SLEEP, 2024, 47, 1–17
https://doi.org/10.1093/sleep/zsae133
Advance access publication 15 June 2024
Original Article

Original Article

When to sleep and consume caffeine to boost alertness

Francisco G. Vital-Lopez[1,2], Tracy J. Doty[3] and Jaques Reifman[1,*]

[1]Department of Defense Biotechnology High Performance Computing Software Applications Institute, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Development Command, Fort Detrick, MD, USA,
[2]The Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc., Bethesda, MD, USA and
[3]Behavioral Biology Branch, Walter Reed Army Institute of Research, Silver Spring, MD, USA

[*]Corresponding author. Jaques Reifman, Senior Research Scientist and Director, Department of Defense Biotechnology High Performance Computing Software Applications Institute, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Development Command, ATTN: FCMR-TT, 504 Scott Street, Fort Detrick, MD 21702-5012, USA. Email: jaques.reifman.civ@health.mil.

Abstract

Study Objectives: Sleep loss can cause cognitive impairments that increase the risk of mistakes and accidents. However, existing guidelines to counteract the effects of sleep loss are generic and are not designed to address individual-specific conditions, leading to suboptimal alertness levels. Here, we developed an optimization algorithm that automatically identifies sleep schedules and caffeine-dosing strategies to minimize alertness impairment due to sleep loss for desired times of the day.

Methods: We combined our previous algorithms that separately optimize sleep or caffeine to simultaneously identify the best sleep schedules and caffeine doses that minimize alertness impairment at desired times. The optimization algorithm uses the predictions of the well-validated Unified Model of Performance to estimate the effectiveness and physiological feasibility of a large number of possible solutions and identify the best one. To assess the optimization algorithm, we used it to identify the best sleep schedules and caffeine-dosing strategies for four studies that exemplify common sleep-loss conditions and compared the predicted alertness-impairment reduction achieved by using the algorithm's recommendations against that achieved by following the U.S. Army caffeine guidelines.

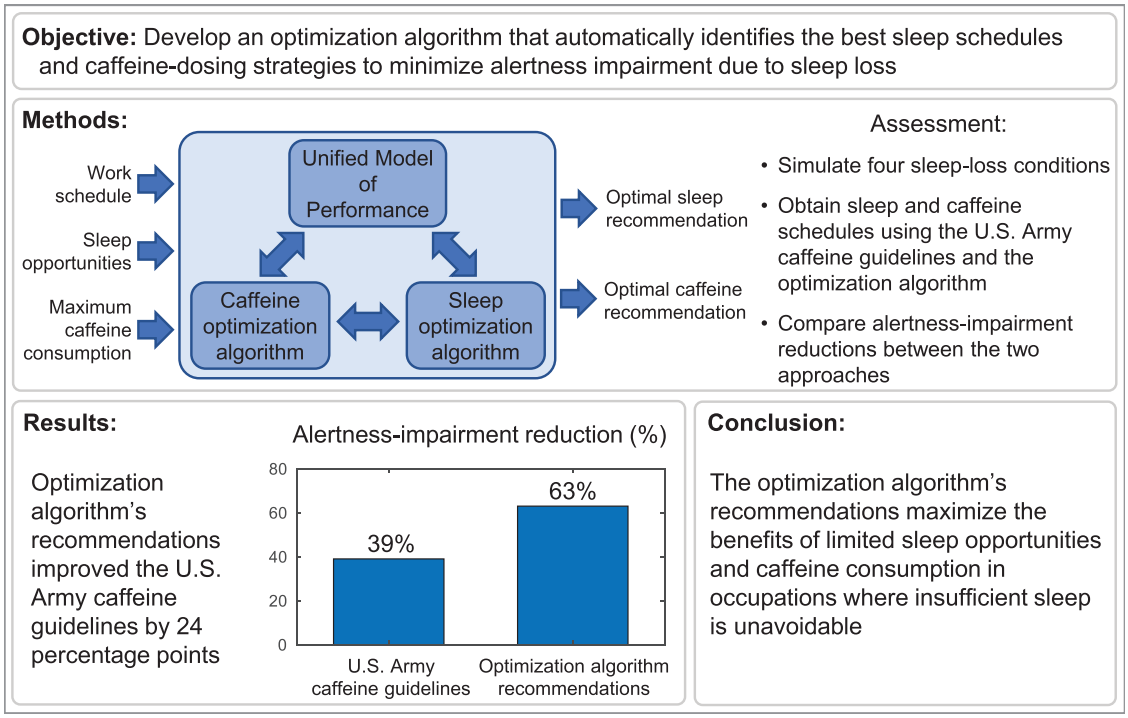
Results: Compared to the alertness-impairment levels in the original studies, the algorithm's recommendations reduced alertness impairment on average by 63%, an improvement of 24 percentage points over the U.S. Army caffeine guidelines.

Conclusions: We provide an optimization algorithm that simultaneously identifies effective and safe sleep schedules and caffeine-dosing strategies to minimize alertness impairment at user-specified times.

Key words: alertness; caffeine; fatigue management; optimization algorithm; sleep schedules

Submitted for publication: January 4, 2024; **Revised:** May 31, 2024

Graphical Abstract



Statement of Significance

In the modern 24/7 society, individuals are often required to perform their daily activities with cognitive abilities impaired by insufficient sleep, reducing productivity and compromising safety. To provide effective sleep-loss countermeasures, here, we developed and assessed an optimization algorithm to simultaneously identify the best sleep schedule and best caffeine-dosing strategy to minimize alertness impairment for a specific situation. By allowing sleep-deprived individuals to maximize the benefits of limited sleep opportunities and consume the least required amount of caffeine, this algorithm can serve as an important component of fatigue-management systems. Future efforts should focus on experimentally assessing the effectiveness of the algorithm's recommendations to mitigate alertness impairment under sleep-loss conditions.

Approximately one-third of the U.S. adult civilian population [1] and two-thirds of active-duty military personnel [2] sleep less than 6.0 hours per day, considerably less than the minimum of 7.0 to 8.0 hours that is recommended for sustaining overall health and well-being [3]. In addition to negative health consequences, insufficient sleep also impairs cognitive performance [4] and leads to fatigue levels associated with an increased risk of accidents in operational settings [5, 6]. For example, during the war in Afghanistan, about 22% of Service members who reported sleeping less than 3.0 hours per day also reported having a mishap, compared to only 1% of those who reported sleeping 7.0 hours per day [7]. One potential solution to these problems is to make accommodations to prevent insufficient sleep, for instance, by reducing workload or creating appropriate conditions to facilitate sleep, such as suitable lighting, reduced noise, and adequate ambient temperature. However, when insufficient sleep is unavoidable, individuals need effective sleep-loss countermeasures to mitigate cognitive performance impairment.

Increasing the amount and quality of sleep, targeted caffeine consumption [8], and exposure to bright light [9] are common sleep-loss countermeasures to reduce alertness impairment. Less common, however, is the availability of tools to determine, for an arbitrary situation, the best sleep schedule, when and how much caffeine to consume, or the time and intensity of light exposure required to sustain a target alertness level for a desired length of time, while considering the accumulated sleep debt, time of day, and operational constraints. The few tools available rely on computational mathematical models that predict when users will reach fatigue levels that compromise safety and, ideally, suggest possible countermeasures [10–13]. For example, using historical data, the commercially available smartphone application Jeppesen CrewAlert estimates the most likely sleep schedule for a given work itinerary and the corresponding time course of alertness for a population of crew members in the aviation industry [10, 14]. This allows crew members to anticipate whether their itinerary has an increased risk of fatigue and prepare accordingly by managing their sleep opportunities [10]. The tool also provides mitigation recommendations, such as an ideal sleep schedule, light exposure, or caffeine consumption [15], however, the suggestions are tailored to maximize alertness at a specific point in time (chosen by the user) rather than during a length of time corresponding to the entire itinerary or the most cognitively demanding periods of the work schedule. Another example is the commercially available PRISM Fatigue Management System

[12], which predicts the time course of alertness based on the three-process model of alertness regulation [16] and allows users to anticipate periods with high levels of alertness impairment, providing an opportunity to engage in suggested countermeasures. However, as with other commercial tools [17, 18], the underlying models and algorithms used to obtain alertness predictions and countermeasures are not publicly available, nor are peer-reviewed publications reporting the validity of their effectiveness.

To address the need for tools to identify effective sleep-loss countermeasures to reduce alertness impairment, we previously developed and publicly released *2B-Alert* web, an application accessible through multiple web browsers that predicts the average alertness level of a group of individuals as a function of time of day, sleep history, and caffeine consumption, and automatically recommends safe and effective caffeine interventions that lead to optimal alertness levels at user-specified times under any sleep-loss condition [11]. We also developed the smart phone application *2B-Alert* app, which, in addition to the capabilities of the web version to predict alertness and recommend caffeine interventions, can measure the alertness level of the user via the psychomotor vigilance test (PVT) and use the test measurements to learn the user's trait-like response to sleep loss to provide individualized predictions and recommendations [13]. Recently, we validated the ability of the *2B-Alert* app to identify effective individualized caffeine interventions in a 62-hour total sleep deprivation (TSD) study and showed that by following the app's recommendations, participants sustained a pre-specified alertness level 80% of the time, regardless of their sensitivity to sleep loss [13].

Here, we describe an optimization algorithm that simultaneously identifies the best sleep schedule and the best caffeine dosage (i.e. time and amount) so as to minimize alertness impairment at desired times of the day for the desired length of time. To create the optimization algorithm, we combined our previously developed algorithms to optimize caffeine consumption [19] and sleep schedules [20]. As in its original conceptualization, the resulting algorithm relies on the well-validated Unified Model of Performance (UMP) [21–23] to predict the time course of alertness for a large number of potential sleep schedules and caffeine-dosing strategies and identify the best one. Importantly, the algorithm also uses the UMP extensions that predict sleep latency and duration [24] to assess whether potential sleep schedules are physiologically feasible and avoid schedules that cannot be followed in practice. We assessed the optimization algorithm by simultaneously identifying the best sleep schedule and the best caffeine-dosing strategy to achieve a target alertness level at the desired times in four different sleep-deprivation studies that exemplify common sleep-loss conditions (i.e. restricted sleep, night-shift work, and sustained operations). To provide a benchmark, we compared the alertness-impairment reduction achieved by using the algorithm's recommendations with that achieved by following the U.S. Army caffeine guidelines for sleep-loss countermeasures [25], both for the original sleep schedules used in the studies and in combination with the sleep schedules obtained when optimizing sleep alone.

Materials and Methods

Experimental studies

We assessed the potential benefits of simultaneously optimizing caffeine consumption and sleep schedules using four diverse field and laboratory studies (*studies 1–4*), which we previously used to assess the benefits of optimizing sleep schedules alone [20]. The studies involved simulated sustained operations with a period of TSD and daytime sleep, chronic sleep restriction (CSR), CSR with night-shift work, and CSR with day-night split sleep. For each study, we used the algorithm to first optimize sleep schedules alone, and then to simultaneously optimize both caffeine consumption and sleep schedule, with the goal of quantifying and comparing the benefits of combining both countermeasures to reduce alertness impairment. In addition, we compared the reduction in alertness impairment achieved by the algorithm's recommendations against that achieved by following the U.S. Army caffeine guidelines [25].

Study 1 [26]. Ten male Special Forces personnel (mean age: 28.6 years, range: 19.0 to 32.0 years) participated in a field study of sustained operations. After an overnight 8.0-hour sleep period, starting at 07:00 on day 1, participants underwent 31.0 hours of TSD followed by 2 days of restricted daytime sleep (time in bed [TIB] from 13:30 to 17:30 on days 2 and 3). The study ended at 09:30 on day 4. Participants completed 31 sessions of a 5-minute PVT at intervals of 0.2 to 2.8 hours during wakefulness.

Study 2 [27]. Twelve participants (eight women; mean age: 26.0 years, standard deviation [SD]: 7.1 years) took part in a CSR study. As instructed, during the initial home phase of the study, participants followed their habitual sleep schedule for 14 days and monitored their sleep patterns using actigraphy and sleep diaries. Then, participants spent eight nights in the laboratory, maintaining their habitual sleep (mean duration: 7.1 hours, SD: 0.7 hours), waking up at 07:00, and leaving the laboratory during the day to perform daily activities. After this phase, participants started the full-time, in-laboratory phase of the study, consisting of one night of habitual sleep (waking up at 07:00) and seven nights of sleep restriction (TIB from 04:00 to 07:00). Participants performed a 5-minute PVT every hour between 08:00 and 18:00 during the sleep-restriction days. During PVT assessment and sleep periods, participants stayed in individual sound-attenuated rooms, where ambient temperature was maintained at 23°C and lighting was kept at 500 lux during wake periods, with background white noise kept at 65 dB at all times.

Study 3 [28]. Twelve male participants (mean age: 26.8 years, range: 18.0 to 32.0 years) took part in a study of simulated night-shift work. Prior to the in-laboratory simulated night-shift work, participants reported habitual sleep-onset times between 22:00 and 02:00 and a total sleep time ranging from 6.0 to 9.0 hours. The simulated period started at 07:00 on day 1 and ended at 23:00 on day 5. Participants had sleep opportunities with TIB from 08:00 to 12:00 on days 2 to 5 and performed 5-minute PVTs every 1.5 hours throughout most of the time awake. During PVT assessment and sleep periods, participants stayed in individual sound-attenuated rooms, where ambient temperature was maintained at 23°C and lighting was kept at 500 lux during wake periods, with background white noise kept at 65 dB at all times.

Study 4 [29]. Twelve male participants (mean age: 28.0 years, range: 21.0 to 47.0 years) took part in a CSR study with split sleep. During the week prior to the study, participants slept from 23:30 to 07:30, as verified by actigraphy and sleep diaries. Then, participants started the full-time, in-laboratory phase of the study. After three baseline nights with TIB from 23:30 to 07:30, participants started a period of 88.0 hours (starting at 07:00 on day 1 and ending at 23:00 on day 4) in which they were allowed to sleep for 2.0 hours every 12.0 hours starting at 15:00 on day 1 (seven TIB sleep opportunities in total). Participants performed a 10-minute PVT every 2.0 hours throughout most of the time awake. Throughout the in-laboratory phase, participants were isolated from time cues and ambient light was kept to less than 50 lux.

Metrics of alertness impairment

In the four studies, participants used the PVT, a well-validated and widely used reaction-time test [30, 31], to assess changes in their alertness-impairment levels caused by sleep loss. To perform a PVT, participants press a button (or tap on a touch screen) as quickly as possible immediately after they see a visual stimulus, which repeatedly appears on the screen at random time intervals between 1 and 10 seconds, over a typical 5- or 10-minute PVT session. In essence, the PVT measures the response time (RT) from the presentation of a stimulus to a participant's response to it, and from these measurements, we computed two PVT statistics: mean RT over the number of responses collected during a PVT session and PVT lapses, which indicate the number of RTs greater than 500 milliseconds. Note that the larger the values of mean RT and PVT lapses, the greater the alertness impairment.

The UMP

Based on the seminal two-process model postulated by Borbély [32], we previously developed the UMP to quantitatively predict the effect of sleep history, time of day, and caffeine consumption on alertness [21–23]. The inputs to the UMP are a sleep–wake schedule and a caffeine-consumption schedule (dose and time), and the output is the corresponding prediction of the time course of the expected alertness impairment P representative of a group of individuals:

$$P_0(t) = S(t) + \kappa C(t), \tag{1}$$

$$P(t) = P_0(t) \times g_{PD}(t, c), \tag{2}$$

$$g_{PD}(t, c) = \left[1 + M_c \frac{k_a}{k_a - k_c} \{\exp[-k_c(t - t_0)] - \exp[-k_a(t - t_0)]\}\right]^{-1} \tag{3}$$

$$M_c = M_0 c \; and \; k_c = k_0 \exp(-z\,c) \tag{4}$$

where the homeostatic process S represents the need for sleep, process C represents the sleep-independent effect of the circadian rhythm on alertness, κ denotes the circadian amplitude, t represents time (in hours), and P_0 represents the model's estimate of the mean RT statistics (or PVT lapses) in the absence of caffeine. In equation (2), g_{PD} represents the caffeine-effect factor at time t for caffeine dose c, and in equations (3) and (4), M_c and k_c denote the amplitude factor and elimination rate, respectively, for a caffeine dose c administered at time t_0, and M_0, k_0, z, and k_a denote the slope of the amplitude factor, basal elimination rate, decay constant, and absorption rate, respectively.

In this formulation, we assumed that the entire caffeine dose c is consumed at time point t_0 and absorbed and cleared by the body following a one-compartment pharmacokinetics model, and we represented the pharmacodynamics (i.e. the effect of the caffeine concentration on alertness) using the Hill equation [33]. In the absence of caffeine, $g_{PD} = 1.0$ and, thus, $P = P_0$. After caffeine consumption, g_{PD} decreases (towards a theoretical lower bound of 0.0), reaching a minimum after about 45 minutes, and then increases again towards 1.0 as the caffeine is cleared by the body, reaching 0.95 in about 7.4 hours for a 100-mg dose and 13.7 hours for a 300-mg dose. We refer the reader to Priezjev et al. [34] for a complete list of equations and parameter values of the UMP.

We have extensively validated the UMP by comparing its predictions against data collected under various sleep conditions, including CSR (3.0 to 5.0 hours of sleep per night for up to 7 days), TSD (28.0 to 88.0 hours), combinations of TSD and CSR, daytime sleep, and sleep extension. In particular, using a comprehensive set of 12 studies, including 22 sleep and caffeine conditions and a total of 301 unique participants, we recently showed that 80% of the time the UMP predictions were indistinguishable from experimental data (i.e. the predictions fell within the 95% confidence interval [CI] of the group-average data) [34]. Notably, for seven study conditions that used one or more caffeine doses ranging from 100 to 600 mg, 81% of the predictions fell within the 95% CI of the data, compared to 79% for the caffeine-free conditions, demonstrating that the UMP has a similar accuracy under caffeine and caffeine-free conditions [34].

We also extended the UMP to predict sleep latency and sleep duration as a function of sleep history and time of day [24], a necessary capability to assess the physiological feasibility of potential sleep schedules. We have validated these extensions to the UMP by comparing their predictions against data collected under various sleep conditions spanning the entire circadian cycle and different levels of sleep debt. Overall, we used data from 23 studies (309 participants, mostly healthy adults), achieving average prediction errors of 4.0 min for sleep latency and 0.8 hours for sleep duration [24].

Optimization algorithm

Previously, we separately developed two algorithms, one to identify optimal caffeine-dosing strategies [19] and the other to identify optimal sleep times [20], with the goal of mitigating alertness impairment caused by sleep loss. Here, we combined these two algorithms to obtain recommendations that simultaneously optimize caffeine consumption and sleep schedule to minimize alertness impairment during user-specified work periods, while reducing impairment during non-work periods (i.e. the time an individual is awake but not working) to the greatest extent possible.

In this formulation, the algorithm finds the time (t_c) and amount (D_c) of each caffeine dose c (with $c = 1, 2, \ldots, C$, where C denotes the total number of doses) and the start time (t_s) and duration (D_s) of each sleep period s (with $s = 1, 2, \ldots, S$, where S denotes the total number of sleep periods) that minimize alertness impairment during work periods m (with $m = 1, 2, \ldots, M$, where M denotes the total number of work periods) and non-work periods n (with $n = 1, 2, \ldots, N$, where N denotes the total number of non-work periods; Figure 1A). To achieve this goal, the algorithm iteratively identifies optimal values of t_s, D_s, t_c, and D_c through two sequential steps, as summarized below and illustrated in Figure 2. We refer the reader to Vital-Lopez et al. [19, 20] for detailed descriptions of the caffeine- and sleep-optimization algorithms.

Algorithm inputs

The algorithm requires six inputs: (1) work periods, i.e. the length of time when it is desired to sustain peak alertness (Figure 1A); (2) sleep window, i.e. the time frame within which the algorithm can define sleep periods (Figure 1A); (3) maximum sleep duration during each sleep window; (4) work alertness threshold, i.e. the desired maximum alertness-impairment level during work periods (Figure 1B); (5) non-work alertness threshold, i.e. the desired maximum alertness-impairment level during non-work periods (Figure 1C); and (6) maximum amount of caffeine to be consumed in a 24.0-hour period.

For each of the four studies, we set the work alertness threshold to 274 milliseconds for PVT mean RT or four lapses for PVT lapses, which corresponds to the highest alertness impairment during well-rested conditions. We obtained these values by using the UMP to predict the time course of alertness for a schedule of 8.0 hours

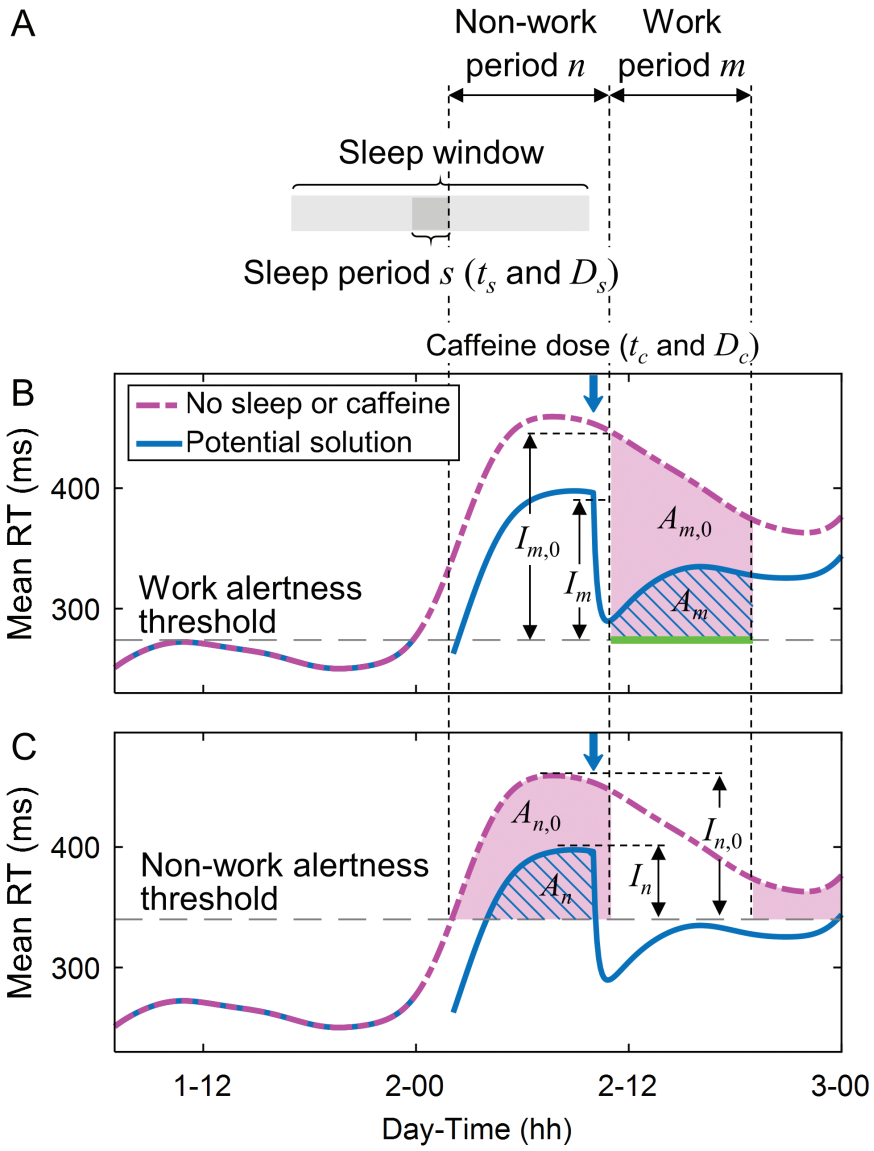


Figure 1. Definition of nomenclature and parameters used in the optimization algorithm. (A) A sleep schedule is comprised of sleep periods, where each sleep period *s* (dark shaded rectangle) is defined by its start time t_s and duration D_s. A sleep window (light shaded rectangle) defines the timespan for allocating sleep periods. Wakefulness is composed of work periods and non-work periods, where a non-work period is defined as the time when the individual is awake but not working and is indirectly inferred by the algorithm. (B) Quantities used to define the objective function Z_1 in Table 1. The graph shows the Unified Model of Performance (UMP) predictions for the no-sleep, no-caffeine condition (dashed line) and the current solution (solid line), which is indicated by the dark shaded rectangle representing sleep (panel A) and the arrow representing a caffeine dose (panel B). A caffeine dose *c* is defined by the consumption time t_c and amount D_c. For the work period *m*, $A_{m,0}$ (for the no-sleep, no-caffeine condition) and A_m (for the current solution) denote the areas under the UMP prediction curves above the user-specified work alertness threshold (horizontal dashed line). $I_{m,0}$ (for the no-sleep, no-caffeine condition) and I_m (for the current solution) denote the difference between the peak of the UMP predictions and the work alertness threshold for work period *m*. (C) Quantities used to define the objective function Z_2 in Table 1. The graph shows the two UMP predictions as in panel B. For the non-work period *n*, $A_{n,0}$ (for the no-sleep, no-caffeine condition) and A_n (for the current solution) denote the areas under the UMP prediction curves above the user-specified non-work alertness threshold (horizontal dashed line). Likewise, $I_{n,0}$ (for the no-sleep, no-caffeine condition) and I_n (for the current solution) denote the differences between the peak of the UMP predictions and the non-work alertness threshold for the non-work period *n*. RT: reaction time.

of sleep per night (from 23:00 to 07:00) and computed the highest alertness impairment during the wake time. Similarly, we set the non-work alertness threshold to 360 milliseconds for mean RT or seven lapses, which is equivalent to the alertness impairment caused by a blood alcohol concentration (BAC) of 0.06% [35, 36]. *Studies 1–3* reported results in terms of PVT mean RT. For *study 4*, once we obtained the algorithm's recommendations, we used the UMP to predict alertness impairment in terms of PVT lapses, as this was the statistic reported in this study (see Ramakrishnan et al. [22] for the UMP parameter values for PVT lapses).

Objective functions

The algorithm uses a different objective function in each of the two steps. In step 1, the algorithm minimizes the objective function Z_1, which scores alertness impairment during work periods (Table 1, Equation (5)) and has its variables graphically illustrated in Figure 1B. Z_1 uses two terms to penalize suboptimal solutions that result in alertness-impairment levels above the work alertness threshold for each work period *m*: the area under the UMP-predicted curve A_m and the maximum alertness impairment I_m.

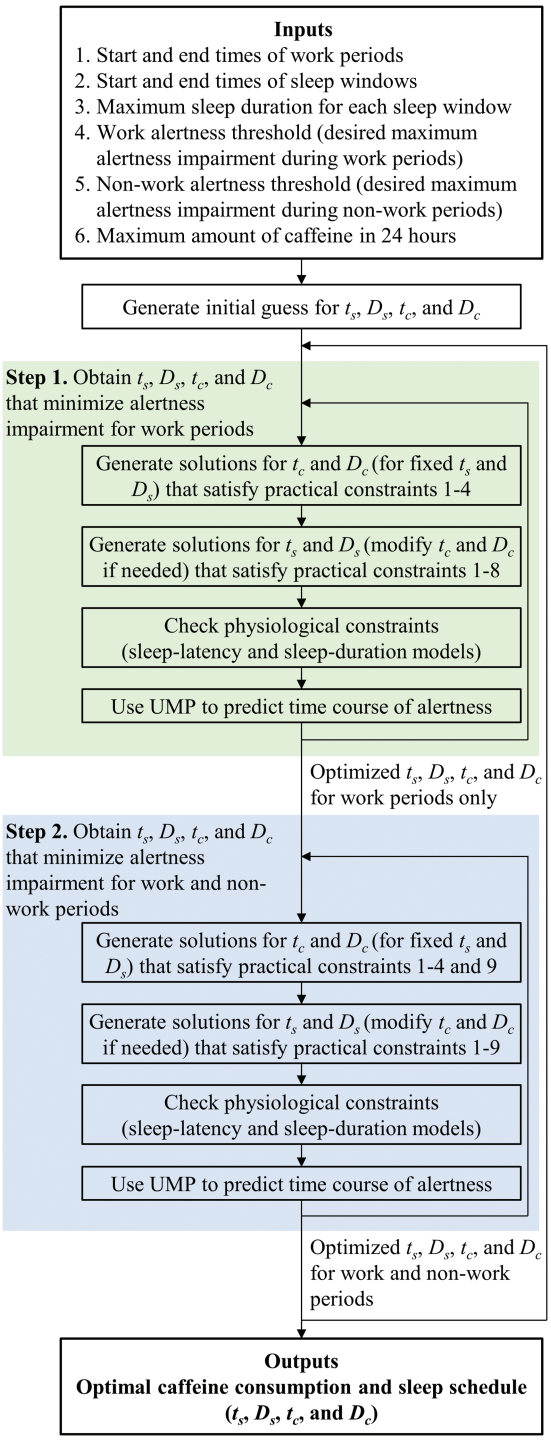


Figure 2. Summary of the two steps of the optimization algorithm used to identify caffeine consumption and sleep schedules that minimize alertness impairment during work periods, while reducing impairment to the greatest extent possible during non-work periods. Throughout the algorithm, we used the Unified Model of Performance (UMP) and its extensions to predict the time course of alertness and the physiological feasibility of each potential sleep schedule. t_s and D_s denote the start time and duration, respectively, of each sleep period s, while t_c and D_c denote the time and amount of each caffeine dose c. For a given set of six user-provided inputs, the algorithm generates initial guesses for these four parameters. In step 1, the algorithm iteratively identifies first t_c and D_c, and then t_s and D_s that minimize alertness impairment during work periods, so as to meet a user-specified work alertness threshold, while satisfying practical constraints 1–8 and physiological constraints 1 and 2 (see Methods). In step 2, the algorithm iteratively updates first t_c and D_c, and then t_s and D_s that minimize alertness impairment during both work and non-work periods, so as to meet user-specified alertness thresholds for work and non-work periods, while satisfying practical constraints 1–9 and physiological constraints 1 and 2 (see Methods). The algorithm iteratively repeats steps 1 and 2 by adding and re-allocating sleep and caffeine until the algorithm identifies a solution that achieves the desired alertness-impairment level or until the user-input maximum amount of caffeine is reached, whichever comes first.

We computed A_W as the sum of A_m ($m = 1, 2, \ldots, M$) and I_W as the largest I_m ($m = 1, 2, \ldots, M$) across all M work periods (Table 1, Equations (7) and (8), respectively), where we normalized A_W and I_W by their corresponding values $A_{W,0}$ and $I_{W,0}$ for the no-sleep, no-caffeine case (i.e. $S = 0$ and $C = 0$). Accordingly, the values for the objective function Z_1 range from zero (when the predicted alertness impairment is below the work alertness threshold for each of the M work periods) to 1.0 (when the predicted alertness impairment is equal to the one achieved with no sleep and no caffeine).

Similarly, in step 2, the algorithm minimizes the objective function Z_2, which balances alertness impairment during non-work periods and work periods (Table 1, Equation (6)). Figures 1, B and C graphically illustrate the variables used to compute the objective function Z_2. In addition to the two terms used to define Z_1, Z_2 has two additional terms to penalize suboptimal solutions for non-work periods that result in alertness levels above their threshold for each non-work period n: the area under the curve A_n and the maximum alertness impairment I_n. We computed A_R as the sum of A_n ($n = 1, 2, \ldots, N$) and I_R as the largest I_n ($n = 1, 2, \ldots, N$) across all N non-work periods (Table 1, Equations (9) and (10), respectively). We normalized A_R and I_R by their corresponding values $A_{R,0}$ and $I_{R,0}$ for the no-sleep, no-caffeine case (i.e. $S = 0$ and $C = 0$). Accordingly, the values for the objective function Z_2 range from zero (when the predicted alertness is below the non-work alertness threshold for each of the N non-work periods and below the work alertness threshold for each of the M work periods) to 1.0 (when the predicted alertness is equal to the one achieved with no sleep and no caffeine).

Constraints

To obtain practical and safe solutions, we imposed the following constraints: (1) D_c was restricted to 100, 200, or 300 mg of caffeine; (2) the dosing time t_c was restricted to occur on the hour, e.g. 18:00, 22:00, and 24:00; (3) the minimum time between doses was 2.0 hours; (4) the accumulation of caffeine in the blood, which could result in unsafe consumption [37], should be less than the maximum level achieved by a single 400-mg dose [38]; (5) D_s must be a multiple of 0.5 hours; (6) t_s should start on the hour or at the half-hour mark, e.g. 23:00 or 23:30; (7) the time lapse between sleep periods should be at least 6.0 hours; (8) the time between a sleep period and a work period should be at least 2.0 hours; and (9) for step 2 only, the average alertness impairment during each work period should be no more than 30 milliseconds larger than that of the solution obtained in step 1. Previously, we estimated the within-participant variability under well-rested conditions to be 30 milliseconds, in terms of the PVT mean RT [39].

In addition, for each sleep period s, the potential solutions must satisfy two physiological constraints: (1) D_s cannot exceed the maximum sleep duration predicted by the sleep-duration model and (2) D_s should be at least 1.0 hours, if the sleep latency predicted by the sleep-latency model is greater than 15 min. We used the latter constraint to avoid scheduling a 30-minute nap at a time when an individual would spend most of the sleep period trying to fall asleep. Note that when we predict alertness

Table 1. Optimization Algorithm Objective Functions Used to Identify the Best Caffeine Doses and Best Sleep Schedules That Minimize Alertness Impairment During Work Periods, While Reducing Impairment to the Greatest Extent Possible During Non-work Periods

Objective functions
Step 1: $\min_{t_s,D_s,t_c,D_c} Z_1 = \frac{1}{2}\left(\frac{A_W}{A_{W,0}} + \frac{I_W}{I_{W,0}}\right)$ (5)
Step 2: $\min_{t_s,D_s,t_c,D_c} Z_2 = \frac{2}{3}\left(\frac{A_W}{A_{W,0}} + \frac{I_W}{I_{W,0}}\right) + \frac{1}{3}\left(\frac{A_R}{A_{R,0}} + \frac{I_R}{I_{R,0}}\right)$ (6)
where Z_1 and Z_2 denote the objective functions we wish to minimize in each of the two steps of the algorithm; t_s and D_s represent, respectively, the start time and duration (in hours) of sleep period s, with $s = 1, 2, \ldots, S$ (the total number of sleep periods); and t_c and D_c denote, respectively, the dosing time (in hours) and amount (in mg) of caffeine dose c, with $c = 1, 2, \ldots, C$ (the total number of doses). A_W, I_W, A_R, and I_R are defined as follows:
$A_W = \sum_{m=1}^{M} A_m$ (7)
$I_W = \max(I_m,\ m = 1, \ldots,\ M)$ (8)
$A_R = \sum_{n=1}^{N} A_n$ (9)
$I_R = \max(I_n,\ n = 1, \ldots,\ N)$ (10)
where A_m and A_n denote, respectively, the area under the predicted psychomotor vigilance test (PVT) mean response time (RT) curve above the alertness threshold for work period m, with $m = 1, 2, \ldots, M$ (the total number of work periods) and non-work period n, with $n = 1, 2, \ldots, N$ (the total number of non-sleep periods). I_m and I_n denote the difference between the peak of the predicted mean RT curve and the threshold for work period m and non-work period n, respectively. $A_{W,0}$ and $A_{R,0}$ denote the total area under the mean RT curve (predicted assuming a no-sleep, no-caffeine condition, i.e. assuming $S = 0$ and $C = 0$) above the alertness threshold for the M work periods and the N non-work periods, respectively. $I_{W,0}$ and $I_{R,0}$ denote the largest difference between the peak of the mean RT curve (predicted assuming a no-sleep, no-caffeine condition) and the alertness threshold across all M work periods and across all N non-work periods, respectively. For the study using PVT lapses, we obtained these parameters based on the prediction of the PVT lapse curves as opposed to the mean RT curves.

impairment corresponding to a given sleep schedule, the predicted sleep latency is not subtracted from the duration of the sleep period because the UMP was developed using TIB as an input rather than total sleep time.

Algorithm output

The outputs of the algorithm are the optimal values of t_s, D_s, t_c, and D_c that minimize alertness impairment during work periods, while reducing impairment during non-work periods to the greatest extent possible.

Optimization steps

The algorithm starts the iterations in steps 1 and 2 with an initial guess for the no-sleep, no-caffeine condition (i.e. $S = 0$ and $C = 0$). In the first iteration, the maximum amount of caffeine is set to 0 mg, and it is increased in subsequent iterations in increments of 100 mg until no more caffeine is needed or the user-specified maximum amount is reached.

In step 1 of each iteration, the algorithm aims to identify the values of t_s, D_s, t_c, and D_c that minimize alertness impairment across the M work periods, as scored by the objective function Z_1 (Table 1, Equation (5)). To obtain the solution for step 1, first, the algorithm uses the UMP to predict the effect on Z_1 of possible modifications to the current caffeine solution (t_c and D_c), while keeping the sleep solution (t_s and D_s) fixed, and selects the modification that yields the lowest Z_1. Then, with the updated caffeine solution, the algorithm uses the UMP to predict the effect (and physiological feasibility) of possible modifications to the current sleep solution, and selects the modification that yields the lowest Z_1. (If a sleep period in the modified solution overlaps with a caffeine dose, the latter is changed to accommodate the sleep solution.) The algorithm repeats these inner steps until a maximum number of iterations is reached (600 iterations in the examples used here) or until there are no more possible modifications that improve the current solution. The algorithm assesses 12 possible modifications to the caffeine solution (e.g. adding a new caffeine dose, adding more caffeine to an existing dose, or changing the time of a current dose) and eight to the sleep solution (e.g. adding a new sleep period, adding time to an existing sleep period, or moving the start time of an existing sleep period). Please refer to Vital-Lopez et al. for additional details [19, 20].

Starting with the solution obtained in step 1, in step 2, the algorithm aims to identify the values of t_s, D_s, t_c, and D_c that minimize alertness impairment across the M work periods and the N non-work periods, as scored by the objective function Z_2 (Table 1, Equation (6)). As in step 1, the algorithm uses the UMP to predict the effect on Z_2 of possible modifications to the current caffeine solution, while keeping the sleep solution fixed, and selects the modification that yields the lowest Z_2, while keeping the average alertness impairment during each work period within 30 milliseconds of the level achieved in step 1 (constraint 9). Then, with the updated caffeine solution, the algorithm uses the UMP to predict the effect (and physiological feasibility) of possible modifications to the current sleep solution, and selects the modification that yields the lowest Z_2, while satisfying constraint 9. As in step 1, the algorithm repeats these inner steps until a maximum number of iterations is reached (600 iterations in the examples used here) or until there are no more possible modifications that improve the current solution. The possible modifications to the current caffeine and sleep solutions are the same as those in step 1. The algorithm uses the solution obtained in step 2 as the initial guess for the next iteration of step 1, and this iterative procedure continues until the algorithm identifies a solution that achieves the desired alertness-impairment level or until the user-specified maximum amount of caffeine is reached, whichever comes first. In principle, we could have directly identified a solution that minimizes alertness impairment across both work and non-work periods in step 2 without first solving for step 1. However, such an approach could lead to unacceptable work-period solutions, which we avoided by first identifying the best solution for the work period alone in step 1 and then considering the solution in step 2 (i.e. using constraint 9) to be no more than 30 milliseconds larger than the work-period solution obtained in step 1.

Alertness-impairment improvement metric

To quantify the benefit achieved by using the optimization algorithm, we compared the alertness impairment for the optimal solution and that for the original study (based on the sleep schedule used in the study and no caffeine). Specifically, based on the optimal solution for each study, we computed the percentage reduction of the area under the UMP-predicted alertness-impairment curve for work and non-work periods, separately, and then averaged the two percentages. Similarly, we computed the percentage reduction of the maximum alertness impairment

across all work and non-work periods, separately, and then averaged the two percentages. Finally, to obtain the total alertness-impairment reduction, we averaged the two results. In addition, we computed the alertness-impairment reduction achieved by the U.S. Army caffeine guidelines for the original study and by the algorithm when optimizing sleep alone, with and without the U.S. Army caffeine guidelines.

Results

We used the four sleep-deprivation studies described above to assess the benefits of simultaneously optimizing sleep and caffeine to mitigate alertness impairment. For each study, we used the UMP to perform four types of simulations: (1) considering the study's sleep schedule, we added caffeine consumption according to the U.S. Army caffeine guidelines, as a benchmark to compare the optimization algorithm's sleep and caffeine recommendations; (2) we used the optimization algorithm to obtain the best sleep schedule alone, without caffeine consumption; (3) we combined the best sleep schedule obtained in the simulation (2) above with the U.S. Army caffeine guidelines; and (4) we used the optimization algorithm to simultaneously obtain the best sleep schedule as well as the best amount and time of caffeine consumption.

For the optimizations, we defined the timing of the work periods primarily by the design of the original studies. For example, for the simulated night-shift work in *study* 3, we assumed that individuals work 8.0 hours a day from 22:00 to 06:00, for each of the 5 days of the study. Similarly, we defined three 8.0-hour work periods from 00:00 to 08:00 in *study* 1 and seven 8.0-hour work periods from 16:00 to 00:00 in *study* 2. In *study* 4, we considered a special case where the entire awake time is the work period, therefore, it does not have non-work periods. In *studies* 1–3, we defined the sleep windows based on the timing of the work periods, i.e. the sleep windows started 2.0 hours after and ended 2.0 hours prior to the work periods (to consider the time that it takes an individual to prepare for and commute back and forth to work). In *study* 4, we defined the sleep window as the entire study period. In the four studies, we constrained the total sleep duration to that of the original schedule in each study and constrained the maximum caffeine consumption per day to that recommended by the U.S. Army caffeine guidelines. For each of the four simulations in each study, we computed the predicted alertness-impairment reduction with respect to the impairment level sustained in the original study (without caffeine).

Accuracy of UMP predictions

The ability of the algorithm to optimize caffeine consumption and sleep schedules rests on the UMP's ability to make accurate predictions of alertness impairment. Therefore, first we evaluated the model by comparing its alertness predictions with the experimental data of the original study condition. For *study* 1, Figure 3A shows the measured group-average PVT mean RT values (dots) and their associated two standard errors of the mean (SEM) during wakefulness (dark shaded rectangles indicate sleep), as well as the corresponding UMP prediction (solid line). In this study, 74% of the model predictions fell within two SEM of the group-average data. Similarly, the percentage of predictions that fell within two SEM of the group-average data was 82% for *study* 2 (Figure 4A), 80% for *study* 3 (Figure 5A), and 73% for *study* 4 (Figure 6A), for an average of 77% over the four studies. These results indicate that the group-average predictions were largely indistinguishable from the experimentally measured PVT data and that the UMP can be reliably used by the optimization algorithm to predict the time course of alertness impairment of potential solutions.

U.S. Army caffeine guidelines to mitigate alertness impairment

We assessed the effectiveness of the U.S. Army caffeine guidelines [25] as a countermeasure to sleep deprivation in each of the four studies. To this end, for each study, we defined the work and non-work periods and associated alertness thresholds, with a lower impairment threshold for work periods, as discussed in Methods. Within a study, we used the same work periods in all four simulations. In *study* 1, we defined three nocturnal work periods from 00:00 to 08:00 (Figure 3B, dotted lines). For this sustained-operations condition, the U.S. Army guidelines recommend a 200-mg dose of caffeine at 00:00, 04:00, and 08:00, each night, for a total of 1800 mg over the three days of the study (Figure 3B, arrows). Although consuming caffeine according to the guidelines did not completely bring the predicted alertness impairment (solid lines) to the desired levels throughout the duration of the work (dotted lines) or non-work periods (dashed lines), it did reduce alertness impairment on average by 69%, compared to the alertness impairment of the original study (Table 2). For *study* 2, which represented a restricted sleep condition, the U.S. Army caffeine guidelines of 200 mg at 08:00 and 12:00, for each of the seven days of the study for a total of 2800 mg, reduced alertness impairment by 28% (Figure 4B). For *study* 3, which represented a night-shift condition, the U.S. Army guidelines of 200 mg caffeine at the beginning of each shift at 22:00, for each of the 5 days of the study, reduced alertness impairment by 13% (Figure 5B), whereas for *study* 4, which represented another restricted sleep condition, the guidelines of 200 mg at 06:00 and 10:00, for each of the last three days of the study, reduced alertness impairment by 45% (Figure 6B). As expected, the use of the U.S. Army caffeine guidelines as a countermeasure for sleep loss consistently reduced alertness impairment by an average of 39% across the four studies (Table 2).

Optimization of sleep recommendations

We assessed the effectiveness of the optimization algorithm to find the best times to sleep so as to minimize alertness impairment during work periods, while reducing impairment during non-work periods to the greatest extent possible. For *study* 1 (Figure 3C), we defined three sleep windows, providing opportunities to sleep from 10:00 to 22:00 each day (light-shaded rectangles), during which the algorithm allocated a total of 8.0 hours of sleep across the three sleep windows, the same total sleep duration as in the original study. The optimized sleep schedule resulted in three naps of 1.5, 3.0, and 3.5 hours (dark-shaded rectangles) on days 1, 2, and 3 of the study, respectively, each ending at the end of each sleep window at 22:00. When compared to the sleep schedule of the original study (Figure 3A), the optimized sleep schedule alone reduced the overall alertness impairment (solid line) by only 9% (Table 2). Because of the constraint we imposed to simultaneously reduce alertness impairment during both work and non-work periods, allocating sleep to reduce impairment during work periods would necessarily increase impairment during the non-work periods, and vice versa. For similar reasons, the optimized sleep schedules did not reduce alertness impairment in *studies* 2 (0%) and 3 (3%; Table 2,

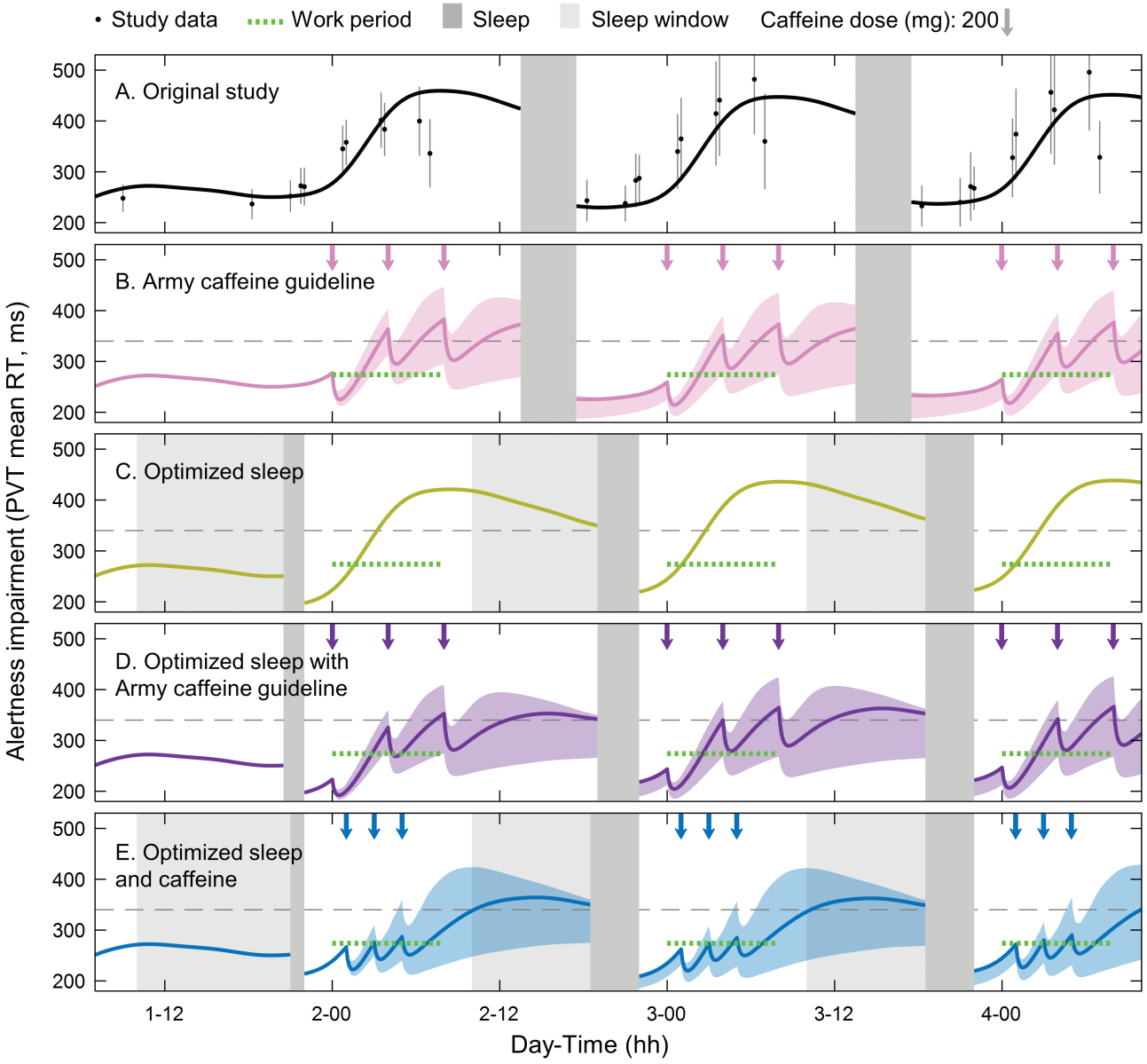


Figure 3. Experimental data and simulations for *study* 1. (A) Group-average values ($N = 10$) of the experimentally measured mean reaction time (RT) for the psychomotor vigilance test (PVT mean RT; dots), associated two standard errors of the mean (vertical lines), and Unified Model of Performance (UMP) prediction (solid line) for the original sleep schedule (dark shaded rectangles) in the study. (B) UMP prediction (solid line) for the original sleep schedule plus caffeine (three 200-mg doses at 00:00, 04:00, and 08:00 each day; arrows) based on the U.S. Army caffeine guidelines for sustained operations. (C) UMP prediction (solid line) for the optimized sleep schedule alone (three naps of 1.5, 3.0, and 3.5 hours, each ending at 22:00; dark shaded rectangles). The light shaded rectangles indicate the sleep windows within which sleep was optimized. (D) UMP prediction (solid line) for the combination of the optimized sleep schedule in panel C and the U.S. Army caffeine guidelines as in panel B (arrows). (E) UMP prediction (solid line) for the simultaneously optimized sleep schedule (three naps of 1.0, 3.5, and 3.5 hours, each ending at 22:00; dark shaded rectangles) and caffeine consumption (three 200-mg doses at 01:00, 03:00, and 05:00, each day; arrows). The horizontal dotted lines indicate both the time length and the desired alertness threshold during work periods, and the dashed lines indicate the desired alertness threshold during non-work periods (equivalent to a blood alcohol concentration of 0.06%). In all simulations, the total amount of sleep was 8.0 hours, and the total amount of caffeine in each of panels B, D, and E was 1800 mg. The shaded areas surrounding the UMP predictions (solid lines) in panels B, D, and E represent the 95% confidence intervals (CIs) of predicted alertness impairment, which we estimated based on 1000 random realizations. The lower bound of the CIs represents the response of a long caffeine half-life in the plasma and a strong caffeine effect on alertness, whereas the upper bound represents a short caffeine half-life and a weak caffeine effect on alertness.

and Figures 4C and 5C, respectively). However, in *study* 4, by consolidating sleep into three nocturnal naps of 4.5–5.0 hours at around 01:30, instead of having diurnal naps at times when the circadian rhythm boosted alertness, the optimized sleep schedule reduced alertness impairment by 33% (Figure 6C). For the four studies, the optimized sleep schedules reduced alertness impairment by an average of 11% (Table 2).

Optimization of sleep recommendations with caffeine consumption per the U.S. Army caffeine guidelines

Our results above suggested that optimizing sleep alone was minimally effective and consuming caffeine as recommended by the U.S. Army guidelines was moderately effective in reducing alertness impairment (Table 2). Thus, we examined the extent to which

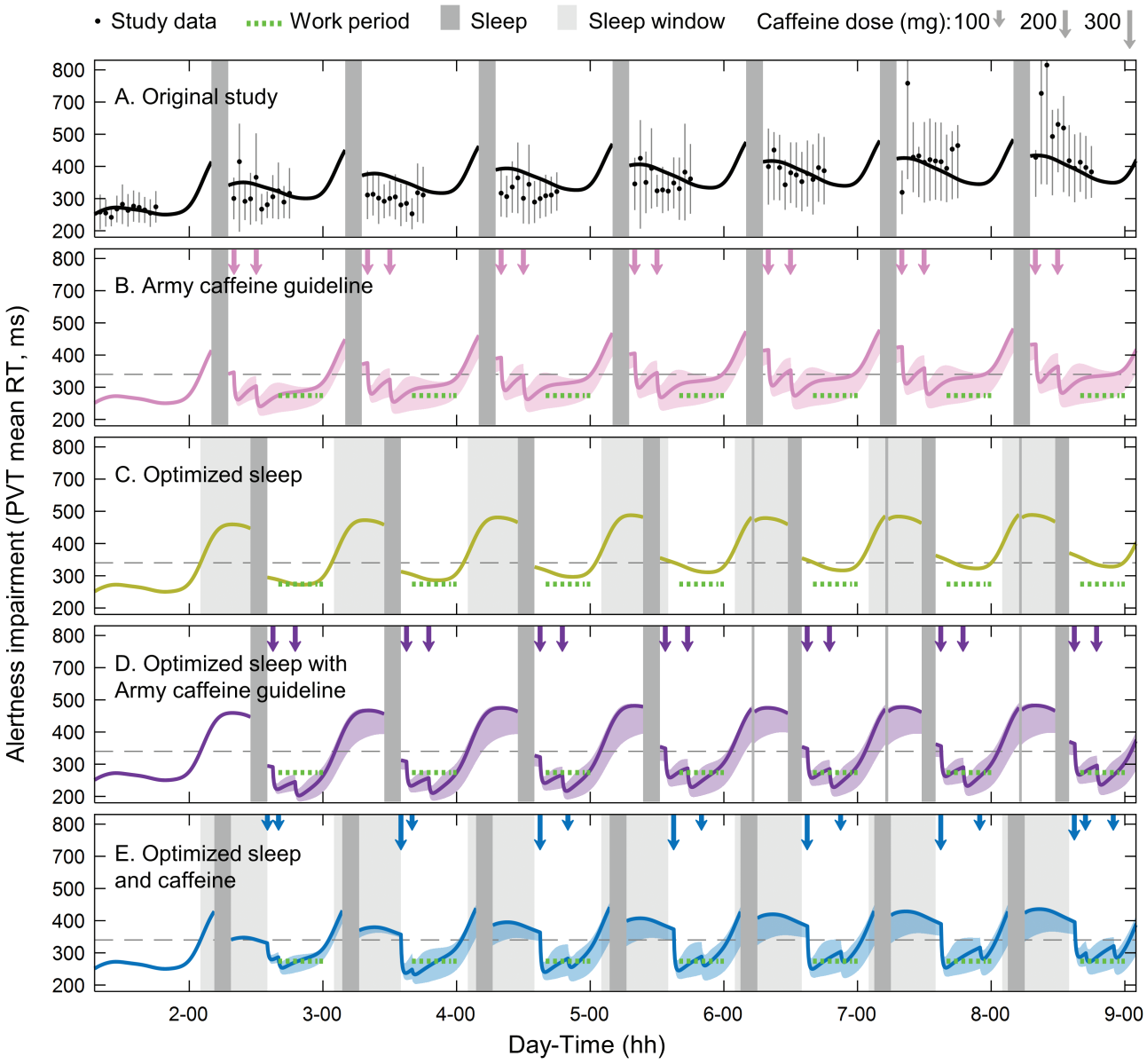


Figure 4. Experimental data and simulations for *study* 2. (A) Group-average values ($N = 12$) of the experimentally measured mean reaction time (RT) for the psychomotor vigilance test (PVT mean RT; dots), associated two standard errors of the mean (vertical lines), and Unified Model of Performance (UMP) prediction (solid line) for the original sleep schedule (dark shaded rectangles) in the study. (B) UMP prediction (solid line) for the original sleep schedule plus caffeine (two 200-mg doses at 08:00 and 12:00 each day; arrows) based on the U.S. Army caffeine guidelines for restricted sleep. (C) UMP prediction (solid line) for the optimized sleep schedule alone (a 3.0-hour nap for each of the first four days and two naps of 0.5 and 2.5 hours for each of the last 3 days). The light shaded rectangles indicate the sleep windows within which sleep was optimized. (D) UMP prediction (solid line) for the combination of the optimized sleep schedule in panel C and the U.S. Army caffeine guidelines (one 200-mg dose between 13:30 and 15:00 and one 200-mg dose between 17:30 and 19:00, each day; arrows). (E) UMP prediction (solid line) for the simultaneously optimized sleep schedule (one 3.0-hour nap each night, starting between 03:00 and 04:30) and caffeine consumption (two 100-mg doses before the start of the first work period, one 300-mg dose before the work period and one 100-mg dose during the work period for days 3 to 7, and one 200-mg dose before the work period and two 100-mg doses on the last day; arrows). The horizontal dotted lines indicate both the time length and the desired alertness threshold during work periods, and the dashed lines indicate the desired alertness threshold during non-work periods (equivalent to a blood alcohol concentration of 0.06%). In all simulations, the total amount of sleep was 21.0 hours, and the total amount of caffeine was 2800 mg in each of panels B and D and 2600 mg in panel E (because only 200 mg was recommended in the first day instead of the 400 mg prescribed by the U.S. Army guidelines). The shaded areas surrounding the UMP predictions (solid lines) in panels B, D, and E represent the 95% confidence intervals (CIs) of predicted alertness impairment, which we estimated based on 1000 random realizations. The lower bound of the CIs represents the response of a long caffeine half-life in the plasma and a strong caffeine effect on alertness, whereas the upper bound represents a short caffeine half-life and a weak caffeine effect on alertness.

combining the optimized sleep schedules with caffeine consumption as recommended by the guidelines would reduce alertness impairment. In *study* 1 (Figure 3D), the combination reduced alertness impairment by 74% relative to the impairment level sustained in the original study, an improvement of 5 percentage points over the guidelines with the original sleep schedule (Figure 3B). We observed similar improvement trends for the other studies, yielding an average 46% reduction in alertness impairment across the four studies and only a slight improvement of 7 percentage points (39% vs. 46%) compared to the U.S. Army caffeine guidelines (Table 2).

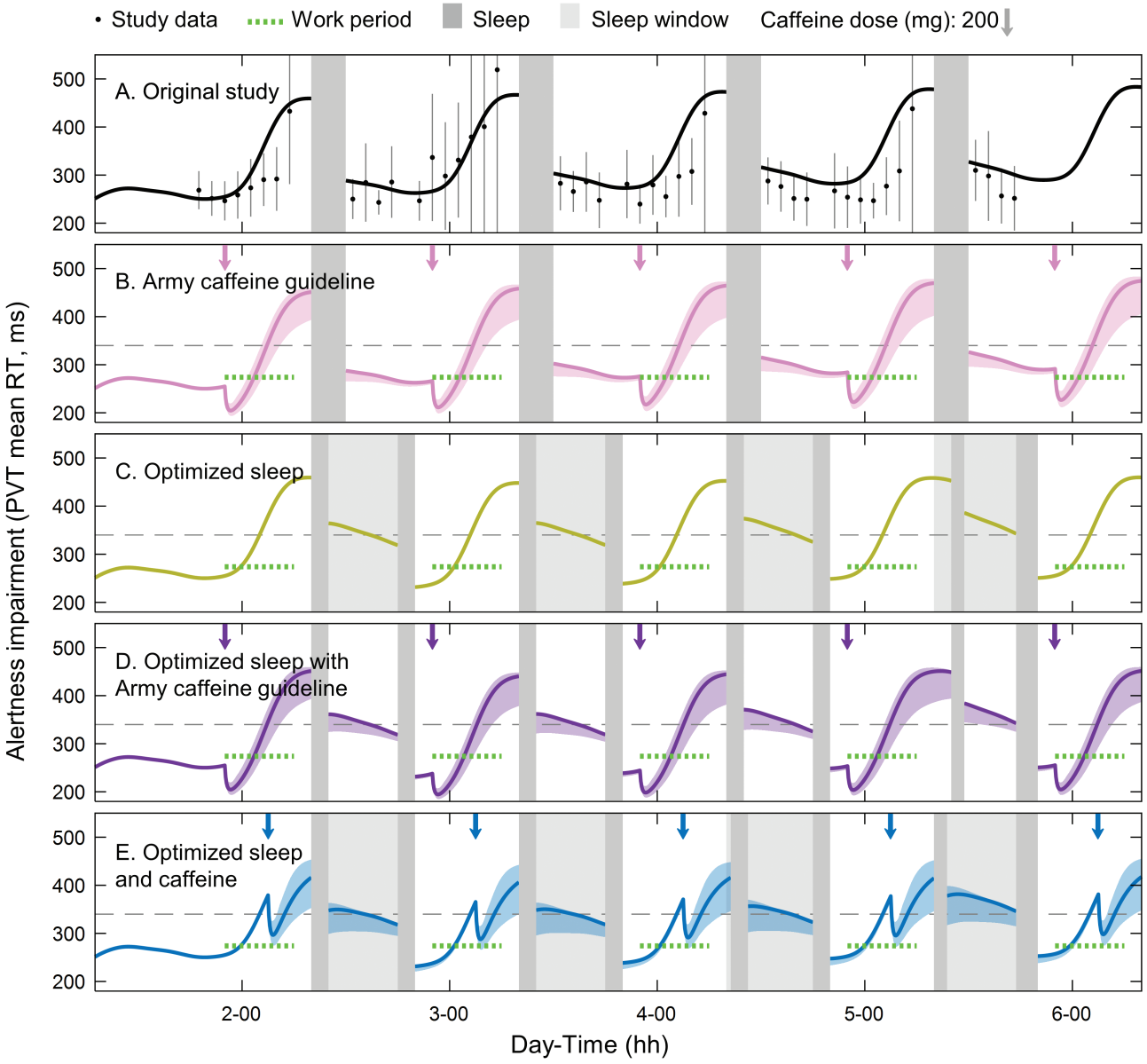


Figure 5. Experimental data and simulations for *study* 3. (A) Group-average values ($N = 12$) of the experimentally measured mean reaction time (RT) for the psychomotor vigilance test (PVT mean RT; dots), associated two standard errors of the mean (vertical lines), and Unified Model of Performance (UMP) prediction (solid line) for the original sleep schedule (dark shaded rectangles) in the study. (B) UMP prediction (solid line) for the original sleep schedule plus caffeine (one 200-mg dose at 22:00 each day; arrows) based on the U.S. Army caffeine guidelines for night shift. (C) UMP prediction (solid line) for the optimized sleep schedule alone (two naps of 2.0 hours at 08:00 and 18:00 for days 2 to 4 and two naps of 1.5 and 2.5 hours at 10:00 and 17:30, respectively, on day 5). The light shaded rectangles indicate the sleep windows within which sleep was optimized. (D) UMP prediction (solid line) for the combination of the optimized sleep schedule in panel C and the U.S. Army caffeine guidelines as in panel B (arrows). (E) UMP prediction (solid line) for the simultaneously optimized sleep schedule (two naps of 2.0 hours at 08:00 and 18:00 for days 2 and 3, two naps of 2.0 hours at 08:30 and 18:00 on day 4, and two naps of 1.5 and 2.5 hours at 08:00 and 17:30, respectively, on day 5) and caffeine consumption (one 200-mg dose at 03:00 each day; arrows). The horizontal dotted lines indicate both the time length and the desired alertness threshold during work periods, and the dashed lines indicate the desired alertness threshold during non-work periods (equivalent to a blood alcohol concentration of 0.06%). In all simulations, the total amount of sleep was 16.0 hours, and total amount of caffeine in each of panels B, D, and E was 1000 mg. The shaded areas surrounding the UMP predictions (solid lines) in panels B, D, and E represent the 95% confidence intervals (CIs) of predicted alertness impairment, which we estimated based on 1000 random realizations. The lower bound of the CIs represents the response of a long caffeine half-life in the plasma and a strong caffeine effect on alertness, whereas the upper bound represents a short caffeine half-life and a weak caffeine effect on alertness.

Simultaneous optimization of sleep schedule and caffeine consumption

Finally, we assessed the potential benefits of the optimization algorithm by simultaneously optimizing sleep schedules and caffeine consumption. In these simulations, for each study, we used the same work periods, sleep windows, and total sleep durations as in the sleep-optimization simulations described above and the same maximum amount of caffeine per 24-hour period as that prescribed by the U.S. Army caffeine guidelines. Accordingly, for *study* 1, we constrained the optimization algorithm to use 600 mg of caffeine per day. In this study (Figure 3E), the optimized sleep schedule resembled the schedule obtained when sleep was optimized alone, consisting of three naps of 1.0, 3.5, and 3.5 hours (dark shaded rectangles) on days 1, 2, and 3 of the study, respectively,

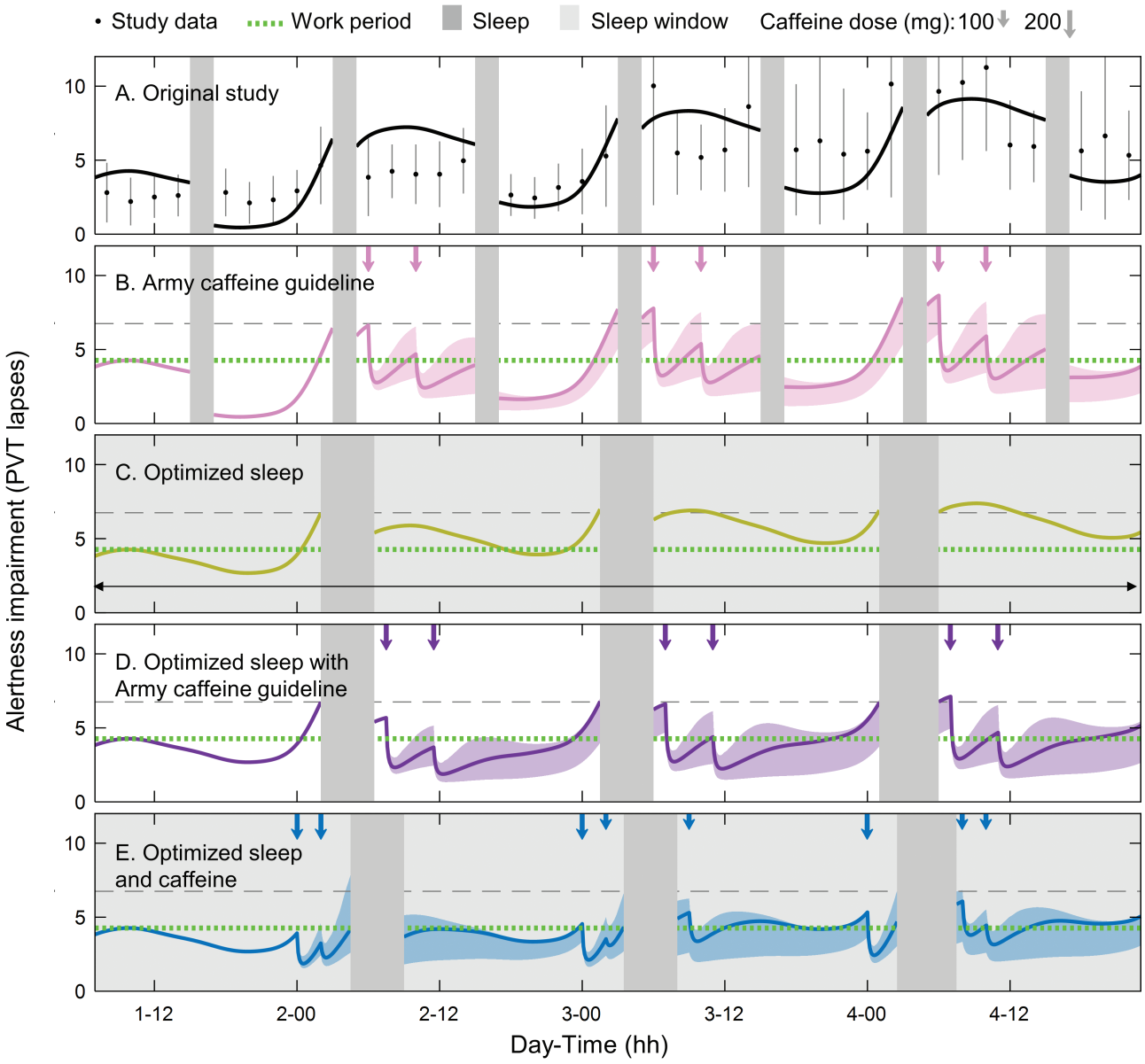


Figure 6. Experimental data and simulations for *study* 4. (A) Group-average values ($N = 12$) of the experimentally measured lapses for the psychomotor vigilance test (PVT lapses; dots), associated two standard errors of the mean (vertical lines), and Unified Model of Performance (UMP) prediction (solid line) for the original sleep schedule (dark shaded rectangles) in the study. (B) UMP prediction (solid line) for the original sleep schedule plus caffeine (two 200-mg doses at 06:00 and 10:00 each day; arrows) based on the U.S. Army caffeine guidelines for restricted sleep. (C) UMP prediction (solid line) for the optimized sleep schedule alone (three naps of 4.5, 4.5, and 5.0 hours starting at 02:00, 01:30, and 01:00, respectively, on days 2–4; dark shaded rectangles). The light shaded rectangles indicate the sleep windows within which sleep was optimized. (D) UMP prediction (solid line) for the combination of the optimized sleep schedule in panel C and the U.S. Army caffeine guidelines (two 200-mg doses at 07:30 and 11:30 on day 2 and two 200-mg doses at 07:00 and 11:00 on days 3 and 4; arrows). (E) UMP prediction (solid line) for the simultaneously optimized sleep schedule (three naps of 4.5, 4.5, and 5.0 hours starting at 04:30, 03:30, and 02:30, respectively, on days 2–4) and caffeine consumption (two 200-mg doses at 00:00 and 02:00 on day 2, one 200-mg and two 100-mg doses at 00:00, 02:00, and 09:00, respectively, on day 3, and one 200-mg and two 100-mg doses at 00:00, 08:00, and 10:00, respectively, on day 4; arrows). The horizontal dotted lines indicate both the time length and the desired alertness threshold during work periods, and the dashed lines indicate the desired alertness threshold during non-work periods (equivalent to a blood alcohol concentration of 0.06%). In all simulations, the total amount of sleep was 14.0 hours, and total amount of caffeine in each of panels B, D, and E was 1200 mg. The shaded areas surrounding the UMP predictions (solid lines) in panels B, D, and E represent the 95% confidence intervals (CIs) of predicted alertness impairment, which we estimated based on 1000 random realizations. The lower bound of the CIs represents the response of a long caffeine half-life in the plasma and a strong caffeine effect on alertness, whereas the upper bound represents a short caffeine half-life and a weak caffeine effect on alertness.

each ending at the end of each sleep window at 22:00. As in the U.S. Army caffeine guidelines, the algorithm recommended three caffeine doses of 200 mg each (arrows), but prescribed the doses 2.0 hours apart (instead of 4.0 hours), starting at 01:00 instead of 00:00. Simultaneously optimizing sleep and caffeine reduced alertness impairment by 85%, an improvement of 16 percentage points as compared to the reduction achieved following the U.S. Army caffeine guidelines with the original sleep schedule (Table 2).

For *studies* 2, 3, and 4, simultaneously optimizing sleep and caffeine reduced alertness impairment by 46%, 45%, and 78%, respectively, as compared with the impairment level in the original studies (Figures 4E, 5E, and 6E, respectively, and Table 2). Thus,

across the four studies, simultaneous optimization of sleep and caffeine reduced alertness impairment on average by 24 percentage points beyond the benefits of using the U.S. Army caffeine guidelines and 17 percentage points over the combined optimization of sleep schedule and U.S. Army caffeine guidelines (Table 2). These results show that, by tailoring sleep timing and caffeine consumption to each condition, we can enhance the benefits of each of these countermeasures, while using the same total amount of sleep and consumed caffeine.

Effectiveness of the optimal recommendations for individuals with different caffeine responses

Because there is considerable variability in how individuals respond to caffeine, due to both variations in caffeine metabolism and its effect on alertness, we sought to investigate the effectiveness of the group-average optimal caffeine recommendations for individuals with different caffeine responses. Starting with the time and amount of caffeine recommended by the group-average optimization algorithm provided above for each of the four studies, we performed 1000 separate simulations for each study by running the UMP with modified values of the parameters that determine caffeine metabolism (k_0 and z) and its effect on alertness (M_0) in Equation (4). To obtain the modified values, we randomly sampled the parameters from normal distributions with means equal to the parameters' group-average values and SDs equal to their standard errors estimated by fitting the model to experimental data collected in a sleep-loss study involving repeated doses of caffeine of different concentrations [23, 40]. Then, we estimated the 95% CI for the alertness impairment from the 1000 simulations.

The shaded areas surrounding the average alertness-impairment predictions (solid lines) in Figures 3–6 (panels B, D, and E) show the estimated 95% CIs resulting from the variability in individual responses to caffeine. The lower bound represents the response to a long caffeine half-life (6.8 hours) in the plasma and a strong caffeine effect on alertness, whereas the upper bound represents a short caffeine half-life (1.2 hours) and a weak caffeine effect on alertness. This wide range of caffeine half-life values (1.2 to 6.8 hours) is consistent with the findings of a recent systematic review of 141 caffeine studies involving over 4700 participants, which showed that most individuals metabolize caffeine within this range [41], suggesting that our simulations captured the observed between-individual variability.

As expected, the lower bound resulted in larger alertness-impairment reductions than the average response (76% to 90% vs. 39% to 63%) for each caffeine-dosing strategy in each of the four studies (Table 2). In contrast, the upper bound resulted in smaller alertness-impairment reductions than the average response (18% to 30% vs. 39% to 63%) for each of the four studies.

As in the case of the group-average prediction, for the lower bound, the optimized sleep and caffeine schedules were more effective in reducing alertness impairment than the U.S. Army caffeine guidelines applied to the original sleep schedule or the optimized sleep schedule (90% vs. 76% in both cases), with a similar pattern for the upper bound (Table 2). For the lower-bound solution in *study 1*, the estimated accumulation of caffeine in the plasma exceeded the level reached with a single 400-mg dose for an average individual by as much as 50%.

Discussion

Insufficient sleep is a common stressor in the modern 24/7 society of industrialized countries and unavoidable in some occupations. Recognizing this problem, U.S. agencies, such as the Occupational Safety and Health Administration and the Department of the Army, have issued guidelines for napping [42] and caffeine consumption [25] to mitigate the effects of insufficient sleep. However, by and large, these represent "one-size-fits-all" solutions that may not result in the full benefits of such fatigue countermeasures for every situation. Here, we developed a "clear-box" algorithm to tailor sleep time and caffeine consumption to an individual's situation so as to minimize alertness impairment at the desired times of the day for the desired duration. To this end, we combined our previous algorithms for optimizing caffeine consumption [19] and sleep schedule [20] such that the resulting algorithm can optimize sleep schedule without caffeine, optimize caffeine dosing for a fixed (i.e. given) sleep schedule, or simultaneously optimize sleep schedule and caffeine dosing.

As the algorithm leverages the UMP [21, 23, 34] to predict the time course of alertness, we first showed that the model adequately predicted the measured alertness-impairment levels of the original conditions of each of the four sleep-deprivation studies. Then, we used the UMP predictions to assess the effectiveness of the algorithm's recommendations for sleep schedule and caffeine dosing. The results over the four studies showed that, although optimizing sleep alone reduced alertness impairment during *work periods* by an average of 25%, the average reduction across both *work and non-work periods* was only 11% (Table 2), because allocating fixed sleep resources to mitigate impairment at certain times necessarily increased impairment at other times, resulting in minor overall alertness improvements.

To alleviate this situation, individuals could consume caffeine. However, it is not trivial to determine the optimal caffeine dose (time and amount) that results in the desired alertness level at the desired times of the day. Moreover, self-administration could result in a vicious cycle in which fatigued individuals consume excessive amounts of caffeine, leading to sleep disruptions followed by subsequent increases in fatigue and caffeine consumption [43]. For example, during deployment to Afghanistan, Service members who consumed large amounts of caffeine were more likely to fall asleep during briefings and on guard duty [44], and the cycle continued after deployment, with one in six Service members continuing to consume large amounts of caffeine and experience sleep problems and fatigue [45]. Certainly, as a countermeasure to sleep loss, military personnel could follow the U.S. Army caffeine guidelines, which provide recommendations for three general cases: sustained operations, restricted sleep, and night-shift work [25]. Applying these guidelines for each of the four studies reduced alertness impairment by an average of 39%, an improvement of 28 percentage points over optimizing sleep alone, which was further improved by another 7 percentage points to 46% when combined with sleep optimization (Table 2).

As expected and supported by our simulations, consuming caffeine, even at suboptimal times and amounts, such as those recommended by the U.S. Army guidelines, results in meaningful reductions in alertness impairment. Yet, when sleep schedules and caffeine doses are tailored to an individual's specific condition, alertness can be substantially improved. For example, in *study 1*, the U.S. Army guidelines recommended three 200-mg caffeine doses for each work period. However, at the time of the second dose, the alertness-impairment level was already above the work alertness threshold and the third dose was given at the end of the work period, reducing impairment during non-work periods but not reducing it enough during the most critical work periods (Figure 3B). In contrast, the optimization algorithm also prescribed three 200-mg caffeine doses for each work period, however, they were given at times when the alertness-impairment

Table 2. Alertness-Impairment Reduction Obtained in the Four Types of Simulations, Each Using a Different Strategy to Allocate Sleep and Caffeine Countermeasures

Study	U.S. Army caffeine guidelines (CI*), %	Optimized sleep without caffeine, %	Optimized sleep with caffeine based on the U.S. Army guidelines (CI), %	Optimized sleep and caffeine (CI), %
1	69 (28, 94)	9	74 (35, 99)	85 (27, 100)
2	28 (12, 81)	0	30 (10, 65)	46 (23, 80)
3	13 (2, 57)	3	15 (4, 63)	45 (16, 79)
4	45 (30, 73)	33	64 (54, 75)	78 (52, 100)
Average	39 (18, 76)	11	46 (26, 76)	63 (30, 90)

*We computed the 95% confidence interval (CI) based on 1000 random realizations for simulations that involved caffeine with a range of model parameter values. The lower bound of the CI represents the response of a long caffeine half-life in the plasma and a strong caffeine effect on alertness, whereas the upper bound represents a short half-life and a weak effect on alertness.

level had just reached the work alertness threshold, optimizing their benefit. In addition, while in the original study, no sleep was scheduled before the first work period, the optimization algorithm properly allocated a nap before each work period. Thus, the simultaneous optimization of sleep and caffeine resulted in an impairment reduction of 85% (vs. 69% for the U.S. Army caffeine guidelines with the original sleep schedule).

Overall, across the four studies, the simultaneous optimization of sleep schedules and caffeine doses exceeded the alertness-impairment reductions obtained by following the U.S. Army guidelines with the original sleep schedule by 16–33 percentage points, improvements that may have practical consequences. For example, during work periods, on average, alertness impairment following the U.S. Army guidelines was above the 0.06% BAC threshold in the four studies 21% of the time, whereas for the optimized sleep schedules and caffeine doses, the impairment level only crossed this threshold in *study* 3, corresponding to 6% of the time across the four studies.

Our results also show that, although combining optimal sleep schedules obtained by optimizing sleep alone with the U.S. Army caffeine guidelines reduced the overall alertness impairment more than the guidelines alone in every study (Table 2), in some cases it may produce undesirable consequences. *Study* 2 is a case in point. In this study (Figure 4D), the combined optimized sleep and caffeine guideline recommendations reduced alertness impairment during the *work periods* by 84% (vs. 27% for the guidelines alone in Figure 4B), however, this benefit came at the cost of increasing impairment during *non-work periods* by 24% as compared to that of the original study (vs. a 28% reduction for the guidelines alone). Importantly, this solution resulted in 37.0 hours during which the impairment level was above that caused by a BAC of 0.08%, i.e. 458 milliseconds (vs. 2.0 hours for the U.S. Army caffeine guidelines alone; Figure 4B). These results highlight the limitations of general-purpose guidelines and the importance of simultaneously optimizing sleep and caffeine, for both work- and non-work periods.

As an additional advantage, the algorithm automatically determines the minimum amount of caffeine required to sustain a desired alertness level for each specific situation. For example, for night shift, the U.S. Army guidelines recommended a single 200-mg dose at the beginning of each shift. However, in *study* 3, this dose was insufficient, reducing alertness impairment by only 13% (Table 2 and Figure 5B). In contrast, optimizing the timing of the 200-mg dose together with the sleep schedule reduced alertness impairment by 45%. However, we could have used the algorithm with a higher caffeine constraint to determine the minimum amount of caffeine required to sustain alertness impairment at the desired level. In this case, additional simulations showed that to achieve the desired alertness level, the algorithm recommended 600 mg of caffeine over three or four doses each work period, reducing the overall alertness impairment by 95%.

The algorithm uses group-average predictions to provide optimal caffeine recommendations. However, there is a large between-individual variability in the sensitivity to caffeine, driven by genetic factors as well as lifestyle decisions, such as previous use of caffeine, smoking, or the use of oral contraceptives [41]. Thus, we assessed the extent to which individual differences in caffeine metabolism resulted in discrepancies in the model's predictions. In our model, the effect of caffeine on alertness initially increases as caffeine is absorbed by the body and then decays as the caffeine is cleared. For individuals with a slow caffeine metabolism (i.e. caffeine half-life of 6.8 hours), the discrepancy in model predictions as compared to an average individual reached a maximum at around 8 hours after caffeine consumption, whereas for individuals with a fast caffeine metabolism (i.e. caffeine half-life of 1.2 hours), the discrepancy reached a maximum at about 3 hours after caffeine consumption. After they peaked, the discrepancies decreased and became negligible after reaching values comparable to the within-participant variability of 30 milliseconds in PVT mean RT [39], which occurred 16 hours after caffeine consumption for a slow caffeine metabolism and 10 hours for a fast caffeine metabolism. These model discrepancies influenced the effectiveness of the algorithm's recommendations. For example, we found that for individuals who metabolize caffeine quickly and where caffeine weakly affects their alertness (upper bounds in panels B, D, and E in Figures 3–6), the algorithm's recommendations would result in a 33-percentage point smaller alertness-impairment reduction than the reduction predicted for an average individual (30% vs. 63%, Table 2). In contrast, for individuals who metabolize caffeine slowly and where caffeine strongly affects their alertness (lower bounds in panels B, D, and E in Figures 3–6), the group-average recommendations would result in a 27-percentage point larger alertness-impairment reduction than for an average individual (90% vs. 63%, Table 2). Thus, individual variability in response to caffeine could considerably alter the effectiveness of the algorithm's recommendations based on group-average predictions.

In general, the optimization algorithm's recommendations will be no worse than those provided by the U.S. Army caffeine guidelines (Table 2) because the recommendations provided by the guidelines are part of the solution space explored by the optimization algorithm. For example, when the U.S. Army guidelines suggest optimal or near-optimal solutions, both approaches yield very similar results. This is observed for the lower bound of the 95% CI in *study* 1, which resulted in a 28% alertness-impairment reduction for the U.S. Army guidelines versus 27% for the optimization

algorithm (Table 2), and for the upper bound in *study 2* (81% vs. 80%). However, the algorithm's recommendations could be worse than the U.S. Army guidelines when the number of potential solutions is very large and we do not allow the algorithm to exhaustively explore all feasible solutions, because for practical reasons we wish to obtain an answer in less than 60 seconds. This could be the case in complex schedules spanning more than 10 work and non-work periods in as many days requiring conservative maximum alertness-impairment threshold levels during these periods. In less demanding schedules, even with time constraints, we expect that the algorithm's recommendations will be more effective than those provided by the U.S. Army caffeine guidelines.

Our work has limitations. First, the algorithm's recommendations optimize alertness impairment at the group-average level and do not account for individual differences, such as sensitivity to sleep deprivation or caffeine, which could considerably alter their effectiveness. For example, for individuals who metabolize caffeine quickly, the group-average recommendations would result in a lower amount of caffeine than necessary to achieve the desired alertness-impairment level. In contrast, for individuals who metabolize caffeine slowly, the group-average recommendations would result in a larger amount of caffeine than necessary and a high concentration of caffeine in the plasma. Overuse of caffeine may lead to undesirable side effects, such as anxiety, irritability, headache, tachycardia, nausea, and sleep disturbances [46]. This limitation can be lessened by tailoring the UMP to learn an individual's trait-like response to sleep loss and providing personalized caffeine recommendations [13, 47, 48]. In fact, in a recent prospective TSD study, we demonstrated the capability to provide personalized caffeine recommendations in real time that allowed individuals to reach a pre-specified alertness level 80% of the time, regardless of their phenotypical response to sleep loss or caffeine [13].

Second, the optimization algorithm provides sleep and caffeine recommendations that minimize alertness impairment as measured by reaction times using the PVT. However, these recommendations may not be necessarily optimal for other cognitive functions, such as executive functions. Moreover, performance on different tasks may not necessarily monotonically improve with an increase in arousal resulting from a larger consumption of caffeine because, after an initial improvement, performance can potentially deteriorate with higher levels of arousal [49]. Thus, caffeine-dosing strategies that induce high levels of arousal, by prescribing large amounts of caffeine in short periods of time, may be detrimental to cognitive performance. The optimization algorithm partially mitigates this limitation by identifying caffeine-dosing strategies that recommend the minimum amount possible of caffeine required to achieve the desired alertness target. In addition, users can limit the amount of caffeine consumption by setting (in the algorithm) the maximum amount of caffeine to be recommended per day, which would reduce or eliminate the likelihood of detrimental effects. Nevertheless, under sleep-loss conditions, which is the intended use of the algorithm, the achieved arousal levels are expected to be in a range that is beneficial to reduce alertness impairment.

Third, to obtain "optimal" solutions in a reasonable amount of time, the optimization algorithm only samples a fraction of all possible combinations of sleep and caffeine schedules, which may lead to near-optimal solutions. For example, in *study 2*, the algorithm did not sample a solution that would have generated two caffeine doses on day 8, as in the previous days of the study (Figure 4E). However, the three caffeine doses identified by the algorithm for day 8 and the two caffeine doses on the six previous days resulted in the same total amount of caffeine per day (400 mg), and the two solutions (of two or three doses on day 8) yielded reductions in alertness impairment that differed by only 0.3%.

Fourth, the underlying UMP and its extensions to predict sleep latency and sleep duration are based on data from young, healthy adults. We do not know whether the model's predictions would apply to older populations or populations with sleep disorders. Finally, only *study 2* included female participants (eight vs. four male participants). Thus, the conclusions derived from *studies 1*, 3, and 4 apply to men only. However, because we developed the underlying model used by the optimization algorithm based on more than 31 studies, of which 20 included a total of 270 female participants (vs. a total of 553 male participants), we believe that our results are likely valid for women, as well.

In summary, here we present the first computational algorithm to simultaneously optimize sleep schedule and caffeine consumption so as to maximize alertness during work and non-work hours. Complementing other fatigue-management tools, such as the publicly available *2B-Alert* Web [11], this unique capability will be helpful for occupations where insufficient sleep is unavoidable by maximizing the benefits of limited sleep opportunities and the strategic use of caffeine stimulants.

Funding

This work was sponsored by the Military Operational Medicine Program Area Directorate of the U.S. Army Medical Research and Development Command (USAMRDC), Fort Detrick, MD. The Henry M. Jackson Foundation was supported by the USAMRDC under Contract No. W81XWH20C0031.

Disclosure Statements

This was not an industry-supported study. JR and FGVL receive royalties for the licensing of the *2B-Alert* technology to Distritec. Nonfinancial disclosure: The authors indicate no other conflicts of interest. The opinions and assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the U.S. Army, the U.S. Department of Defense, or The Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc. The investigators have adhered to the policies for protection of human participants as prescribed in AR 70–25. This paper has been approved for public release with unlimited distribution.

Author Contributions

FGVL, TJD, and JR designed the study. FGVL performed the computations. FGVL and JR analyzed the results and wrote the manuscript. All authors have reviewed the manuscript and approved the submitted version.

Data Availability

All data will be made available following a written request to the corresponding author, along with a summary of the planned research.

References

1. Liu Y, Wheaton AG, Chapman DP, Cunningham TJ, Lu H, Croft JB. Prevalence of healthy sleep duration among adults--United States, 2014. *MMWR Morb Mortal Wkly Rep.* 2016;**65**(6):137–141. doi: 10.15585/mmwr.mm6506a1

2. Troxel WM, Shih RA, Pedersen ER, *et al*. Sleep in the military: promoting healthy sleep among U.S. Servicemembers. *Rand Health Q*. 2015;**5**(2):19.
3. Hirshkowitz M, Whiton K, Albert SM, *et al*. National Sleep Foundation's updated sleep duration recommendations: final report. *Sleep Health*. 2015;**1**(4):233–243. doi: 10.1016/j.sleh.2015.10.004
4. Killgore WD. Effects of sleep deprivation on cognition. *Prog Brain Res*. 2010;**185**:105–129. doi: 10.1016/B978-0-444-53702-7.00007-5
5. Swaen GM, Van Amelsvoort LG, Bultmann U, Kant IJ. Fatigue as a risk factor for being injured in an occupational accident: results from the Maastricht Cohort Study. *Occup Environ Med*. 2003;**60**(suppl 1):88–92.
6. Williamson A, Lombardi DA, Folkard S, Stutts J, Courtney TK, Connor JL. The link between fatigue and safety. *Accid Anal Prev*. 2011;**43**(2):498–515. doi: 10.1016/j.aap.2009.11.011
7. Mental Health Advisory Team 9. *Operation Enduring Freedom (OEF) 2013 Afghanistan*. Washington, DC: Office of The Surgeon General U.S. Army Medical Command; Office of the Command Surgeon Headquarters, U.S. Army Central Command; Office of the Command Surgeon U.S. Forces Afghanistan; 2013.
8. Wesensten NJ, Killgore WDS, Balkin TJ. Performance and alertness effects of caffeine, dextroamphetamine, and modafinil during sleep deprivation. *J Sleep Res*. 2005;**14**:255–266. doi: 10.1111/j.1365-2869.2005.00468.x
9. 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**(2):161–168.
10. Jeppesen. *Jeppesen CrewAlert for iOS*. 2019. https://ww2.jeppesen.com/wp-content/uploads/2019/05/CrewAlert-product-sheet.pdf, Accessed January 2, 2024.
11. Reifman J, Kumar K, Hartman L, *et al*. *2B-Alert* Web 2.0, an open-access tool for predicting alertness and optimizing the benefits of caffeine: utility study. *J Med Internet Res*. 2022;**24**(1):e29595. doi: 10.2196/29595
12. Predictive Safety. *PRISM Fatigue Management System*. https://predictivesafety.com/prism-fatigue-management-system/, Accessed January 2, 2024.
13. Vital-Lopez FG, Doty TJ, Anlap I, Killgore WDS, Reifman J. *2B-Alert* App 2.0: personalized caffeine recommendations for optimal alertness. *Sleep*. 2023;**46**(7):zsad080. doi: 10.1093/sleep/zsad080
14. Jeppesen. *The Boeing Alertness Model Technical Fact Sheet*. 2019. https://ww2.jeppesen.com/wp-content/uploads/2019/06/BAM_Tech_Fact_Sheet_2.3.pdf, Accessed January 2, 2024.
15. Jeppesen. *CrewAlert Fatigue Mitigation Functionality*. 2019. https://ww2.jeppesen.com/wp-content/uploads/2019/07/CrewAlert-Fatigue-Mitigation-Functionality.pdf, Accessed January 2, 2024.
16. Akerstedt T, Folkard S. Validation of the S and C components of the three-process model of alertness regulation. *Sleep*. 1995;**18**(1):1–6. doi: 10.1093/sleep/18.1.1
17. Branford K, Lowe A, Hayward B, Folkard S, Cabon P. *Biomathematical Fatigue Models Guidance Document*. Albert Park, Australia: Australian Government Civil Aviation Safety Authority; 2014.
18. Tucker P, Folkard S. *Working Time, Health and Safety: A Research Synthesis Paper*. Geneva, Switzerland: International Labor Office; 2012.
19. Vital-Lopez FG, Ramakrishnan S, Doty TJ, Balkin TJ, Reifman J. Caffeine dosing strategies to optimize alertness during sleep loss. *J Sleep Res*. 2018;**27**(5):e12711. doi: 10.1111/jsr.12711
20. Vital-Lopez FG, Doty TJ, Reifman J. Optimal sleep and work schedules to maximize alertness. *Sleep*. 2021;**44**(11):zsab144. doi: 10.1093/sleep/zsab144
21. Rajdev P, Thorsley D, Rajaraman S, *et al*. A unified mathematical model to quantify performance impairment for both chronic sleep restriction and total sleep deprivation. *J Theor Biol*. 2013;**331**:66–77. doi: 10.1016/j.jtbi.2013.04.013
22. Ramakrishnan S, Wesensten NJ, Balkin TJ, Reifman J. A unified model of performance: validation of its predictions across different sleep/wake schedules. *Sleep*. 2016;**39**(1):249–262. doi: 10.5665/sleep.5358
23. 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**(10):1827–1841. doi: 10.5665/sleep.6164
24. Vital-Lopez FG, Balkin TJ, Reifman J. Models for predicting sleep latency and sleep duration. *Sleep*. 2021;**44**(5);zsaa263. doi: 10.1093/sleep/zsaa263
25. United States Army Medical Department Center and School United States Army Health Readiness Center of Excellence. *A Leader's Guide to Soldier Health and Fitness. Army Techniques Publication 6-22.5*. Washington, DC: Headquarters, Department of the Army; 2015.
26. Kamimori GH, McLellan TM, Tate CM, Voss DM, Niro P, Lieberman HR. Caffeine improves reaction time, vigilance and logical reasoning during extended periods with restricted opportunities for sleep. *Psychopharmacology (Berl)*. 2015;**232**(12):2031–2042. doi: 10.1007/s00213-014-3834-5
27. Rupp TL, Wesensten NJ, Bliese PD, Balkin TJ. Banking sleep: realization of benefits during subsequent sleep restriction and recovery. *Sleep*. 2009;**32**(3):311–321. doi: 10.1093/sleep/32.3.311
28. Wesensten NJ, Reichardt RM, Balkin TJ. Ampakine (CX717) effects on performance and alertness during simulated night shift work. *Aviat Space Environ Med*. 2007;**78**(10):937–943. doi: 10.3357/asem.2055.2007
29. Doran SM, Van Dongen HP, Dinges DF. Sustained attention performance during sleep deprivation: evidence of state instability. *Arch Ital Biol*. 2001;**139**(3):253–267.
30. Dinges DF, Powell JW. Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behav Res Methods, Inst, Comp*. 1985;**17**:652–655. doi: 10.3758/bf03200977
31. Dorrian J, Rogers NL, Dinges DF. Psychomotor vigilance performance: neurocognitive assay sensitive to sleep loss. In: Kushida CA, ed. *Sleep Deprivation. Clinical Issues, Pharmacology, and Sleep Loss Effects*. New York, NY: Marcel Dekker; 2005: 39–70.
32. Borbély AA. A two process model of sleep regulation. *Hum Neurobiol*. 1982;**1**(3):195–204.
33. Ramakrishnan S, Laxminarayan S, Wesensten NJ, Kamimori GH, Balkin TJ, Reifman J. Dose-dependent model of caffeine effects on human vigilance during total sleep deprivation. *J Theor Biol*. 2014;**358**:11–24. doi: 10.1016/j.jtbi.2014.05.017
34. Priezjev NV, Vital-Lopez FG, Reifman J. Assessment of the unified model of performance: accuracy of group-average and individualised alertness predictions. *J Sleep Res*. 2023;**32**(2):e13626. doi: 10.1111/jsr.13626
35. Dawson D, Reid K. Fatigue, alcohol and performance impairment. *Nature*. 1997;**388**(6639):235. doi: 10.1038/40775
36. Williamson AM, Feyer AM, Mattick RP, Friswell R, Finlay-Brown S. Developing measures of fatigue using an alcohol comparison to validate the effects of fatigue on performance. *Accid Anal Prev*. 2001;**33**(3):313–326. doi: 10.1016/s0001-4575(00)00045-2
37. Killgore WD, Rupp TL, Grugle NL, Reichardt RM, Lipizzi EL, Balkin TJ. Effects of dextroamphetamine, caffeine and modafinil on psychomotor vigilance test performance after 44 h of

continuous wakefulness. *J Sleep Res.* 2008;**17**(3):309–321. doi: 10.1111/j.1365-2869.2008.00654.x

38. Institute of Medicine. *Caffeine for the Sustainment of Mental Task Performance: Formulations for Military Operations*. Washington, DC: The National Academies Press; 2001.
39. Reifman J, Priezjev NV, Vital-Lopez FG. Can we rely on wearable sleep-tracker devices for fatigue management? *Sleep.* 2024;**47**(3):zsad288. doi: 10.1093/sleep/zsad288
40. Kamimori GH, Johnson D, Thorne D, Belenky G. Multiple caffeine doses maintain vigilance during early morning operations. *Aviat Space Environ Med.* 2005;**76**(11):1046–1050.
41. Grzegorzewski J, Bartsch F, Koller A, Konig M. Pharmacokinetics of caffeine: a systematic analysis of reported data for application in metabolic phenotyping and liver function testing. *Front Pharmacol.* 2021;**12**:752826. doi: 10.3389/fphar.2021.752826
42. United States Occupational Safety and Health Administration. *Long work hours, extended or irregular shifts, and worker fatigue*. https://www.osha.gov/worker-fatigue. Accessed June 24, 2024.
43. Roehrs T, Roth TC. Sleep and daytime sleepiness. *Sleep Med Rev.* 2008;**12**(2):153–162.
44. Toblin RL, Clarke-Walper K, Kok BC, Sipos ML, Thomas JL. Centers for Disease Control and Prevention. Energy drink consumption and its association with sleep problems among U.S. service members on a combat deployment - Afghanistan, 2010. *MMWR Morb Mortal Wkly Rep.* 2012;**61**(44):895–898.
45. Toblin RL, Adrian AL, Hoge CW, Adler AB. Energy drink use in U.S. Service members after deployment: associations with mental health problems, aggression, and fatigue. *Mil Med.* 2018;**183**(11-12):e364–e370. doi: 10.1093/milmed/usy205
46. Nawrot P, Jordan S, Eastwood J, Rotstein J, Hugenholtz A, Feeley M. Effects of caffeine on human health. *Food Addit Contam.* 2003;**20**(1):1–30. doi: 10.1080/0265203021000007840
47. Liu J, Ramakrishnan S, Laxminarayan S, Balkin TJ, Reifman J. Real-time individualization of the unified model of performance. *J Sleep Res.* 2017;**26**(6):820–831. doi: 10.1111/jsr.12535
48. Reifman J, Ramakrishnan S, Liu J, *et al. 2B-Alert* App: a mobile application for real-time individualized prediction of alertness. *J Sleep Res.* 2019;**28**(2):e12725. doi: 10.1111/jsr.12725
49. Nehlig A. Is caffeine a cognitive enhancer? *J Alzheimers Dis.* 2010;**20**(Suppl 1):S85–S94. doi: 10.3233/JAD-2010-091315