Temporal dynamics of ocular indicators of sleepiness across sleep restriction

Suzanne Ftouni[1,2,3], **Shadab A. Rahman**[1,2], **Kate E. Crowley**[1,2], **Clare Anderson**[1,2,3], **Shantha M.W. Rajaratnam**[1,2,3], and **Steven W. Lockley**[1,2,3]

[1]Division of Sleep Medicine, Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA

[2] Division of Sleep Medicine, Department of Medicine, Harvard Medical School, Boston, MA, USA

[3]School of Psychology and Psychiatry, Monash University, Melbourne, Victoria, Australia

Abstract

The current study characterized the temporal dynamics of ocular indicators of sleepiness during extended sleep restriction. Ten male participants (mean age ± SD = 23.3 ± 1.6 years) underwent 40-hours of continuous wakefulness under constant routine (CR) conditions, and completed the Karolinska Sleepiness Scale (KSS) and a 10-minute auditory psychomotor vigilance task (aPVT) hourly. Waking electroencephalography (EEG) and ocular measures were recorded continuously throughout the CR. Infrared-reflectance oculography was used to collect the ocular measures positive and negative amplitude/velocity ratio, mean blink duration, the percentage of eye closure, and a composite score of sleepiness levels (Johns Drowsiness Scale). All ocular measures except blink duration, displayed homeostatic and circadian properties. Only circadian effects were detected in blink duration. Significant, phase-locked cross-correlations ($p < 0.05$) were detected between ocular measures and aPVT reaction time (RT), aPVT lapses, KSS, and EEG delta-theta (0.5-5.5 Hz), theta-alpha (5.0-9.0 Hz) and beta (13.0-20.0 Hz) activity. Receiver Operating Characteristic (ROC) curve analysis demonstrated reasonable sensitivity and specificity of ocular measures in correctly classifying aPVT lapses above individual baseline thresholds (initial 16 h of wakefulness). Under conditions of sleep restriction, ocular indicators of sleepiness paralleled performance impairment and self-rated sleepiness levels, and demonstrated their potential to detect sleepiness-related attentional lapses. These findings, if reproduced in a larger sample, will have implications on the use of ocular based sleepiness-warning systems in operational settings.

Corresponding author: Steven W. Lockley, Ph.D., Division of Sleep Medicine, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, 221 Longwood Avenue, Boston, MA 02115. Telephone: 617–732–4977, slockley@hms.harvard.edu.

Keywords

Sleep deprivation; ocular measures; oculography; electroencephalography; performance; constant routine; circadian rhythms

Introduction

Eye and eyelid movements have proven useful in quantifying levels of sleepiness that can reflect performance or attentional impairment (Cajochen et al. 1999; Lockley et al. 2004; Torsvall and Åkerstedt 1987). Eye blinks associated with the onset of sleep or microsleeps, often experienced by sleep-deprived individuals, display unique properties which reflect sleep- and wake-related central nervous system (CNS) changes (Stern et al. 1984). This consequential change to distinct blink characteristics is linked to the close proximity of neural centers regulating the facial nerves that control eye movements, and the medullary structures associated with the sleep/wake process (Morris and Miller 1996; Stern et al. 1984). Independent studies have assessed the sleepiness-related changes in ocular measures, such as the amplitude/velocity ratio of blinks (Anderson et al. 2013; Johns et al. 2008; Michael et al. 2008; Picot et al. 2012; Vadeby et al. 2010), blink duration (Anund et al. 2008; Caffier et al. 2003; Häkkänen et al. 1999; Picot et al. 2012; Schleicher et al. 2008), blink closing and re-opening durations (Caffier et al. 2003; Picot et al. 2012), and percentage of eye closures (PERCLOS) (Abe et al. 2011; Chua et al. 2012; Dinges and Grace 1998; Schleicher et al. 2008; Wierwille and Ellsworth 1994), in relation to well-researched methods of sleepiness quantification. These ocular measures correspond to increases in a marker of sleep onset, slow eye movements (Cajochen et al. 1999; De Gennaro et al. 2005; Henn et al. 1984; Picot et al. 2012), and measures of sleep propensity, such as the multiple sleep latency test and maintenance of wakefulness test (De Gennaro et al. 2005; Häkkänen et al. 1999). As such, sleepiness-detection systems have been developed around ocular measures (Bergasa et al. 2006; Johns et al. 2008; Johns et al. 2007) based on their temporal relationship with sleep onset and associated performance decline (Anderson et al. 2013; Chua et al. 2012; Michael et al. 2008; Van Dongen et al. 2003).

While ocular measures regulated by the CNS are described as reliable state measures of sleepiness (Stern et al. 1984), it is yet to be established whether ocular measures are regulated by the same physiological processes that regulate the sleep-wake cycle. The temporal dynamics of alertness are regulated by the endogenous rhythm of sleep propensity which peaks at night and is generated by the circadian pacemaker (Process C) – and sleep pressure, which is generated by a homeostatic process (Process S) and predicts increased sleepiness with increased time awake (Borbely 1982). Although various papers have examined the many ocular measures available in relation to sleepiness, fewer have assessed the temporal variations in ocular measures derived from sleepiness-detection devices developed for operational settings, as well as their association with changes in alertness and neurobehavioural performance (Abe et al. 2011; Anderson et al. 2013; Chua et al. 2012; Dinges and Grace 1998). Therefore, the aim of this study was to assess changes in ocular measures derived from a sleepiness-detection device (Optalert™), over 40 hours of sleep deprivation under highly-controlled laboratory conditions, to characterize the circadian and

time awake-dependent (homeostatic) changes in these measures. The time course of these variables was correlated with the time course of attentional performance, EEG based measures of alertness, and subjective sleepiness. In addition, the study aimed to assess the predictability of ocular measures in detecting lapses in performance.

Methods

Ethical approval

The study was approved by the Partners Human Research Committee and the Monash University Human Research Ethics Committee. All participants gave written informed consent prior to the study.

Participants and pre-study conditions

Ten healthy males (mean age ± SD = 23.3 ± 1.6 years; range 20-25 years) were studied in the Intensive Physiology Monitoring Unit in the Centre for Clinical Investigation (CCI) at Brigham and Women's Hospital. All had comprehensive physical, psychological and ophthalmologic examinations, including a negative Ishihara color-blindness test (Ishihara 1996). For at least 3 weeks prior to entering the Intensive Physiology Monitoring Unit, participants maintained a self-selected, constant 8 h sleep/rest/dark schedule confirmed with calls to a time- and date-stamped voicemail at bedtime and wake time for 3 weeks, and with actigraphy (Actiwatch-L, Mini-mitter, Inc., Bend, OR) for at least 7 days prior to entering the unit. Participants were asked to refrain from use of any prescription or non-prescription medications, supplements, recreational drugs, caffeine, alcohol, or nicotine. Compliance with these instructions was verified by urine and blood toxicology during screening, and urine toxicology upon entry to the Unit.

Study Protocol

Participants were studied for 7 days in an environment free of time cues (no access to windows, clocks, watches, live TV, radio, internet, telephones, and newspapers, and continually supervised by staff trained not to reveal information about the time of day). The schedule consisted of a 3-day baseline (8:16 h sleep:wake cycle based on average sleep times in the 7 days prior to study entry), a 40 h constant routine (CR), a 16 h experimental light exposure day, followed by 8 h sleep and then discharged. Only data recorded during the CR are reported here. During the CR, participants were asked to remain awake while supervised under constant dim light, in a semi-recumbent posture, with daily nutritional intake divided into hourly portions (150 mEq Na+/100 mEq K+ (± 20%) controlled nutrient, isocaloric [basal energy expenditure × 1.3] diet, 2500 mL fluids/24 h).

Sleepiness and Performance Assessments

Subjective sleepiness was rated using the Karolinska Sleepiness Scale (KSS; 9-point scale from 1-"very alert" to 9-"very sleepy, fighting sleep")(Åkerstedt and Gillberg 1990) by pressing the appropriate number on a computer keyboard when prompted. Performance was assessed using an auditory 10-minute psychomotor vigilance task (aPVT), in which an auditory signal was presented at random intervals (1-9 seconds), and the participant asked to

press a button as soon as possible after hearing the sound. No simultaneous visual stimulus was presented.

The auditory version of the PVT was used as it was more appropriate for administration during the light exposure phase of the study that followed the CR (data not included here). The aPVT is a validated measure of sustained attention and a reliable index of objective sleepiness (Jung et al. 2010; Lisper and Kjellberg 1972), and correlates with the visual PVT during sleep deprivation (Jung et al. 2010). KSS and aPVT were assessed every 60 minutes beginning 2 hours after waking.

Oculography

Participants' eyelid movements were monitored by infrared (IR) reflectance oculography to record sleepiness levels continuously throughout the CR (Optalert™, Melbourne, Australia (Johns et al. 2007)). IR transducers fitted to spectacle frames were positioned towards the top eyelid to capture three phases of the eye blink: i) the closing phase, ii) closed phase, iii) and re-opening phase. The system provides thirteen ocular measures sampled each minute, of which the following were selected based on previous work (Åkerstedt et al. 2010; Caffier et al. 2003; Ftouni et al. 2013) (detailed description of each measure provided in Table 1): positive and negative amplitude/velocity ratio of each blink (PosAVR, NegAVR); Johns Drowsiness Scale (JDS) score; the percentage of time with eyes closed (%TEC); and mean total blink durations (BD). Increasing values indicate increasing sleepiness.

Waking EEG Recordings

Polysomnographic recordings were made continuously throughout the CR using a portable, modular, battery-operated, ambulatory, digital polysomnographic recorder (Vitaport-3 digital recorder, TEMEC Instruments B.V., Kerkrade, The Netherlands). Recordings consisted of electroencephalogram (EEG), electrooculogram, and a 2-lead electrocardiogram. Electrodes were positioned according to the International 10-20 System, with linked mastoid references (Ax) used for wake recordings from the z-line, Fz-Ax, Cz-Ax, Pz-Ax, and Oz-Ax. Only data from the Cz-Ax derivation (central position on the nasion-inion midline) are presented in this report. All EEG signals were high-pass filtered (time constant: 0.33 seconds), low-pass filtered (−6 dB at 70 Hz, 24 dB/octave), and digitized (resolution: 12-bit, sampling rate: 256 Hz, storage rate: 128 Hz). The raw signals were stored on a Flash RAM Card (SanDisk, Sunnyvale, CA) and downloaded off-line. Electrode impedances were checked using a GRASS F-EZM4 impedance meter (Grass-Telefactor, Astro-Med, Inc., West Warwick, RI) every 8 h throughout the CR. Electrode impedances were documented and electrode applications were repeated until the impedances were all < 10 kΩ. Participants were asked to complete the Karolinska Drowsiness Test (KDT) hourly, following the KSS and aPVT test battery, during which participants were instructed to relax and fixate on a 5 cm black dot attached to a computer screen 1 m away for 3 minutes with their eyes open. EEG data in the report was sampled during the 3 minute KDT.

Data Analysis

Mean aPVT reaction time (RT) and number of lapses (responses >500ms) over the 10 minute test were calculated for each session. aPVT RT scores were transformed with a

logarithmic function, and a square root transformation was applied to aPVT lapses (Dinges and Kribbs 1991). Waking EEG signals derived from Cz/Ax during the KDT were visually inspected, and 2-second epochs containing muscle artifact, eye blinks, and eye movements were discarded from further analysis. Artifact-free 2-second epochs were subjected to off-line spectral analysis using a fast-Fourier transformation and a 10% cosine window. Data were reduced by discarding spectra above 20 Hz. Absolute power densities were log-transformed, and to minimize the effects of inter-individual variation, log-transformed power density at each sampling point was expressed as a percentage of the log-transformed power density of the entire CR for each participant individually before comparing between individuals. EEG power densities were grouped into the following frequency bands: delta-theta (0.5 – 5.5 Hz), theta-alpha (5.0 – 9.0 Hz), high-alpha (9.5 – 12.0 Hz) and beta (13.0 – 20.0 Hz) (Cajochen et al. 2002; Lockley et al. 2006).

Ten minute recordings of each ocular measure were extracted during the 10-minute aPVT and used for cross-correlation and Receiver Operating Characteristic (ROC) Area Under the Curve (AUC) analyses. Data were aligned based on recorded aPVT start and end times. Ocular measures recorded during aPVT were selected for analysis to reduce variance across participants as participants were instructed to remain still with eyes directed at the screen ahead and concentrate on the task. Ocular measures recorded during the KDT could not be used in the analysis as participants were instructed to blink as little as possible during the 3-minute task. As ocular measures are derived from eye blinks, ocular measures recorded during this task may be inaccurate representations of the participants' sleepiness level.

Data were averaged within the individual for each session and then averaged across the group. Group averages were calculated by aligning individuals with respect to their habitual wake time for analysis of homeostatic effects, and by the core body temperature (CBT) minimum for circadian effects.

Linear mixed model (SAS 4.3; SAS Institute Inc., Cary, NC, USA) analyses were used to assess the elapsed time-dependent changes of each variable in 2 hour bins, as this approach accommodates missing data. Participant was modeled as a random effect, and time awake modeled as a fixed effect, using the Kenward-Rogers method to calculate degrees of freedom. An auto-regressive co-variance structure was used for repeated measures (Littell et al. 2000). Post-hoc comparisons for any main effect of time awake were made using pairwise comparisons with Bonferroni-Holm corrections for multiple comparisons. Each 2 hour bin was compared to the mean of the first 16 hours (Anderson et al. 2013; Van Dongen et al. 2003) to reduce type 1 error. To assess the circadian component of the ocular measures, a cosinor model with a fixed period of 24-h was fitted to the data (Nelson et al. 1979).

Cross-correlation analyses were performed to assess the phase correlations between ocular measures with aPVT RT and lapses, KSS, and EEG power density waveforms. Data were grouped by habitual wake time in 1 hour bins with natural logarithm transformation. Significance was set at $p < 0.05$.

ROC curve analysis was performed to evaluate and compare the performance of ocular measures (JDS, PosAVR, %TEC, NegAVR, and BD) and subjective sleepiness (KSS) at predicting aPVT lapses above three threshold increases of each subject's baseline level of lapses (first 16 hours of wakefulness): 25%, 50%, and 75% (Chua et al. 2012). Three threshold levels were selected to assess the ability of ocular measures to predict from mild to severe levels of performance impairment. KSS was not used as the predictor as an objective measure of sleep-related neurobehavioural impairment was preferred rather than a subjective measure of sleepiness. ROC curve analysis was performed using SigmaPlot 12.3 software (ROC Curves Module; Systat Software Inc., San Jose, CA). Optimal cut-off values were selected based on maximum sensitivity and specificity values. The χ^2 statistic and associated p value for pairwise AUC comparisons are presented to assess the relative performance of the ocular measures in relation to subjective sleepiness reports at correctly classifying whether the participant had performed above or below their baseline threshold on the aPVT.

Results

Of the 380 total testing sessions (38 sessions $\times$ 10 participants), 319 (83.9%) KDT sessions, and 379 (99.7%) KSS sessions were retained. A total of 353 (92.9%) aPVT sessions were recorded, and 314 (82.6%) with coincident ocular data. Ocular data loss was primarily due to missing data (glasses off), excluded system-generated interpolated data (when no signal is identified for two minutes), and blinks not recognized resulting in zero values. Intervals with poor signal quality were removed on an individual basis.

Time course of ocular measures

The time course of ocular measures for 40 hours under CR conditions is presented in Figure 1. The first 16 hours were averaged to form one value in order to examine the effect of elapsed time awake. Linear mixed model analysis revealed that ocular measures displayed significant main effects of time awake across the 40 hour CR (JDS scores: $F_{12,78.9} = 1.91$, $p = 0.046$; PosAVR: $F_{12,80.9} = 3.32$, $p = 0.001$; %TEC: $F_{12,81} = 2.09$, $p = 0.026$; NegAVR: $F_{12,80.6} = 2.39$, $p = 0.011$). No significant main effect for elapsed time awake was detected for BD. Ocular measures remained low during the first ~16 h of CR, after which all began to increase. Paired t-tests showed that ocular measures were significantly higher after 24 – 25 hours of wakefulness compared to the first 16-hour (baseline) mean (JDS: $t_{47.8} = 4.17$, $p < 0.001$; TEC: $t_{52.9} = 3.80$, $p < 0.001$; PosAVR: $t_{79.3} = 4.40$, $p < 0.001$; NegAVR: $t_{65.2} = 4.42$, $p < 0.001$; Figure 1A-E). Following this peak, ocular measures decreased until approximately 36 h of CR before scores began to increase again.

Cosinor analysis with a fixed period of 24 h was used to assess the circadian component of each ocular variable (Figure 1F-J). All ocular measures demonstrated significant cosinor regression as determined by a non-zero slope (JDS adjusted $r^2 = 0.60$, $p < 0.001$; PosAVR adjusted $r^2 = 0.71$, $p < 0.001$; %TEC adjusted $r^2 = 0.51$, $p < 0.001$; NegAVR adjusted $r^2 = 0.66$; BD adjusted $r^2 = 0.40$, $p = 0.002$).

Time course of performance, subjective sleepiness and EEG power densities

The time courses of aPVT RT and lapses, and KSS are presented in Figure 2. The first 16 hours were averaged to form one value in order to examine the effect of elapsed time awake. Linear mixed model analyses detected a significant main effects of time awake in KSS scores, and aPVT RT and lapses (KSS: $F_{12,87.8}$= 6.54, p < 0.001; mean aPVT RT: $F_{12,80.6}$= 5.56, p < .001; aPVT lapses: $F_{12,86.2}$= 5.52, p < .001). KSS scores remained low over the first 16 h awake and peaked at 26 h of CR (Figure 2A). Mean scores showed a transient decline and then remained relatively high for the remainder of the CR and did not recover to baseline levels. This difference was reflected in the paired t-tests which showed significantly higher KSS scores consistently over 17 – 40 hours of wakefulness compared to first 16 hours of wakefulness, peaking at 24 – 25 hours of wakefulness ($t_{49.1}$= 7.55, p < 0.001). Paired t-tests of aPVT RT and lapses revealed a similar pattern across the 40 h of wakefulness (Figure 2B-C), both peaking between 26 – 27 hours (aPVT RT: $t_{49.9}$= 6.56, p < 0.001; aPVT lapses: $t_{81.5}$= 6.16, p < 0.001), followed by a decrease in RT and lapses after approximately 32 hours of wakefulness.

The time course of EEG power density frequency bands are presented in Figure 3. The first 16 hours were averaged to form one value in order to examine the effect of elapsed time awake. Linear mixed model analyses detected a significant main effect of elapsed time awake in delta-theta ($F_{12,77}$ = 4.65, p < 0.001), theta-alpha ($F_{12,81}$= 4.97, p < 0.001), and beta ($F_{12,89.8}$ = 2.36, p = 0.011) frequency bands. No significant main effect for elapsed time awake was detected for the high-alpha frequency band. Delta-theta and theta-alpha frequency bands remained low over the initial 16 hours of CR, and peaked after 26 – 27 hours of wake (delta-theta: $t_{88.8}$= 4.96, p < 0.001; theta-alpha: $t_{86.1}$= 5.04, p < 0.001), then did not decrease to baseline levels over the remaining hours. The beta frequency band demonstrated a similar pattern, with beta peaking after 24 – 25 hours of wake (t_{105}= 4.14, p < 0.001).

Correlations between ocular measures and performance, sleepiness, and EEG power densities

The phase relationships between subjective sleepiness, aPVT RT, aPVT lapses, and EEG power densities versus ocular measures were quantified by cross-correlation analyses. The peaks in KSS scores, and aPVT RT and lapses versus ocular measures demonstrated significant maximal correlations with all ocular measures at zero to −1 h lag (r range = 0.66 – 0.88; See Supplementary Online Material [SOM] Figure S1). Significant maximal correlation coefficients for EEG power density in the delta-theta range (0.5 – 5.5 Hz) and theta-alpha range (5.0 – 9.0 Hz) with all ocular measures were observed at lags of −1 h (r range = 0.66 – 0.81; See SOM Figure S2), suggesting that ocular measures of sleepiness lag EEG power densities. While maximal correlation coefficients for EEG power density in the beta range (13.0 – 20.0 Hz) and ocular measures were observed at −1 to 1 h lag. No significant correlation coefficients were detected at any time lag for any of the ocular measures with EEG power density in the high-alpha range (9.5 – 12.0 Hz).

Predictability of ocular measures

To assess the performance of ocular measures and subjective sleepiness in predicting increases in aPVT lapses in performance, ROC curves for JDS, PosAVR, %TEC, NegAVR, and BD, and KSS were constructed (See SOM Table S1). aPVT lapses were dichotomized based on individual participants' increase in lapses above thresholds defined as >25%, >50%, and >75% during the 40 hours of wakefulness, relative to the first 16 hours of wakefulness.

For predicting a 25% increase in aPVT lapses, ocular measures were not significantly different from KSS. Mean PosAVR yielded the greatest AUC (0.80 ± 0.03), with a cut off of 1.70 correctly classifying a lapse in attention with 80% sensitivity and 76.5% specificity. KSS yielded similar results with correct classification sensitivity of 80.9% and specificity of 63.7% at reported KSS scores of 6.5 and above.

Similar results were produced for the 50% and 75% threshold, with PosAVR, JDS, KSS, and BD the strongest predictors. Mean PosAVR was a significantly stronger predictor of aPVT lapses compared to KSS at both 50% and 75% thresholds ($p = .02$). Mean PosAVR correctly classified lapses above the 50% and 75% thresholds with more than 80% sensitivity and 74% specificity. JDS scores demonstrated less precision with 71% sensitivity and 65% specificity at a cut off of 4.2. BD showed similar predictability with 71% sensitivity and specificity at a cut off of 658 msec mean blink duration.

Discussion

The study demonstrated the potential utility of ocular measures as indicators of sleepiness during extended wakefulness. From a well-rested state during 40 hours of sustained wakefulness, ocular measures - JDS, amplitude/velocity ratios of the opening and closing phase of blinks, and the percentage of time spent with eyes closed - displayed homeostatic and circadian influences, and correlated highly with other measures of sleepiness (KSS) and performance (aPVT). Ocular measures also demonstrated their ability to predict sleepiness-related attentional lapses with reasonable accuracy.

Elapsed time-dependent influences were observed in ocular measures across the 40 hours of extended wakefulness. As expected, ocular-derived sleepiness measures, particularly JDS scores, PosAVR and NegAVR, and %TEC, increased with elapsed time awake, consistent with EEG-derived measures of sleepiness. A modest improvement in alertness in JDS, PosAVR, and NegAVR was noted at approximately 2000 h following 12 hours of wakefulness, likely reflecting the wake maintenance zone (Shekleton et al. 2013). The results corroborate previous findings demonstrating significant increases in the duration of the eyelid closing, closed, and re-opening phases of a blink with time awake, in line with other well-established sleepiness indicators (Åkerstedt et al. 2010; Anund et al. 2008; Caffier et al. 2003; Häkkänen et al. 1999; Schleicher et al. 2008; Stern et al. 1996). Mean blink duration per minute was not found to differ significantly with elapsed time awake. Similar results were found by Åkerstedt et al. (2010) who observed only a weak effect of time awake and time on task of blink duration. This finding may in part be due to the

variability in blink duration with increasing sleepiness (Anund et al. 2008) and thus no significant differences are detected over time.

All ocular measures displayed strong temporal relationships with subjective sleepiness (KSS) levels and attentional performance (aPVT RT and lapses), with the exception of NegAVR and PosAVR. The latter two measures demonstrated lags of within one hour from subjective sleepiness and attentional performance. This lag suggests that measures of AVR may be delayed in the detection of sleepiness compared to other measures, and requires further examination with more frequent sampling to determine the true extent of any lag.

Strong cross-correlations between auditory attentional performance and ocular measures provide a unique examination of the ocular-sleepiness-performance paradigm. The use of a visual task to assess the utility of oculometrics as measures of sleepiness and associated performance impairment can be confounded by the occlusion of vision; whereby, increases in the frequency and duration of eye closures will increase the number of task-related errors due to undetected stimuli (Anderson, Wales, & Horne, 2010). Concurrently, greater eye closures will increase oculometric sleepiness ratings. By contrast herein, the use of an auditory task to assess performance circumvents this confounding factor of visual occlusion. The strong cross-correlations between oculometrics and an auditory performance task further substantiate the distinct link between ocular measures and attentional performance.

Overall, oculometrics demonstrated comparable increases in subjective sleepiness and attentional performance deterioration with time awake (Cajochen et al. 1999; Dijk et al. 1992). Participants who are sleep-deprived under controlled laboratory conditions have repeatedly demonstrated awareness of their physiological level of sleepiness, particularly when allowed the opportunity to sit quietly prior to providing their assessments (Åkerstedt et al. 2008; Horne and Burley 2010; Yang et al. 2004). This relationship has not been found reliably outside the laboratory, however (Anund et al. 2009; Anund et al. 2008; Horne and Baulk 2004). Furthermore, self-assessment of sleepiness in situations where people are often chronically and variably sleep restricted rather than totally sleep deprived, does not reliably reflect neurobehavioral performance impairment (Van Dongen et al. 2003). Consequently, this incongruence limits the use of self-ratings of sleepiness in occupational settings, whereas objective sleepiness measures by nature do not rely upon subjective introspection. We suggest, therefore, that objective measures of sleepiness be further tested in field settings with the aim of developing systems to predict fitness to perform occupational and other safety critical tasks.

Cosinor analyses demonstrated circadian modulation in all ocular measures. JDS, PosAVR, %TEC, and NegAVR displayed strong circadian influences in addition to homeostatic patterns over time. Cross-correlations with EEG frequency bands provided indirect evidence for circadian- and wake-dependent modulation of these ocular measures. Aeschbach and colleagues (1999) demonstrated the influence of the circadian pacemaker and duration of time awake in the various EEG frequency bands across 40 hours of sustained wakefulness. EEG frequencies below 9 Hz (delta, 0.75 – 4.0 Hz; theta, 4.25 – 8.0 Hz) displayed strong circadian and wake-dependent components, while higher frequencies above 12 Hz (beta: 13.25 – 20.0 Hz) displayed only a significant wake-dependent component (Aeschbach et al.

1999). In contrast, the high-alpha frequency band (10.25 – 12.0 Hz) displayed only a strong circadian component. All ocular measures correlated significantly with delta-theta (0.50 – 5.5 Hz), theta-alpha (5.0 – 9.0 Hz), and beta (13.0 – 20.0 Hz) frequency bands, providing additional evidence to suggest that ocular measures of sleepiness reflect both circadian and wake-dependent properties. The absence of a significant correlation of any ocular measures with the high-alpha frequency band (9.5 – 12.0 Hz) may suggest that ocular measures are more strongly influenced by a wake-dependent process relative to circadian variation. Supporting this suggestion is the lack of a significant effect of the high-alpha band with elapsed time awake. A forced desynchrony protocol could be used in future studies to quantify the relative contributions of the two processes to ocular measures. Nonetheless, the current data display a basis for homeostatic and circadian rhythms in ocular measures of sleepiness across extended wakefulness.

Overall, ocular measures demonstrated reasonable accuracy in predicting performance lapses above the threshold set during the first 16 hours of wakefulness. This finding has applied implications to driving performance, which requires vigilance and psychomotor reaction (Doran et al. 2001; Jung et al. 2010; Van Dongen et al. 2003). Positive amplitude/velocity ratio (PosAVR) was the strongest predictor of performance lapses at threshold levels of 25%, 50%, and 75% in the current study, classifying lapses as well as, or better than, subjective reports of sleepiness. A number of recent studies have used the AVR of blinks as indicators of sleepiness and prediction of performance (Johns et al. 2007; Picot et al. 2010; Vadeby et al. 2010). The amplitude and velocity of a blink are known to be closely related; with increases in blink amplitude, there is also an increase in the velocity of the closing and re-opening phase of a blink (Evinger et al. 1991). Johns and colleagues demonstrated increases in the AVR of the closing and re-opening phases of a blink with sleep deprivation (2003; 2007). Specifically, however, the blink re-opening duration (Caffier et al. 2003) and re-opening AVR (Johns et al. 2007) are believed to change more reliably with sleepiness compared to the closing phase properties. The current study found that both PosAVR (closing phase) and NegAVR (re-opening phase) increased significantly from an alert state (less than 16 hours of wakefulness) into a drowsy state (wakefulness of 16 hours or more) in similar proportions. Previous studies, however, primarily found a proportionally greater difference in the duration and amplitude/velocity ratio of the re-opening phase, rather than in the closing phase (Caffier et al. 2003; Johns et al. 2007). In the current study, PosAVR was the strongest predictor of performance lapses, while NegAVR classified lapses with moderate accuracy. At all levels of impairment, PosAVR at a cut-off of 1.7 correctly detected attentional impairment with at least 80% sensitivity, and correctly identified below-threshold performance no less than 74% of the time. Conversely, NegAVR demonstrated more modest detection levels near 67% and 62% sensitivity and specificity, respectively. A study by Vadeby and authors (2010) used both the closing and re-opening AVR in the prediction of performance on a driving simulator. The study found that the use of peak opening phase AVR combined with driving performance measures was only a slightly better indicator of impairment than the peak closing AVR combined with driving performance measures. Additionally, Åkerstedt and colleagues (2010) found that the closing phase AVR was a predictor of severe subjective sleepiness (KSS   8) with 70% sensitivity and 73% specificity. PosAVR also demonstrated the strongest phase-locked relationships with

subjective sleepiness ratings (KSS), reaction time (aPVT RT) and performance lapses (aPVT lapses) above all other ocular variables. The closing phase AVR should be considered as a potentially useful indicator of sleepiness levels, in line with other indicators of sleep onset.

The Johns Drowsiness Scale (JDS) score, a composite score of sleepiness levels by a propriety algorithm of ocular indicators such as AVR and blink durations, demonstrated reasonable accuracy in detecting performance lapses. Though this combination of ocular measures did not yield as high sensitivity and specificity as PosAVR, it was not significantly different from self-reported sleepiness. The recommended cut-off level of 4.5 suggested by Optalert™ is indicated as a cautionary level of sleepiness, with an increased risk of experiencing severe lane excursions (Stephan et al. 2006). The current study demonstrated that at a JDS score cut-off of 4.2, lapses over the 50% threshold were correctly classified with 71.4% sensitivity, and 64.8% specificity. Similar results were seen at the 25% and 75% thresholds. Higher JDS scores may yield higher sensitivity (correctly classifying lapses), yet at the expense of specificity (false positives). Within a driving context, high sensitivity and specificity is essential to correctly identify a driver's risk of attentional lapses, which may affect driving performance and safety, but also to correctly identify no risk in order to maintain a driver's confidence in the accuracy and safety of a sleepiness detection system.

A number of studies have found that blink duration and the percentage of time with eyes closed to be promising measures of sleepiness level (Anund et al. 2009; Anund et al. 2008; Caffier et al. 2003; Chua et al. 2012; Häkkänen et al. 1999; Ingre et al. 2006; Wierwille and Ellsworth 1994). Chua and authors (2012) found that PERCLOS (percentage of eyelid closure over the pupil over time) was a strong predictor of PVT lapses at 25%, 50%, and 75% thresholds, and by comparison outperformed self-reported sleepiness at all threshold levels. The current study, however, did not find that %TEC (a measure of the percentage of eye closure, similar to PERCLOS) to be the strongest indictor of performance. This may be related to two important differences in methodologies: the use of visual PVT in the Chua study; and slight differences in eye closure measurement, with PERCLOS a measure of the percentage per minute that the eyes are at least 80% closed, rather than full eye closures as used here. Morris and Miller (1996) found high correlations between blink durations and long eye closure rates (eye closures greater than 500 msec per minute) with errors on a flight maneuvers task in sleep-deprived pilots. Stepwise regression analysis revealed that blink durations and long eye closure rates were able to significantly predict error on the flight task. In comparison to other available ocular measures, however, BD and %TEC were not the strongest indicators of performance in the current study, though they classified lapses with reasonable accuracy.

Although the sensitivity and specificity levels found may not reach clinical significance, they display the positive trends in a laboratory setting of the ability of ocular measures to detect attentional lapses. Further studies should build on the physiological basis of ocular measures and their potential to detect the performance decrements assessed here.

While the study demonstrated the association between ocular measures and both neurobehavioural performance and sleepiness, it is acknowledged that the sample involved

only ten young healthy males, and as such, results should be interpreted with caution in relation to women and older adults. In a laboratory setting, ocular measures displayed homeostatic and circadian changes consistent with measures used to assess performance, sleepiness, and attention across the spectrum of wakefulness. Ocular indices further demonstrated their potential in detecting lapses in performance on a sustained attention task that has relevance to driving and other safety-sensitive activities. This study provides a deeper understanding of the changes in ocular measures in parallel with widely used measures of sleepiness, alertness, and performance in the laboratory, and has wider implications with regards to the use of ocular measures in sleepiness detection systems in operational settings.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

Abe T, Nonomura T, Komada Y, Asaoka S, Sasai T, Ueno A, Inoue Y. Detecting deteriorated vigilance using percentage of eyelid closure time during behavioral maintenance of wakefulness tests. Int J Psychophysiol. 2011; 82:269–274. [PubMed: 21978525]

Aeschbach D, Matthews JR, Postolache TT, Jackson MA, Giesen HA, Wehr TA. Two circadian rhythms in the human electroencephalogram during wakefulness. American Journal of Physiology - Regulatory Integrative and Comparative Physiology. 1999; 277:R1771–R1779.

Åkerstedt T, Gillberg M. Subjective and objective sleepiness in the active individual. Int J Neurosci. 1990; 52:29–37. [PubMed: 2265922]

Åkerstedt T, Ingre M, Kecklund G, Anund A, Sandberg D, Wahde M, Philip P, Kronberg P. Reaction of sleepiness indicators to partial sleep deprivation, time of day and time on task in a driving simulator - The DROWSI project. Journal of Sleep Research. 2010; 19:298–309. [PubMed: 20050992]

Åkerstedt T, Kecklund G, Axelsson J. Effects of context on sleepiness self-ratings during repeated partial sleep deprivation. Chronobiol Int. 2008; 25:271–278. [PubMed: 18484365]

Anderson C, Chang AM, Sullivan JP, Ronda JM, Czeisler CA. Assessment of drowsiness based on ocular parameters detected by infra-red reflectance oculography. Journal of Clinical Medicine. 2013; 9:907–920.

Anderson C, Wales AWJ, Horne JA. PVT lapses differ according to eyes open, closed, or looking away. Sleep. 2010; 33:197–204. [PubMed: 20175403]

Anund A, Kecklund G, Kircher A, Tapani A, Åkerstedt T. The effects of driving situation on sleepiness indicators after sleep loss: a driving simulator study. Ind Health. 2009; 47:393–401. [PubMed: 19672013]

Anund A, Kecklund G, Peters B, Forsman A, Lowden A, Åkerstedt T. Driver impairment at night and its relation to physiological sleepiness. Scand J Work Env Hea. 2008; 34:142–150.

Bergasa LM, Nuevo J, Sotelo MA, Barea R, Lopez ME. Real-time system for monitoring driver vigilance. IEEE Transactions on Intelligent Transportation Systems. 2006; 7:63–77.

Borbely AA. A two process model of sleep regulation. Hum Neurobiol. 1982; 1:195–204. [PubMed: 7185792]

Caffier PP, Erdmann U, Ullsperger P. Experimental evaluation of eye-blink parameters as a drowsiness measure. Eur J Appl Physiol. 2003; 89:319–325. [PubMed: 12736840]

Cajochen C, Khalsa SB, Wyatt JK, Czeisler CA, Dijk DJ. EEG and ocular correlates of circadian melatonin phase and human performance decrements during sleep loss. American Journal of Physiology - Regulatory Integrative and Comparative Physiology. 1999; 277:R640–649.

Cajochen C, Wyatt JK, Czeisler CA, Dijk DJ. Separation of circadian and wake duration-dependent modulation of EEG activation during wakefulness. Neuroscience. 2002; 114:1047–1060. [PubMed: 12379258]

Chua ECP, Tan WQ, Yeo SC, Lau P, Lee I, Mien IH, Puvanendran K, Gooley JJ. Heart Rate Variability Can Be Used to Estimate Sleepiness-related Decrements in Psychomotor Vigilance during Total Sleep Deprivation. Sleep. 2012; 35:325–334. [PubMed: 22379238]

De Gennaro L, Devoto A, Lucidi F, Violani C. Oculomotor changes are associated to daytime sleepiness in the multiple sleep latency test. Journal of Sleep Research. 2005; 14:107–112. [PubMed: 15910508]

Dijk DJ, Duffy JF, Czeisler CA. Circadian and sleep/wake dependent aspects of subjective alertness and cognitive performance. Journal of Sleep Research. 1992; 1:112–117. [PubMed: 10607036]

Dinges, DF.; Grace, R. PERCLOS: A valid psychophysiological measure of alertness as assessed by psychomotor vigilance. Office of Motor Carriers, Federal Highway Administration; Washington DC: 1998. p. 1-4.

Dinges, DF.; Kribbs, NB. Performing while sleepy: effects of experimentally-induced sleepiness. In: Monk, TH., editor. Sleep, Sleepiness, and Performance. Wiley; Chichester, U.K.: 1991. p. 97-128.

Doran SM, Van Dongen HP, Dinges DF. Sustained attention performance during sleep deprivation: evidence of state instability. Archives Italiennes de Biologie. 2001; 139:253–267. [PubMed: 11330205]

Evinger C, Manning KA, Sibony PA. Eyelid movements. Mechanisms and normal data. Invest Ophth Vis Sci. 1991; 32:387–400.

Ftouni S, Sletten TL, Howard ME, Anderson C, Lenne MG, Lockley SW, Rajaratnam SWM. Objective and subjective measures of sleepiness and their associations with on-road driving events in shift workers. Journal of Sleep Research. 2013; 22:58–69. [PubMed: 22861524]

Häkkänen H, Summala H, Partinen M, Tiihonen M, Silvo J. Blink duration as an indicator of driver sleepiness in professional bus drivers. Sleep. 1999; 22:798–802. [PubMed: 10505826]

Henn V, Baloh RW, Hepp K. The sleep-wake transition in the oculomotor system. Experimental Brain Research. 1984; 54:166–176. [PubMed: 6698143]

Horne JA, Baulk SD. Awareness of sleepiness when driving. Psychophysiology. 2004; 41:161–165. [PubMed: 14693012]

Horne JA, Burley CV. We know when we are sleepy: Subjective versus objective measurements of moderate sleepiness in healthy adults. Biol Psychol. 2010; 83:266–268. [PubMed: 20064579]

Ingre M, Åkerstedt T, Peters B, Anund A, Kecklund G. Subjective sleepiness, simulated driving performance and blink duration: examining individual differences. J Sleep Res. 2006; 15:47–53. [PubMed: 16490002]

Ishihara, S. The series of plates designed as a test for colour-deficiency. Kanehara & Co. LTD; Tokyo, Japan: 1996.

Johns, MW. The amplitude velocity ratio of blinks: A new method for monitoring drowsiness; 17th Annual Conference, Associated Professional Sleep Societies (APSS); Chicago, USA. 2003;

Johns MW, Chapman R, Crowley K, Tucker A. A new method for assessing the risks of drowsiness while driving. Somnologie. 2008; 12:66–74.

Johns MW, Tucker A, Chapman R, Crowley K, Michael N. Monitoring eye and eyelid movements by infrared reflectance oculography to measure drowsiness in drivers. Somnologie. 2007; 11:234–242.

Jung CM, Ronda JM, Czeisler CA, Wright KP Jr. Comparison of sustained attention assessed by auditory and visual psychomotor vigilance tasks prior to and during sleep deprivation. Journal of Sleep Research. 2010; 20:348–355. [PubMed: 20819145]

Lisper H, Kjellberg A. Effects of 24-Hour Sleep Deprivation on Rate of Decrement in a 10-Minute Auditory Reaction-Time Task. Journal of Experimental Psychology. 1972; 96:287–290. [PubMed: 4645958]

Littell RC, Pendergast J, Natarajan R. Modelling covariance structure in the analysis of repeated measures data. Statistics in medicine. 2000; 19:1793–1819. [PubMed: 10861779]

Lockley SW, Cronin JW, Evans EE, Cade BE, Lee CJ, Landrigan CP, Rothschild JM, Katz JT, Lilly CM, Stone PH, Aeschbach D, Czeisler CA. Effect of reducing interns' weekly work hours on sleep and attentional failures. New England Journal of Medicine. 2004; 351:1829–1837. [PubMed: 15509816]

Lockley SW, Evans EE, Scheer FA, Brainard GC, Czeisler CA, Aeschbach D. Short-wavelength sensitivity for the direct effects of light on alertness, vigilance, and the waking electroencephalogram in humans. Sleep. 2006; 29:161–168. [PubMed: 16494083]

Michael N, Johns M, Owen C, Patterson J. Effects of caffeine on alertness as measured by infrared reflectance oculography. Psychopharmacology (Berl). 2008; 200:255–260. [PubMed: 18537025]

Morris TL, Miller JC. Electrooculographic and performance indices of fatigue during simulated flight. Biol Psychol. 1996; 42:343–360. [PubMed: 8652752]

Nelson W, Tong YL, Lee JK, Halberg F. Methods for cosinor-rhythmometry. Chronobiologia. 1979; 6:305–323. [PubMed: 548245]

Picot, A.; Charbonnier, S.; Caplier, A. Drowsiness detection based on visual signs: blinking analysis based on high frame rate video; International Instrumentation and Measurement Technology Conference; Austin, Texas: United States. 2010;

Picot A, Charbonnier S, Caplier A, Vu NS. Using retina modelling to characterize blinking: comparison between EOG and video analysis. Machine Vision and Applications. 2012; 23:1195–1208.

Schleicher R, Galley N, Briest S, Galley L. Blinks and saccades as indicators of fatigue in sleepiness warnings: looking tired? Ergonomics. 2008; 51:982–1010. [PubMed: 18568959]

Shekleton JA, Rajaratnam SM, Gooley JJ, Van Reen E, Czeisler CA, Lockley SW. Improved neurobehavioral performance during the wake maintenance zone. J Clin Sleep Med. 2013; 9:353–362. [PubMed: 23585751]

Stephan, K.; Hosking, S.; Regan, M.; Verdoorn, A.; Young, K.; Haworth, N. The relationship between driving performance and the Johns Drowsiness Scale as measured by the Optalert System. Monash University Accident Research Centre; Melbourne: 2006. p. 1-43.

Stern, JA.; Boyer, DJ.; Schroeder, D.; Touchstone, RM.; Stoliarov, N. Blinks, saccades, and fixation pauses during vigilance task performance: II. Gender and time-of-day. FAA Office of Aviation Medicine; Washington D.C.: 1996.

Stern JA, Walrath LC, Goldstein R. The endogenous eyeblink. Psychophysiology. 1984; 21:22–33. [PubMed: 6701241]

Torsvall L, Åkerstedt T. Sleepiness on the job: continuously measured EEG changes in train drivers. Electroencephalogr Clin Neurophysiol. 1987; 66:502–511. [PubMed: 2438115]

Vadeby A, Forsman A, Kecklund G, Åkerstedt T, Sandberg D, Anund A. Sleepiness and prediction of driver impairment in simulator studies using a Cox proportional hazard approach. Accid Anal Prev. 2010; 42:835–841. [PubMed: 20380910]

Van Dongen HP, 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; 26:117–126. [PubMed: 12683469]

Wierwille WW, Ellsworth LA. Evaluation of driver drowsiness by trained raters. Accid Anal Prev. 1994; 26:571–581. [PubMed: 7999202]

Yang CM, Lin FW, Spielman AJ. A standard procedure enhances the correlation between subjective and objective measures of sleepiness. Sleep. 2004; 27:329–332. [PubMed: 15124731]
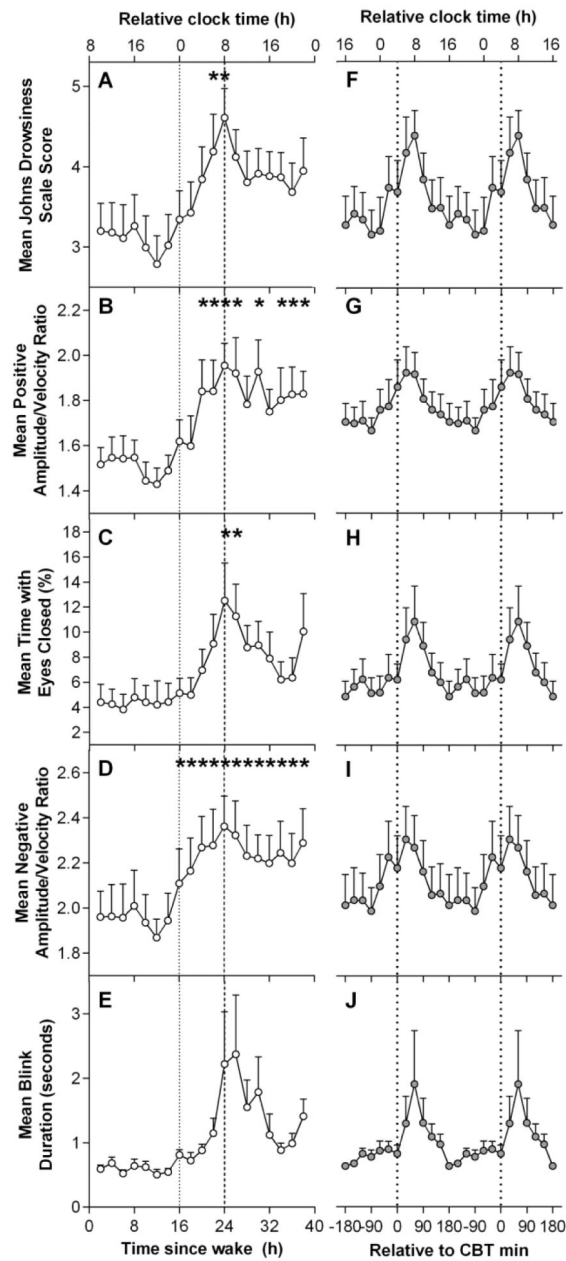
Figure 1. Time course of ocular sleepiness measures during a 40-hour constant routine
A-E: Grouped means of the time course of ocular measures. F-J: Grouped means of ocular measures aligned by core body temperature (CBT) minimum. Relative clock time was calculated based on the mean scheduled wake time. *Indicate significant difference compared to baseline 16 hour mean with Bonferroni-Holm correction ($p < 0.05$).

Figure 2. Time course of subjective sleepiness and performance during a 40-hour constant routine

(A-C) Grouped means of the time course of subjective sleepiness and performance. (D-F) Grouped means of subjective sleepiness and performance aligned by core body temperature (CBT) minimum. Relative clock time was calculated based on the mean scheduled wake time. *Indicate significant difference compared to basel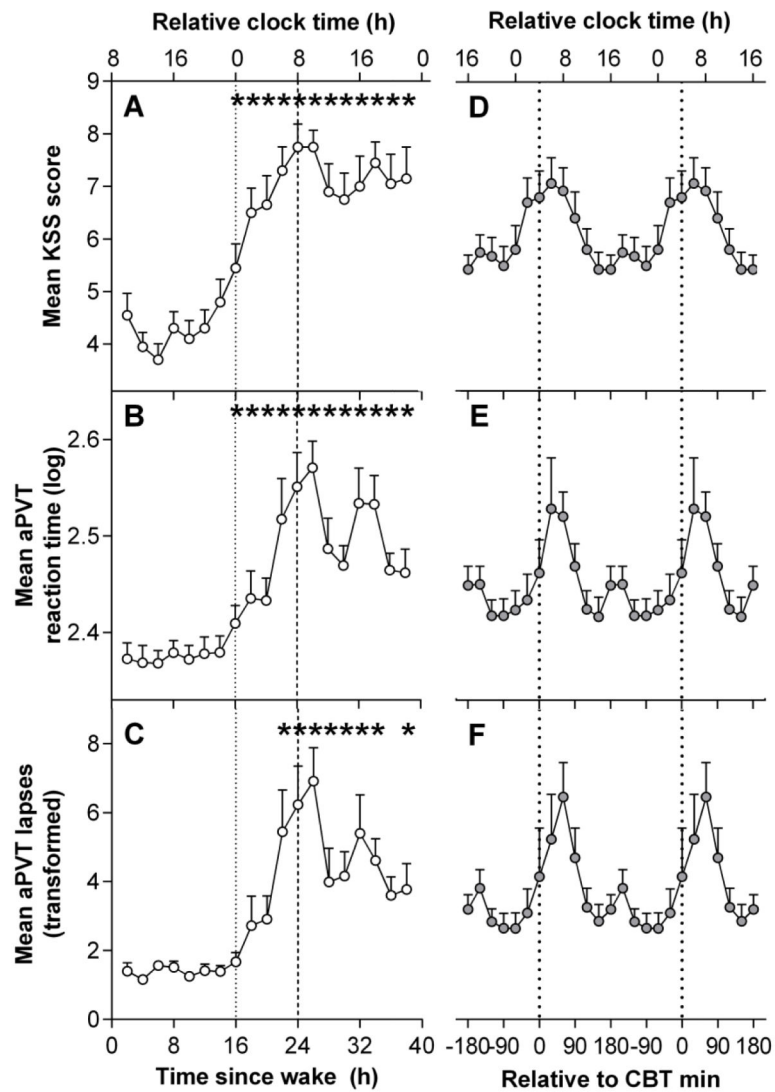ine 16 hour mean with Bonferroni-Holm correction (p < 0.05). KSS: Karolinska Sleepiness Scale; aPVT: Auditory Psychomotor Vigilance Task.
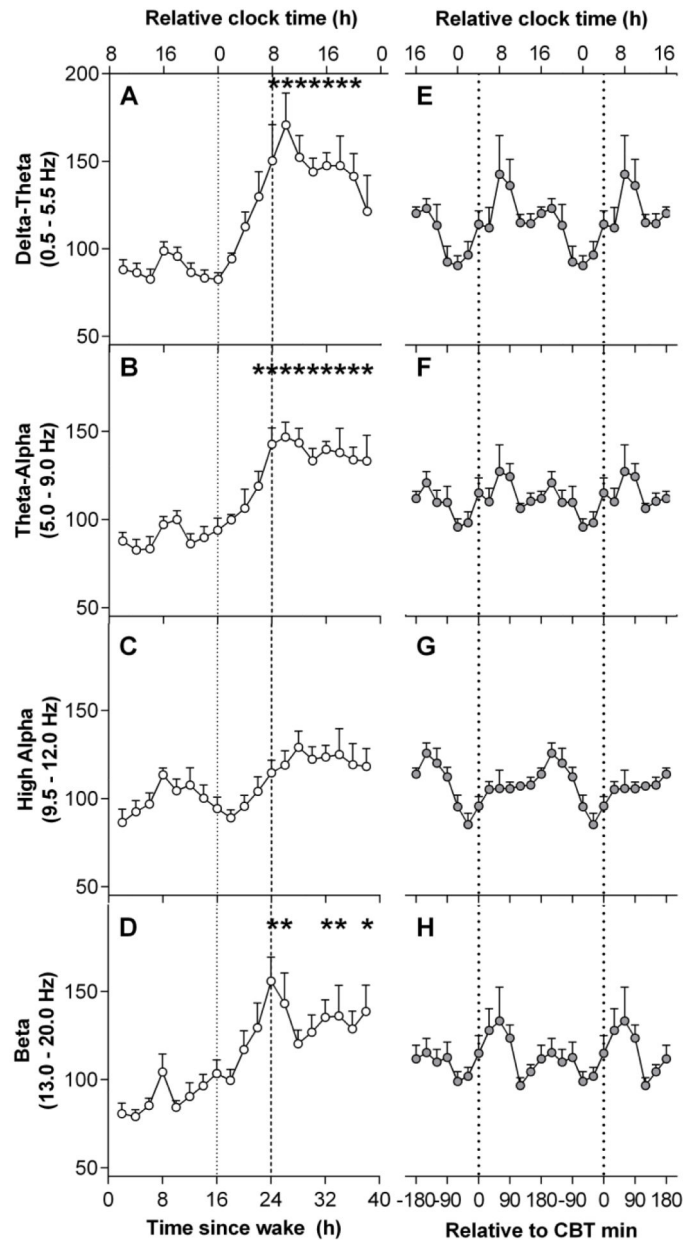
Figure 3. Time course of EEG power during a 40-hour constant routine
(A-D) Grouped means of the time course of EEG power densities. (E-H) Grouped means of
EEG power densities aligned by core body temperature (CBT) minimum. Relative clock
time was calculated based on the mean scheduled wake time. *Indicate significant difference
compared to baseline 16 hour mean with Bonferroni-Holm correction (p < 0.05).

Table 1

Ocular measures

Johns Drowsiness Scale Score (JDS)

JDS is composite score of ocular measures used to calculate an individual's level of sleepiness in real time, providing a minute-to-minute Johns Drowsiness Scale (JDS) rating (see (Johns, Tucker, Chapman, Crowley, & Michael, 2007)). JDS is a continuous scale with scores ranging from 0 to 10 (very alert to very drowsy, respectively). The commercially-available system is designed to emit auditory warnings when individuals reach a JDS score of 4.5 to 4.9 (cautionary level of sleepiness), and a score of 5.0 or above (critical level of sleepiness), which is associated with an increased risk of severe lane excursions on a driving simulator (Johns, Chapman, Crowley, & Tucker, 2008; Johns et al., 2007; Stephan et al., 2006).

Positive Amplitude/Velocity Ratio (PosAVR)

Ratio of the maximum amplitude to the maximum velocity of the closing phase of a blink; a larger ratio indicates slower closing of the eye (Johns et al., 2007).

Percentage of time with eyes closed (%TEC)

Percentage of time that the eyes are deemed closed in each minute. The eyes are deemed closed when the velocity of the eyelid movement following the closing of the eyelid drops below the velocity threshold, and is deemed closed until the velocity increases back above this threshold indicating the beginning of the re-opening of the eyelids.

Negative Amplitude/Velocity Ratio (NegAVR)

Ratio of the maximum amplitude to the maximum velocity of the re-opening phase of a blink; a larger ratio indicates slower re-opening of the eye (Johns et al., 2007)

Blink duration (BD; seconds)

The mean total blink duration of all blinks over the course of one minute. Total blink duration includes the duration of the closing, closed, and re-opening phases of each blink

---