuous wakefulness) performance impairment.

Results: The model predicted that resident physicians working EDWR would spend significantly more time at moderate impairment (p=0.02, effect size=0.2) than those working RCWR; this difference was most pronounced during the circadian night (p<0.001). On both schedules, performance was predicted to decline from weeks 1+2 to weeks 3+4 (p<0.001), but the rate of decline was significantly greater on EDWR (p<0.01). Predicted performance impairment was inversely related to prior sleep duration (p<0.001).

Conclusions: These findings demonstrate the utility of a mathematical model to evaluate the predicted performance profile of schedules for resident physicians and others who experience chronic sleep restriction and circadian misalignment.

Introduction

When individuals perform with during extended time awake ("sleep deprivation"), after multiple nights of insufficient sleep ("chronic sleep restriction"), and/or at an adverse circadian phase, their neurobehavioral performance is impaired[1,2,3,4]. In the context of occupational health, these impairments are associated with increased risk of errors, workplace injuries, and motor vehicle crashes during the commute home[5–10]. Work schedules that require long continuous duty hours and/or night shift work should therefore be designed with sleep and circadian factors in mind. Mathematical models are one tool to facilitate designing these schedules.

The US medical profession requires resident physicians to work extended duration work rosters (EDWRs) that pose challenges for human sleep and circadian biology. Currently standard schedules permit resident physicians to work 80 hours per week (averaged over 4 weeks) including continuous work shifts of 24–28 hours ('on-call' shift). This work pattern results in sleep deprivation, chronic sleep restriction and work during adverse circadian phases (i.e., at times when the circadian system is promoting sleep)[11]. Not surprisingly, EDWRs have been shown to be associated with impairment in medical and neurobehavioral performance[12–14]. When similarly challenging schedules are simulated in laboratory environments, they induce severe neurobehavioral (both objective performance and mood) impairment and metabolic dysfunction[15–19].

To mitigate the increased risk to both patient and physician safety[6,20,21], schedules have been developed that reduce the number of continuous duty hours, weekly work hours, and/or the frequency of extended-duration work shifts[22,23]. One of these schedules was a Rapidly Cycling Work Roster (RCWR) with scheduled continuous duty hours limited to 16 h (RCR-16) consisting of a short day shift (7:00–15:00), a long day shift (7:00–22:00), and a long night shift (21:00–13:00).

The RCWR was tested against a traditional EDWR in a clinical trial design detailed below. Relative to a traditional EDWR, the RCWR intervention was found to significantly increase total sleep time and decrease incidence of serious medical errors when applied to post-graduate year 1 (PGY-1) resident physicians working in the ICU[11,14].

As a complement to quantitative experimental approaches, mathematical models have been widely used for schedule assessment, design, and optimization[24–27,28–35]. Modeling of performance for individuals and teams[26] is valuable for several reasons: (i) modeling allows completely unbiased comparisons of performance between schedules, eliminating any concerns around blinding or expectancy effects; (ii) modeling allows performance to be interrogated at all times, even when it would be logistically difficult to empirically assess (e.g., while participants are performing work tasks or driving); (iii) modeling allows quantitative predictions of performance across multiple possible schedules since empirical assessment of every possible schedule would be impractical; (iv) the model can provide insights into the basis for performance differences between schedules, illuminating a path to effective interventions; and (v) once a model has appraised existing schedules, it can be used in future to predictively assess potential alternative schedules in a safe, relatively inexpensive, and rapid fashion, allowing candidate schedules to be assessed prior to a randomized control trial.

To make accurate predictions on resident physician schedules, a model should ideally be able to predict changes in neurobehavioral performance due to a combination of: (i) the dynamics of the circadian pacemaker (circadian phase is potentially shifted by rotating schedules); (ii) the effects of acute sleep deprivation (since 24–28 h shifts are permitted by the Accreditation Council for Graduate Medical Education [ACGME]); and (iii) the effects of variable timing and chronic sleep restriction that causes accumulated sleep debt over longer timescales. To date, only one model has been applied to resident physician schedules[26]; the model included a dynamic circadian pacemaker and the effects of acute sleep deprivation on performance[36], but did not account for effects of chronic sleep restriction. The model also was only applied to theoretical schedules. Here, we employ a validated model of the human circadian pacemaker[33] combined with a sleep-homeostatic model that has previously been trained and tested using acute sleep deprivation and both stable and variable sleep timing within chronic sleep restriction protocols[19,37]. This model was used to simulate 169 actual EDWR and RCWR schedules to test the hypothesis that RCWR improves neurobehavioral performance compared to EDWR.

Methods

The model is briefly reviewed below, followed by a description of the schedules and their implementation in the model.

Mathematical model

We recently developed a model of human neurobehavioral performance with equations based on the structure of the sleep-promoting adenosine system in the brain[37]. The model includes both circadian and sleep homeostatic variations in performance, including both acute and chronic effects of sleep loss and recovery. Specifically, the model includes a drive,

$$D = H + aC,$$

(1)

where H is the homeostatic sleep drive (based on the concentration of bound adenosine A_1 receptors), a is a constant, and C is the circadian rhythm. The model was previously calibrated against Psychomotor Vigilance Task (PVT) data in healthy young adults under conditions of acute sleep deprivation, chronic sleep restriction, and recovery[38]. The PVT is a visual vigilance task that is sensitive to sleep deprivation, sleep restriction, and circadian phase[16,38,39] and has been correlated with performance levels associated with alcohol intoxication[40,41]. The model generates an estimate of the number of PVT lapses (i.e., responses slower than 500 ms) that would occur during a standard 10-minute test, using a sigmoid function,

$$P = S(D).$$

(2)

The model has to date been used to simulate schedules with night sleep opportunities only[19,37]. Thus, it was previously valid to assume entrainment (i.e., a relatively stable circadian phase), as has been done in other sleepiness and performance models[27,42]: C was modeled as a sinusoid (i.e., a constant amplitude, 24-hour rhythm). Since the model is used here to simulate schedules that involve both day and night sleep opportunities, and included light/dark patterns known to affect timing ("phase") of the circadian pacemaker, it is important to incorporate into the performance model a dynamic model of the circadian pacemaker (i.e., a model that can accurately predict changes in phase or amplitude of the circadian rhythm in response to a light stimulus) rather than a static model (e.g., sinusoidal representation of the circadian pacemaker). For this, we followed an approach used previously in which a dynamic circadian pacemaker model is substituted in place of a sinusoidal rhythm[43,44]. Specifically, the sinusoid is replaced by the best-fitting output of the latest validated version of the Jewett-Kronauer-Forger model of the human circadian pacemaker[33]. This was achieved by setting

$$C = -0.608x + 0.685x_c + 0.075,$$

(3)

where x and x_c are the two pacemaker outputs of the differential-equation-based dynamic circadian pacemaker model[33]. The coefficients were selected based on a least-squares fit to the previously used sinusoid.

The model used two inputs: each resident's recorded sleep/wake patterns, and a modeled light input to the dynamic circadian pacemaker model (measured in lux, method detailed below). We used the model to generate predictions of number of attentional failures (i.e., PVT lapses) at each point in time as output.

Schedules

The schedules used for modeling were created from daily logs of sleep and work (detailed below) of resident physicians from July 2013 to March 2017 as part of the Randomized Order Safety Trial Evaluating Resident Schedules (ROSTERS) study[45]. This was a multi-center cluster-randomized, clinical trial designed to evaluate the effectiveness of replacing EDWRs with RCWRs in six pediatric intensive care units (PICUs). Each condition had a 4-month wash-in period followed by an 8-month data collection period, with the order of the schedule condition being randomly assigned[45]. All post-graduate year 2 (PGY-2) and 3 (PGY-3) resident physicians working in the PICU at each site over the study period were invited to participate. Full details of the study, including PVT results, are reported elsewhere[45–47]. The study was approved by the Institutional Review Board (IRB) at each academic medical center as well as at the University of California San Francisco (Data Coordinating Center, DCC) and by the Partners Human Research Committee (Clinical Coordinating Center, CCC), and is reported on ClinicalTrials.gov (NCT02134847).

A daily electronic log was used to collect data from participants on sleep timing (including sleep latency estimates) and duration, awakenings, and start and end times of work schedules ("rosters"). A paper version of the same sleep log was previously shown to have high epoch-to-epoch agreement with continuous polysomnography (95.6% epoch-to-epoch agreement; r=0.94, p<0.001)[11]. Analyses of experimental sleep and work schedule data are presented in Barger et al.[6] For inclusion in the model, we required at least 3 consecutive weeks of sleep/wake and work data with no missing intervals. A total of 298 schedules had 3 consecutive weeks of sleep/wake and work data collection, of which 169 (57%) had no missing intervals and were suitable for modeling: 95 EDWRs and 74 RCWRs. On logs where individuals failed to report sleep latency (5.0% of nights), that same individual's median reported sleep latency was imputed using median substitution. Schedule data included in the analysis ranged in length from 21.2 days to 33.3 days for EDWR and from 21.2 days to 33.8 days for RCWR.

Simulations

Sleep/wake schedules were implemented by forcing the model to be in sleep or wakefulness during the times recorded in the logs. Light levels were simulated using a previously published method for simulating how individuals self-select light patterns in the real world[48]. Specifically, an indoor light level of 50 lux was used for both work hours and hours of wakefulness before sunrise or after sunset[49–51]. A smooth light profile with a maximum of 300 lux was used for non-work hours (including potential commuting hours) between sunrise and sunset:

$$I = 150\tanh(c(t - s_1)) - 150\tanh(c(t - s_2)),$$

(4)

where t is clock time (time in hours, modulo 24, ranging from 0 to 24). The parameters $s_1 = 7.5$ h, $s_2 = 16.5$ h, and $c = 0.6 \, \text{h}^{-1}$ were chosen to generate an approximately 12-hour photoperiod centered at noon. Light levels were set to zero lux during any times the

individual was asleep; therefore, light levels rise and fall abruptly at sleep offset and onset, respectively.

Prior to simulating the work schedules, the model was entrained to a permanent day-shift schedule for 6 months with two non-work days per week and sleep opportunities from 22:30 to 5:30 on the five work days per week; this was to ensure entrainment of the circadian pacemaker and equilibration of the long timescale sleep homeostatic component to the schedule. This schedule was motivated by the pre-study conditions in Lockley et al.[11] and corresponds to the sleep times used in a previous simulation paper[26]. To obtain PVT thresholds corresponding to moderate impairment and high impairment, a simulation was performed in which the model was held awake from a wake time of 5:30 for 16 hours and 20 hours, respectively, at the end of the baseline schedule; note that the end of these continuous wake durations is during the circadian "night" and therefore during circadian misalignment. The threshold for moderate impairment was defined as the number of PVT lapses predicted by the model at 16–20 hours of continuous wakefulness: 8.7–15.2 lapses. Similarly, the threshold for high impairment was defined as the number of PVT lapses predicted by the model at 20 hours of continuous wakefulness: >15.2 lapses.

Analysis

As primary outcomes for each schedule, we computed the average number of hours per week (i.e., per 168 hours) during working hours corresponding to: (i) moderate impairment; and (ii) high impairment. The EDWR and RCWR conditions were first compared for all weeks of data in all participant schedules. Because the impairment time distributions were highly skewed, a nonparametric comparison was used (Wilcoxon rank sum test).

To measure acute vs. chronic effects of each schedule, number of hours of moderate impairment and high impairment were computed for weeks 1+2 and for weeks 3+4 on each schedule, using all available data within the time range. A mixed-effects regression model was fit to the primary outcomes, including schedule and week block as fixed effects and participants as random effects. The week 1–2 block in the EDWR condition (i.e., first time point in the non-intervention condition) was used as the baseline for contrasts. An interaction between schedule and week block was included. Normality of residuals was checked with Q-Q plots. Mixed-effects models were implemented in MATLAB (R2015a, Mathworks, Natick MA, USA) using the function fitlme. Absolute effect sizes and their 95% confidence intervals were reported.

To examine temporal patterns in performance, the mean of predicted PVT lapses were computed for each schedule in 1-h clock-time bins across the schedule. Two methods of averaging were used for each schedule: (i) using all hours of wakefulness, and (ii) using all work hours. To compare between schedule conditions, these temporal patterns were averaged across schedules within each condition, and 1-h timepoint-wise comparisons between conditions were made using t-tests after running an ANOVA to confirm differences between conditions. Differences were considered significant if $p < 0.002$ (with Bonferroni correction for multiple comparisons).

As a secondary analysis, we investigated whether performance differences within the EDWR or RCWR conditions could be explained by total average sleep time. For each individual schedule, we computed the average percentage of time spent asleep in 1-h clock-time bins from start to end of schedule, so that estimates would not be biased by clock timing of missing data. The average total sleep time for an individual schedule was computed by averaging across clock-time bins. Spearman correlations were computed between average total sleep time and hours per week in a predicted moderate or highly impaired state.

As a further secondary analysis, we investigated whether performance differences within the EDWR or RCWR conditions could be explained by the frequency of short between-shift intervals (defined as intervals <10 h in duration), since they could affect sleep timing and duration. Specifically, we computed the average number of short between-shift intervals per week (168 hours). Spearman correlations were computed between the frequency of short between-shift intervals and hours per week in a moderate or highly impaired state.

Results

Examples of predicted amount of time spent in a moderately or highly impaired state on EDWR and RCWR are shown in Figure 1. On the EDWR, there was a range of 5.0–47.1 hours per week spent in a moderately impaired state (median=15.5 h), while there was a range of 1.4–32.2 hours per week (median 13.5 h) spent in a highly impaired state. On the RCWR, there was a range of 3.6–37.1 hours per week (median=10.8 h) spent in a moderately impaired state, while there was a range of 0–37.6 hours per week (median 12.0 h) spent in a highly impaired state (Figure 2). At this schedule level, time spent in a moderately impaired state was significantly higher for EDWR compared to RCWR (p=0.02; effect size = 0.2). Time spent in a highly impaired state was not significantly different between schedules (p=0.11).

Mixed-effects modeling showed significant main effects of the week block and a significant interaction of condition with week block. For the main effect of week block, we found that compared to the week 1–2 block, time spent in a moderately impaired state per week was increased by 2.6 hours (p=0.001, 95% CI: 1.0–4.2 hours) and time spent in a highly impaired state was increased by 1.7 hours (p=0.001, 95% CI: 0.7–2.8 hours) in the week 3–4 block. For the interaction of condition with week block, we found that compared to RCWR, time spent in a moderately impaired state for EDWR was further increased from weeks 1–2 to weeks 3–4 by 3.4 hours (p=0.003, 95% CI: 1.2–5.7 hours), whereas time spent in a highly impaired state had no significant interaction (0.6 hours, p=0.46, 95% CI: −0.9–2.0 hours). There was no main effect of condition, meaning time spent in a moderately or highly impaired state in the week 1–2 block did not significantly differ between conditions.

Differences in predicted performance with time of day were observed between the schedules (Figure 3). For the analysis of performance across all hours of wakefulness (Figure 3A), there was a main effect of condition (p<0.0001) and time (p<0.0001). In pairwise comparisons, performance was significantly worse on the EDWR than the RCWR in all clock time bins from 21:00 to 04:00. Similarly, for the analysis of performance across working hours (Figure 3B), there was a main effect of condition (p<0.0001) and time

(p<0.0001). In pairwise comparisons, performance was significantly worse on the EDWR than the RCWR in all clock time bins from 19:00 to 05:00, whereas performance was significantly better on the EDWR than the RCWR in all clock time bins from 08:00 to 11:00.

Short between-shift intervals were significantly less common on the EDWR than the RCWR (0.3% vs.8.4%, p<0.0001)[47]. Within the EDWR condition, the frequency of short between-shift intervals was not significantly associated with predicted number of hours in a moderately impaired state (r=−0.15, p=0.07) or highly impaired state (r=−0.02, p=0.84). Within the RCWR condition, however, the frequency of short between-shift intervals was significantly associated with predicted number of hours in a moderately impaired state (r=0.39, p<0.000001) but not with predicted number of hours in a highly impaired state (r=−0.01, p=0.92).

Predicted performance was highly correlated with sleep duration (Figure 4). On average, total subjective sleep time (from daily logs) was significantly lower on the EDWR schedules compared with the RCWR schedules (6.7 ± 0.7 h vs. 7.0 ± 0.7 h; p=0.002, effect size = 0.20). Spearman correlations of total subjective sleep time with work hours spent in a moderately impaired state were significant for the EDWR (r=−0.85, p<0.001) and RCWR (r=−0.81, p<0.001). Spearman correlations of total subjective sleep time with work hours spent in a highly impaired state were also significant for both the EDWR (r=−0.63, p<0.001) and RCWR (r=−0.70, p<0.001).

Discussion

We used a mathematical model to compare traditional EDWR schedules with RCWR schedules (that were designed to ameliorate impacts of working hours on sleep and neurobehavioral performance) using predicted neurobehavioral performance during 169 resident physicians' sleep and work schedules. This is the first example of linking three key components for real-world performance prediction in a single study: (i) a dynamic model of the human circadian pacemaker and its response to light/dark and sleep/wake patterns; (ii) a model of effects of chronic (cumulative) sleep restriction on performance; and (iii) actual schedules of self-selected sleep/wake timing with associated realistic light/dark patterns. Our findings indicate that the RCWR intervention is partially effective in reducing the predicted neurobehavioral performance impairments generated by the traditional EDWR. Specifically, the RCWR intervention was found to be most beneficial during the circadian (biological) night and least beneficial in late morning.

A key strength of this analysis is the use of real sleep and work patterns, rather than idealized or planned schedules. Within both the EDWR and RCWR schedules, we observed considerable interindividual and intra-individual variability in both sleep and work patterns. This variability is representative of real-world differences in compliance, adherence, work/ social constraints, and physiology, as well as schedule implementation[47]. Using the model across these datasets, we were able to identify two concrete areas of future improvement for RCWR schedules. First, we found a specific window of vulnerability in performance during work hours from 08:00 to 11:00, which is not predicted to be present on EDWR schedules. Second, we identified that short between-shift intervals are predictive of a greater

number of hours working in a moderately impaired state specifically for RCWR schedules. This prediction suggests that carefully controlling the duration of between-shift breaks could enhance efficacy of the RCWR intervention.

Our work predicts chronic deterioration of performance on both RCWR and EDWR schedules. Compared to EDWR, the rate of deterioration is predicted to be lower by 57% on the RCWR. As experimental studies have demonstrated, the effects of chronic sleep restriction can accumulate over weeks[16,38,52,53]. Moreover, if chronic sleep restriction is alternated with recovery breaks that are of insufficient duration to achieve full recovery to baseline, the rate of cumulative decline in neurobehavioral performance may not be slowed by the recovery intervals[19]. Studies of neurobehavioral performance in resident physicians to date have all been of less than a month in duration[12,54]. Our findings suggest the need for longitudinal studies, across months or years, to assess the long-term impacts of resident physician schedules on sleep and neurobehavioral performance.

We note that the model used in this study predicts objective performance on the PVT, rather than subjective measures of alertness. Subjective alertness, which is used by individuals to decide if they need a countermeasure (e.g., nap or caffeine) and/or if they can safely perform (e.g., driving home) is not well correlated with objective performance when individuals are exposed to chronic sleep restriction[19,38,53,55,56]. Adding subjective alertness to the model could improve its use in education about the dangers of insufficient sleep, including its effects on objective performance.

There are possible limitations of the method presented here. First, all components of the model have been trained on healthy young adults with intermediate chronotypes (i.e., excluding individuals whose preferred sleep timing is extremely early or extremely late). Predictions may therefore not be representative for individuals of other ages; with sleep, circadian, or other disorders; and/or extreme chronotypes. In the future, the model could be used to investigate sensitivity of the predictions to physiological parameters that determine chronotype, as has been done previously for diurnal schedules[44,48], since different individuals may respond differently to the same work or sleep schedule[57,58]. Second, this model does not yet incorporate the effects of caffeine or other pharmaceuticals on sleep and performance. Models of caffeine's effects on sleep and neurobehavioral performance have been developed[59–61] and should be incorporated in future. Models of the effects of prescription medications need to be developed. Third, the model predictions generated from these real work patterns have not been validated against data. The models, however, have previously been validated on data from other experimental studies of chronic sleep restriction and circadian misalignment of sleep/wake and light/dark patterns[37]. Validation on experimental data from RCWR and EDWR schedules is an important next step to confirm whether some of the differences observed between RCWR and EDWR, such as the worse performance predicted from 08:00 to 11:00 on the RCWR schedule, are observed during actual work shifts. Finally, the model predictions do not currently account for effects of sleep inertia on neurobehavioral performance[62,63]. Notably, many of the participants on EDWR schedules took naps during the extended duration shifts, followed by immediate return to work. The model simulates the restorative properties of napping, but not the sleep inertia, which may be impactful even for short naps[25].

In previous work with the chronic sleep restriction component of the model, the model was also used to generate accurate predictions of sleep/wake patterns[37]. Here, we chose instead to use the reported sleep/wake patterns for each schedule as input to the model. This selection was made for two reasons: (i) the model has only been validated for sleep/wake predictions on daytime work schedules, and sleep during rotating shiftwork will be more difficult to predict; and (ii) to ensure that performance predictions are not systematically inaccurate due to any potential inaccuracies in the model's predictions of sleep/wake timing. In future work, the model's sleep/wake predictions could be trained against this and other datasets. This would allow model-based prospective assessment of candidate schedules in the absence of any worker data to determine which schedules allow the most opportunity for sleep with the least amount of time spent in a moderately or highly impaired state.

Modeling is an essential component of effective occupational risk assessment, yet the use of models in medical settings remains very limited[26,64]. The modeling approach presented here could be employed in an active role in future medical residents or other schedule design and optimization (e.g., transportation [airline pilots, train conductors, ship captains, truck drivers], military/security, nuclear power plant monitoring) for individuals and teams. Although there are many factors – both fiscal and logistical – that constrain the design of resident physician schedules, our findings underscore the importance of considering the effects of sleep loss and adverse circadian phase, especially given the high cost of failure to both patient safety and physician health in the medical setting or operator and public safety in other settings.

Acknowledgments

Ayoub Lahlou, Kathryn Malone, Karen Meyer, Maryam Nuriyeva, Tolulope D. Oyewumi, Kimberly Ralston, Adam Rosenberg, Nabeel D. Sawaged MD, Beth Smith MD, Vanitha K. Varre and Gaurav Yadav.

Iowa City, IA: Rasha Abdalla MD, Safa Abukhalil MD, Ihab Ahmed MBBS, Safa Ahmed MBBS, Shilpa Balikai DO, Maria Ana Canaya-Voskov MD, Gretchen Cress MPH, Gwen Erkonen MD M.Ed, Alla Hamed MD, Janice Jeter RN, Sameer Kamath MD MBBS, Crystal Lamansky RN, Vani Movva MBBS, Geoffrey Obel MD, Angela Platt, Ruthann Schrock RN, Jeffrey Segar MD, Amy Stier MD, Alexandra Paige Volk MD, and Jin Zhou RN.

Seattle, WA: Mouammar Abouagila MD, Canan Akture MD, Maneesh Batra MD MPH, Courtney Donohue, Reid Farris MD MS, Jennifer J. Gile, Richa Kashayap MD, John K. McGuire MD, Carol Mendivil MD, Anas Najjar MD, Gowri Rajendran MBBS, Shahar Robinson MD, Erin Sullivan, and Nastassya West.

References

1. Alhola P, Polo-Kantola P. Sleep deprivation: Impact on cognitive performance. Neuropsychiatr Dis Treat 2007;3:553–67. [PubMed: 19300585]

2. Goel N, Basner M, Rao H, Dinges DF. Circadian rhythms, sleep deprivation, and human performance. Prog Mol Biol Transl Sci 2013;119:155–90. [PubMed: 23899598]

3. Lowe CJ, Safati A, Hall PA. The neurocognitive consequences of sleep restriction: A meta-analytic review. Neuroscience and Biobehavioral Reviews 2017;80:586–604. [PubMed: 28757454]

4. Tkachenko O, Dinges DF. Interindividual variability in neurobehavioral response to sleep loss: A comprehensive review. Neuroscience and Biobehavioral Reviews 2018;89:29–48. [PubMed: 29563066]

5. Åkerstedt T, Fredlund P, Gillberg M, Jansson B. A prospective study of fatal occupational accidents-relationship to sleeping difficulties and occupational factors. J Sleep Res 2002;11:69–71. [PubMed: 11869429]

6. Barger LK, Lockley SW, Rajaratnam SM, Landrigan CP. Neurobehavioral, health, and safety consequences associated with shift work in safety-sensitive professions. Curr Neurol Neurosci Rep 2009;9:155–64. [PubMed: 19268039]

7. Drake C, Roehrs T, Breslau N, et al. The 10-year risk of verified motor vehicle crashes in relation to physiologic sleepiness. Sleep 2010;33:745–52. [PubMed: 20550014]

8. Folkard S, Lombardi DA. Modeling the impact of the components of long work hours on injuries and "accidents". Am J Ind Med 2006;49:953–63. [PubMed: 16570251]

9. Lee ML, Howard ME, Horrey WJ, et al. High risk of near-crash driving events following night-shift work. Proceedings of the National Academy of Sciences of the United States of America 2016;113:176–81. [PubMed: 26699470]

10. Martiniuk AL, Senserrick T, Lo S, et al. Sleep-deprived young drivers and the risk for crash: the DRIVE prospective cohort study. JAMA Pediatr 2013;167:647–55. [PubMed: 23689363]

11. Lockley SW, Cronin JW, Evans EE, et al. Effect of reducing interns' weekly work hours on sleep and attentional failures. NEnglJMed 2004;351:1829–37.

12. Anderson C, Sullivan JP, Flynn-Evans EE, Cade BE, Czeisler CA, Lockley SW. Deterioration of neurobehavioral performance in resident physicians during repeated exposure to extended duration work shifts. Sleep 2012;35:1137–46. [PubMed: 22851809]

13. Gordon JA, Alexander EK, Lockley SW, et al. Does simulator-based clinical performance correlate with actual hospital behavior? The effect of extended work hours on patient care provided by medical interns. Acad Med 2010;85:1583–8. [PubMed: 20881679]

14. Landrigan CP, Rothschild JM, Cronin JW, et al. Effect of reducing interns' work hours on serious medical errors in intensive care units. NEnglJMed 2004;351:1838–48.

15. Buxton OM, Cain SW, O'Connor SP, et al. Adverse metabolic consequences in humans of prolonged sleep restriction combined with circadian disruption. Science Translational Medicine 2012;4:129ra43.

16. Cohen DA, Wang W, Wyatt JK, et al. Uncovering residual effects of chronic sleep loss on human performance. Science Translational Medicine 2010;2:14ra3.

17. Darwent D, Ferguson SA, Sargent C, et al. Contribution of core body temperature, prior wake time, and sleep stages to cognitive throughput performance during forced desynchrony. Chronobiology International 2010;27:898–910. [PubMed: 20636204]

18. Zhou X, Ferguson SA, Matthews RW, et al. Sleep, Wake and Phase Dependent Changes in Neurobehavioral Function under Forced Desynchrony. Sleep 2011;34:931–41. [PubMed: 21731143]

19. St Hilaire MA, Ruger M, Fratelli F, Hull JT, Phillips AJ, Lockley SW. Modeling neurocognitive decline and recovery during repeated cycles of extended sleep and chronic sleep deficiency. Sleep 2016.

20. Lockley SW, Landrigan CP, Barger LK, Czeisler CA. When policy meets physiology: The challenge of reducing resident work hours Clin OrthopRelatRes 2006;449:116–27.

21. Lockley SW, Barger LK, Ayas NT, Rothschild JM, Czeisler CA, Landrigan CP. Effects of health care provider work hours and sleep deprivation on safety and performance. Jt Comm J Qual Patient Saf 2007;33:7–18.

22. Kamine TH, Barron RJ, Lesicka A, Galbraith JD, Millham FH, Larson J. Effects of the new Accreditation Council for Graduate Medical Education work hour rules on surgical interns: a prospective study in a community teaching hospital. American Journal of Surgery 2013;205:163–8. [PubMed: 23331981]

23. Levine AC, Adusumilli J, Landrigan CP. Effects of reducing or eliminating resident work shifts over 16 hours: a systematic review. Sleep 2010;33:1043–53. [PubMed: 20815185]

24. Hursh SR, Eddy DR. Fatigue modeling as a tool for managing fatigue in transportation operations. In: 2005 International Conference on Fatigue Management in Transportation; 2005 September 2005; Seattle: US Federal Motor Carrier Safety Administration; 2005.

25. Hilditch CJ, Dorrian J, Banks S. A review of short naps and sleep inertia: do naps of 30 min or less really avoid sleep inertia and slow-wave sleep? Sleep Medicine 2017;32:176–90. [PubMed: 28366332]

26. Klerman EB, Beckett SA, Landrigan CP. Applying mathematical models to predict resident physician performance and alertness on traditional and novel work schedules. BMC Med Educ 2016;16:239. [PubMed: 27623842]

27. Mallis MM, Mejdal S, Nguyen TT, Dinges DF. Summary of the key features of seven biomathematical models of human fatigue and performance. AviatSpace Environ Med 2004;75:A4–14.

28. Postnova S, Layden A, Robinson PA, Phillips AJ, Abeysuriya RG. Exploring sleepiness and entrainment on permanent shift schedules in a physiologically based model. JBiol Rhythms 2012;27:91–102. [PubMed: 22306977]

29. Postnova S, Postnov DD, Seneviratne M, Robinson PA. Effects of Rotation Interval on Sleepiness and Circadian Dynamics on Forward Rotating 3-Shift Systems. JBiol Rhythms 2014;29:60–70. [PubMed: 24492883]

30. Serkh K, Forger DB. Optimal schedules of light exposure for rapidly correcting circadian misalignment. PLoS Computational Biology 2014;10:e1003523. [PubMed: 24722195]

31. Asgari-Targhi A, Klerman EB. Mathematical modeling of circadian rhythms. Wiley Interdiscip Rev Syst Biol Med 2018:e1439. [PubMed: 30328684]

32. Dean DA II, Forger DB, Klerman EB. Taking the lag out of jet lag through model-based schedule design. PLoS Comput Biol 2009;5:e1000418. [PubMed: 19543382]

33. St Hilaire MA, Klerman EB, Khalsa SBS, Wright KP Jr., Czeisler CA, Kronauer RE. Addition of a non-photic component to a light-based mathematical model of the human circadian pacemaker. J Theoretical Biol 2007;247:583–99.

34. Hursh SR, Redmond DP, Johnson ML, et al. Fatigue models for applied research in warfighting. AviatSpace Environ Med 2004;75:A44–A53.

35. Klerman EB, St Hilaire MA. On mathematical modeling of circadian rhythms, performance, and alertness. JBiol Rhythms 2007;22:91–102. [PubMed: 17440211]

36. Jewett ME, Kronauer RE. Interactive mathematical models of subjective alertness and cognitive throughput in humans. JBiol Rhythms 1999;14:588–97. [PubMed: 10643756]

37. Phillips AJK, Klerman EB, Butler JP. Modeling the adenosine system as a modulator of cognitive performance and sleep patterns during sleep restriction and recovery. PLoS Computational Biology 2017;13:e1005759. [PubMed: 29073206]

38. Van Dongen HPA, Maislin G, Mullington JM, Dinges DF. The cumulative cost of additional wakefulness: Dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. Sleep 2003;26:117–26. [PubMed: 12683469]

39. Lim J, Dinges DF. Sleep deprivation and vigilant attention. Ann N Y Acad Sci 2008;1129:305–22. [PubMed: 18591490]

40. Arnedt JT, Owens J, Crouch M, Stahl J, Carskadon MA. Neurobehavioral performance of residents after heavy night call vs after alcohol ingestion. JAMA 2005;294:1025–33. [PubMed: 16145022]

41. Elmenhorst D, Elmenhorst EM, Luks N, et al. Performance impairment during four days partial sleep deprivation compared with the acute effects of alcohol and hypoxia. Sleep Medicine 2009;10:189–97. [PubMed: 18276188]

42. Fulcher BD, Phillips AJK, Robinson PA. Quantitative physiologically based modeling of subjective fatigue during sleep deprivation. Journal of Theoretical Biology 2010;264:407–19. [PubMed: 20176034]

43. Phillips AJK, Czeisler CA, Klerman EB. Revisiting spontaneous internal desynchrony using a quantitative model of sleep physiology. JBiol Rhythms 2011;26:441–53. [PubMed: 21921298]

44. Swaminathan K, Klerman EB, Phillips AJK. Are Individual Differences in Sleep and Circadian Timing Amplified by Use of Artificial Light Sources? JBiol Rhythms 2017;32:165–76. [PubMed: 28367676]

45. Blackwell T, Kriesel DR, Vittinghoff E, et al. ; ROSTERS Study Group. Design and recruitment of the randomized order safety trial evaluating resident-physician schedules (ROSTERS) study. Contemp Clin Trials. 2019 May;80:22–33. [PubMed: 30885799]

46. Rahman SA, Sullivan JP, Barger LK et al. ; ROSTERS Study Group. Extended Work Shifts and Neurobehavioral Performance in Resident-Physicians. Pediatrics. 2021 Mar;147(3):e2020009936. [PubMed: 33619044]

47. Barger LK, Sullivan JP, Blackwell T, et al. Effects on resident work hours, sleep duration, and work experience in a randomized order safety trial evaluating resident-physician schedules (ROSTERS). Sleep. 2019 Aug 1;42(8):zsz110. [PubMed: 31106381]

48. Skeldon AC, Phillips AJ, Dijk DJ. The effects of self-selected light-dark cycles and social constraints on human sleep and circadian timing: a modeling approach. Sci Rep 2017;7:45158. [PubMed: 28345624]

49. Savides TJ, Messin S, Senger C, Kripke DF. Natural light exposure of young adults. Physiol Behav. 1986;38:571–574. [PubMed: 3823171]

50. Okudaira N, Kripke DF, Webster JB. Naturalistic studies of human light exposure. Am J Physiol. 1983;245:R613–R615. [PubMed: 6624956]

51. Espiritu RC, Kripke DF, Ancoli-Israel S, Mowen MA, Mason WJ: Natural light exposure of adults 40–64 years old. Sleep Res. 1992;21:374.

52. Belenky G, Wesensten NJ, Thorne DR, et al. Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: A sleep dose-response study. J Sleep Res 2003;12:1–12. [PubMed: 12603781]

53. McHill AW, Hull JT, Wang W, Czeisler CA, Klerman EB. Chronic sleep curtailment, even without extended (>16-h) wakefulness, degrades human vigilance performance. Proceedings of the National Academy of Sciences of the United States of America 2018;115:6070–5. [PubMed: 29784810]

54. Basner M, Dinges DF, Shea JA, et al. Sleep and Alertness in Medical Interns and Residents: An Observational Study on the Role of Extended Shifts. Sleep 2017;40.

55. Bermudez EB, Klerman EB, Czeisler CA, Cohen DA, Wyatt JK, Phillips AJ. Prediction of Vigilant Attention and Cognitive Performance Using Self-Reported Alertness, Circadian Phase, Hours since Awakening, and Accumulated Sleep Loss. PLoS ONE 2016;11:e0151770. [PubMed: 27019198]

56. Zhou X, Ferguson SA, Matthews RW, et al. Mismatch between subjective alertness and objective performance under sleep restriction is greatest during the biological night. J Sleep Res 2012;21:40–9. [PubMed: 21564364]

57. Juda M, Vetter C, Roenneberg T. Chronotype modulates sleep duration, sleep quality, and social jet lag in shift-workers. JBiol Rhythms 2013;28:141–51. [PubMed: 23606613]

58. Vetter C, Fischer D, Matera JL, Roenneberg T. Aligning work and circadian time in shift workers improves sleep and reduces circadian disruption. Current biology 2015;25:907–11. [PubMed: 25772446]

59. Puckeridge M, Fulcher BD, Phillips AJK, Robinson PA. Incorporation of caffeine into a quantitative model of fatigue and sleep. Journal of Theoretical Biology 2011;273:44–54. [PubMed: 21176782]

60. Ramakrishnan S, Rajaraman S, Laxminarayan S, et al. A biomathematical model of the restoring effects of caffeine on cognitive performance during sleep deprivation. Journal of Theoretical Biology 2013;319:23–33. [PubMed: 23182694]

61. Ramakrishnan S, Wesensten NJ, Kamimori GH, Moon JE, Balkin TJ, Reifman J. A Unified Model of Performance for Predicting the Effects of Sleep and Caffeine. Sleep 2016;39:1827–41. [PubMed: 27397562]

62. Achermann P, Werth E, Dijk DJ, Borbély AA. Time course of sleep inertia after nighttime and daytime sleep episodes. Archives Italiennes de Biologie 1995;134:109–19. [PubMed: 8919196]

63. Jewett ME, Wyatt JK, Ritz-De Cecco A, Khalsa SB, Dijk DJ, Czeisler CA. Time course of sleep inertia dissipation in human performance and alertness. J Sleep Res 1999;8:1–8.

64. James FO, Waggoner LB, Weiss PM, et al. Does Implementation of Biomathematical Models Mitigate Fatigue and Fatigue-related Risks in Emergency Medical Services Operations? A Systematic Review. Prehosp Emerg Care 2018;22:69–80. [PubMed: 29324070]

Public Health Relevance

This work highlights the utility of mathematical models for evaluating the predicted performance profile of work schedules in safety-sensitive occupations that are susceptible to chronic sleep restriction and circadian misalignment, including resident physicians; individuals working in transportation, military/security, and industry; and others.
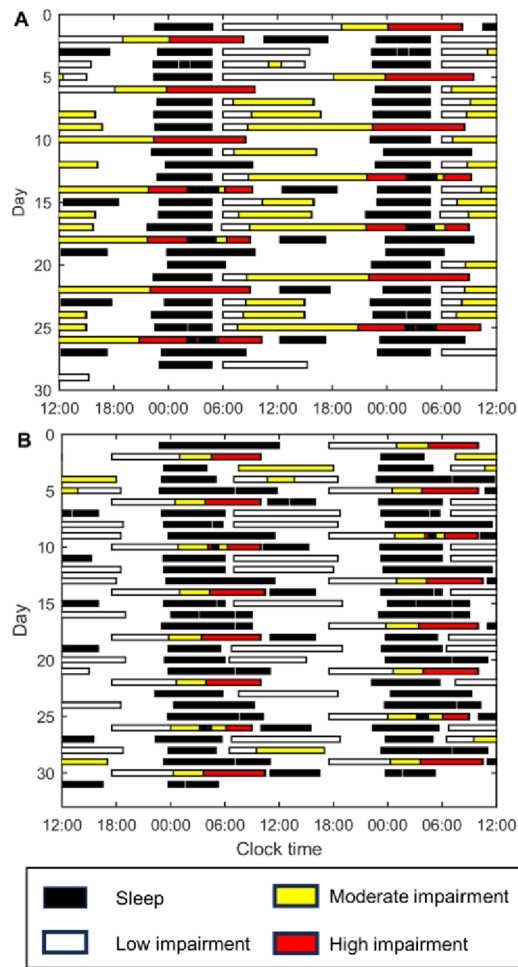
Figure 1:
Double raster plots for examples of (A) the Extended Duration Work Roster (EDWR) and (B) the Rapidly Cycling Work Roster (RCWR). Time intervals corresponding to sleep are marked by black bars. Time intervals corresponding to work are marked by white, yellow, and red bars, which represent times of predicted low, moderate, and high performance impairment during work, respectively.
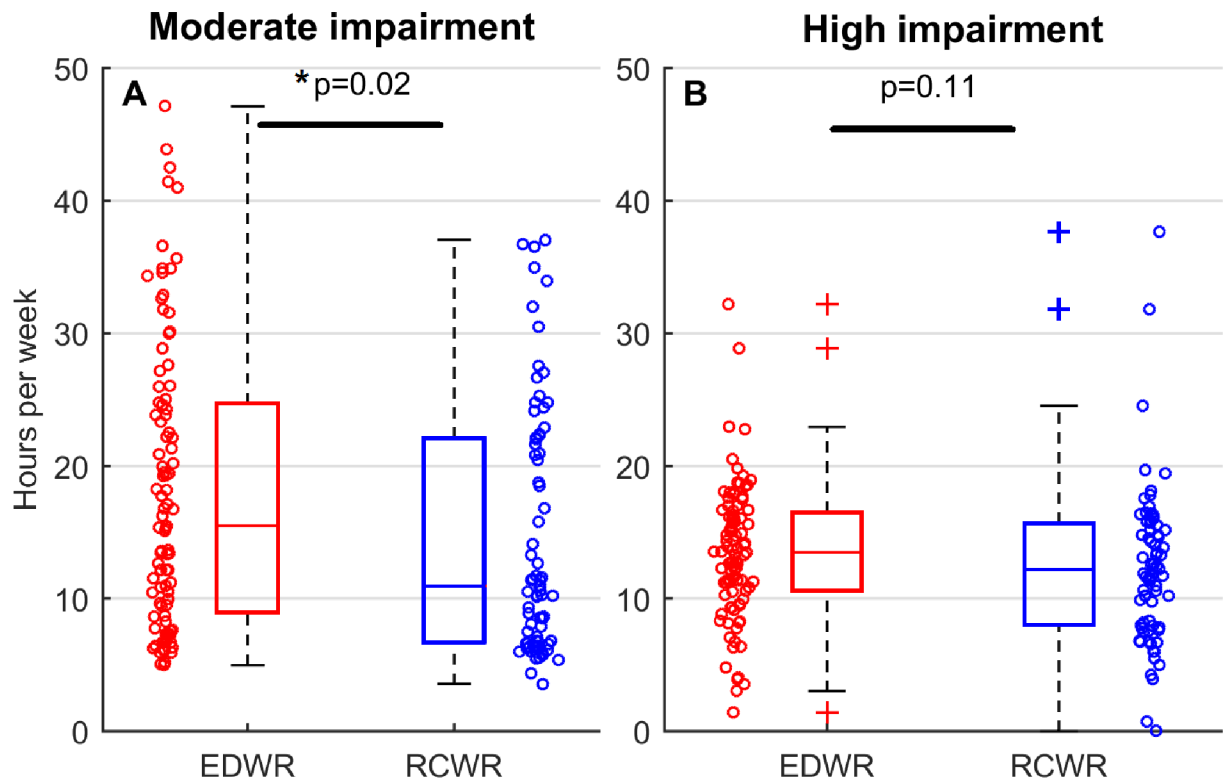
Figure 2:
Predicted hours per week during work hours spent under (A) moderate impairment (equivalent to 16–20 hours of continuous wakefulness); and (B) high impairment (equivalent to >20 hours of continuous wakefulness). Distributions for the Extended Duration Work Roster (EDWR) and the Rapidly Cycling Work Roster (RCWR) are shown as box-and-whisker plots. Individual data points (169 schedules total; 95 EDWRs and 74 RCWRs) are plotted as circles, with x-axis position jittered to reduce overlap of points. Each individual point in each panel is one simulated schedule.
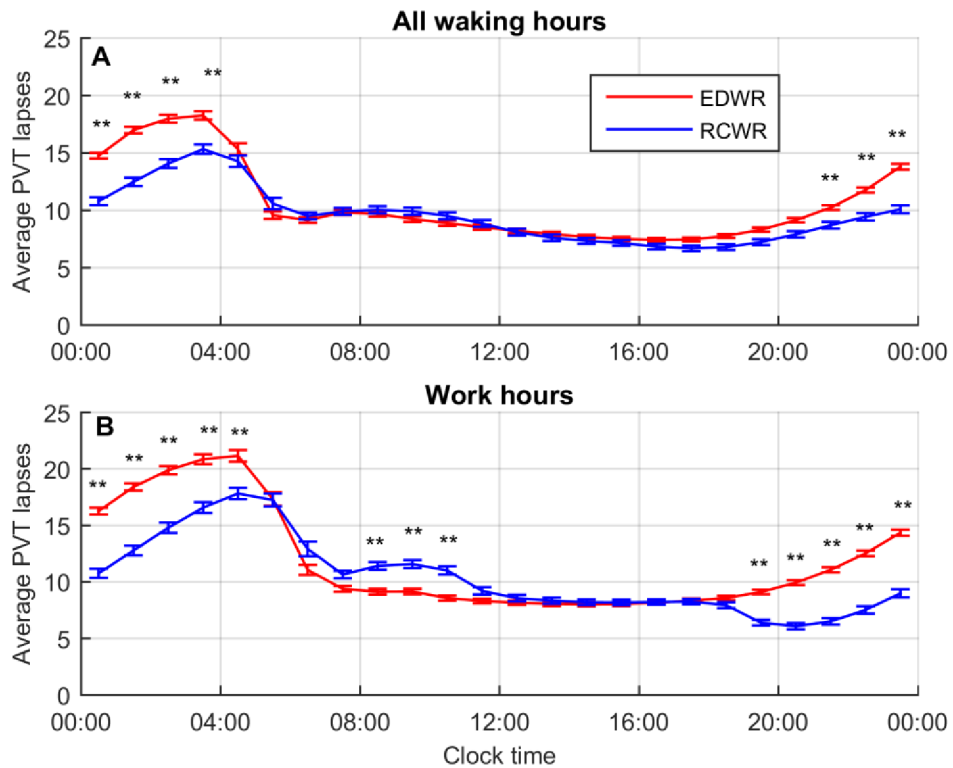
Figure 3:
Mean predicted attentional failures as measured by PVT lapses for Extended Duration Work Roster (EDWR) and Rapidly Cycling Work Roster (RCWR) schedule conditions, plotted by time of day averaged across all data sets (169 schedules total; 95 EDWRs and 74 RCWRs) in 1-hour clock time bins. Data are plotted for (A) all hours of wakefulness, and (B) working hours, and plotted at the midpoint of each clock hour. Error bars show mean ± SEM. Significance is shown for timepoint-wise comparisons: *p<0.002, **p<0.0001.
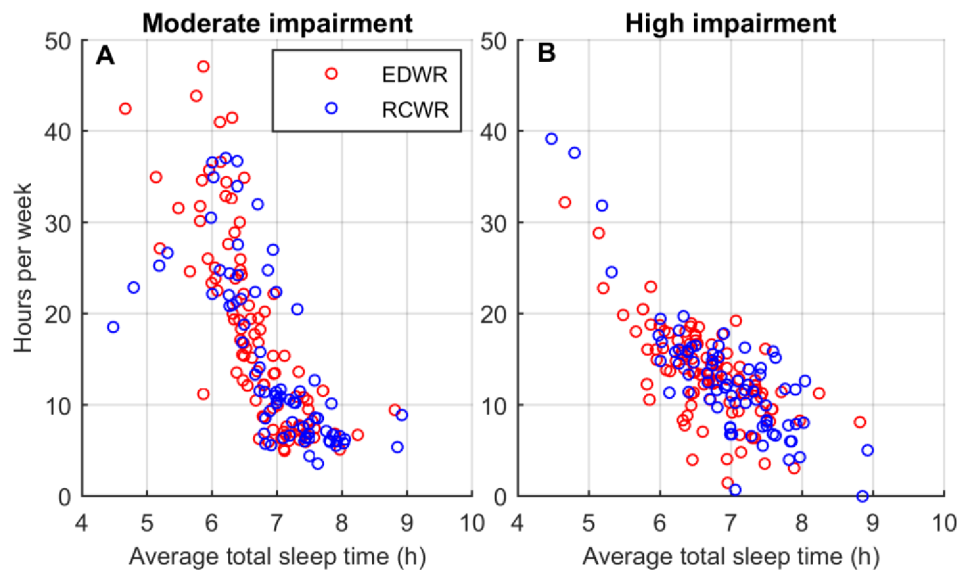
Figure 4:
Associations of average total subjective sleep time across the entire schedule for all 169 schedules (95 EDWRs and 74 RCWRs) with predicted hours per week spent in (A) moderate impairment during work (equivalent to 16–20 hours of continuous wakefulness); and (B) high impairment during work (equivalent to >20 hours of continuous wakefulness). Each individual point in each panel is one simulated schedule.