SLEEP, 2023, 46, 1–13
https://doi.org/10.1093/sleep/zsad080
Advance access publication 29 March 2023
Original Article

Original Article

2B-Alert App 2.0: personalized caffeine recommendations for optimal alertness

Francisco G. Vital-Lopez[1,2], Tracy J. Doty[3], Ian Anlap[4], William D. S. Killgore[4] 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,
[3]Behavioral Biology Branch, Walter Reed Army Institute of Research, Silver Spring, MD, USA and
[4]Department of Psychiatry, University of Arizona College of Medicine, Tucson, AZ, 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: If properly consumed, caffeine can safely and effectively mitigate the effects of sleep loss on alertness. However, there are no tools to determine the amount and time to consume caffeine to maximize its effectiveness. Here, we extended the capabilities of the *2B-Alert* app, a unique smartphone application that learns an individual's trait-like response to sleep loss, to provide personalized caffeine recommendations to optimize alertness.

Methods: We prospectively validated *2B-Alert*'s capabilities in a 62-hour total sleep deprivation study in which 21 participants used the app to measure their alertness throughout the study via the psychomotor vigilance test (PVT). Using PVT data collected during the first 36 hours of the sleep challenge, the app learned the participant's sleep-loss response and provided personalized caffeine recommendations so that each participant would sustain alertness at a pre-specified target level (mean response time of 270 milliseconds) during a 6-hour period starting at 44 hours of wakefulness, using the least amount of caffeine possible. Starting at 42 hours, participants consumed 0 to 800 mg of caffeine, per the app recommendation.

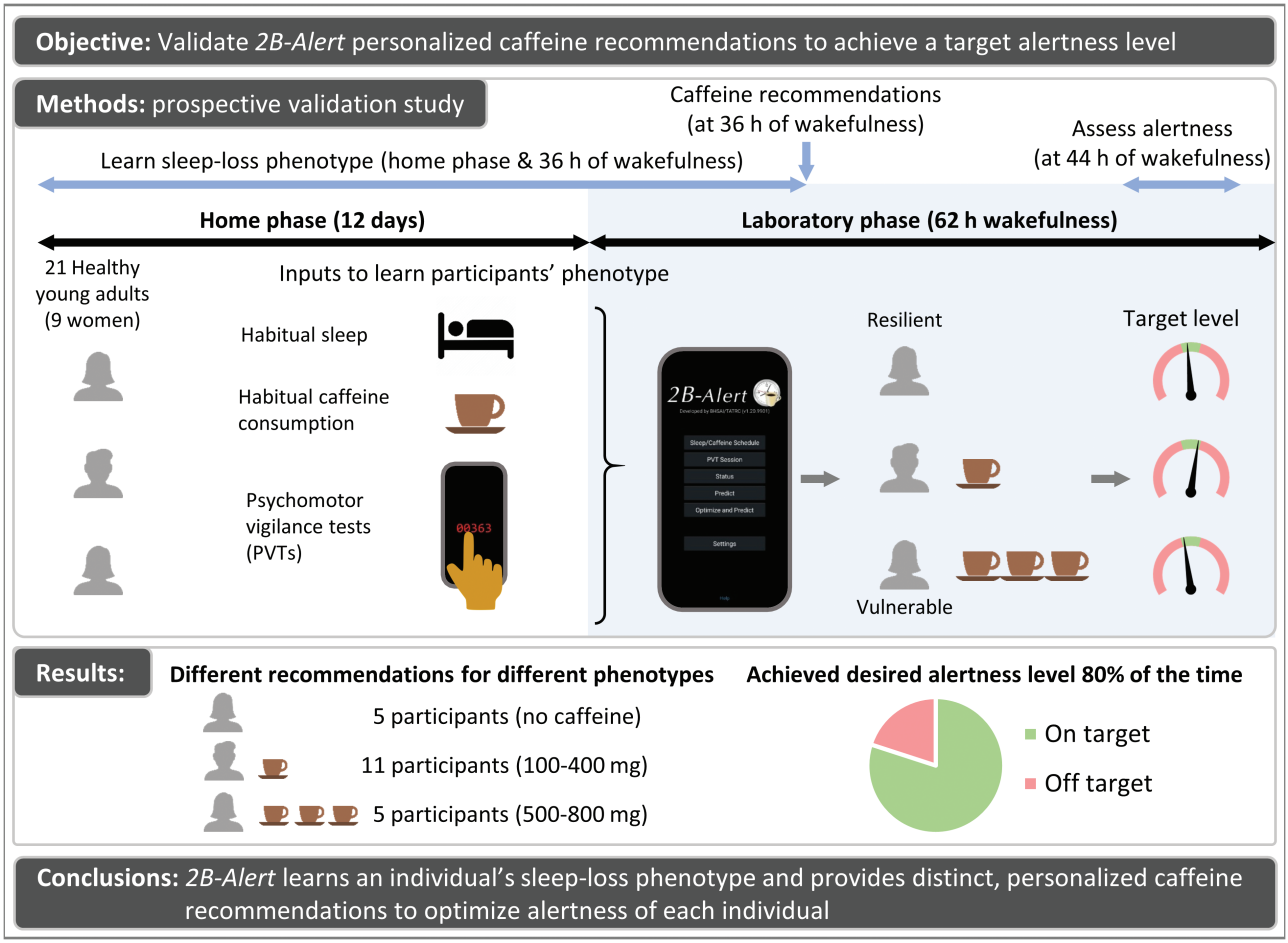
Results: *2B-Alert* recommended no caffeine to five participants, 100–400 mg to 11 participants, and 500–800 mg to five participants. Regardless of the consumed amount, participants sustained the target alertness level ~80% of the time.

Conclusions: *2B-Alert* automatically learns an individual's phenotype and provides personalized caffeine recommendations in real time so that individuals achieve a desired alertness level regardless of their sleep-loss susceptibility.

Key words: alertness; caffeine; mathematical model; personalized predictions; sleep loss; smartphone app

Submitted for publication: December 1, 2022; **Revised:** February 28, 2023

Graphical Abstract



Statement of Significance

Insufficient sleep causes millions of individuals to carry out daily activities with suboptimal alertness, affecting safety and productivity. Caffeine consumption can help mitigate these effects, however, to be effective, caffeine should be consumed at the right time and in the right amount, depending on the individual's sleep history, work schedule, and, importantly, susceptibility to sleep loss. Here, for the first time, we demonstrated that a smartphone application can provide personalized caffeine recommendations to allow individuals under the same sleep-deprivation condition to achieve the same desired level of alertness, regardless of their susceptibility to sleep loss.

Introduction

Sleep loss is a common stressor that negatively affects the health and performance of otherwise healthy individuals [1]. In particular, insufficient sleep is associated with next-day cognitive deficits that may compromise productivity and safety [2–5]. In fact, short sleep duration (e.g. ≤6 hours per day) has been linked to a 35% increased risk of workplace accidents compared with healthy sleep (7 to 9 hours per day) [6, 7] and is prevalent in both civilians (~35% of adults in the United States [8]) and military personnel (~63% of US Service members [9]). Moreover, it is estimated that the economic cost of insufficient sleep in the United States could be up to 2.3% of the gross domestic product (or $411 billion in 2016) [10].

Caffeine is the most widely used stimulant worldwide, with nearly 85% of the US population consuming at least one caffeinated product daily, for an average intake of ~165 mg [11, 12]. With the intent to counteract sleepiness [13], individuals consume caffeine ad libitum according to their perceived needs. Under chronic sleep-loss conditions, this self-prescription may result in a vicious cycle where high caffeine consumption leads to sleep disturbances and increases next-day tiredness, which, in turn, leads to higher caffeine consumption [14]. One example of this vicious cycle was observed in military personnel, where one report by the Centers for Disease Control and Prevention found that US Service members deployed to Afghanistan who consumed large amounts of caffeinated energy drinks were more likely to fall asleep in briefings and on guard duty than those who consumed less caffeine [15]. This vicious cycle may persist even after returning from deployment, as one in six US Service members continues to consume large amounts of energy drinks and experiences sleep problems and fatigue [16]. In addition, if consumed in excess, caffeine can have adverse effects, such as headache, anxiety, nausea,

irritability, dizziness, and, in extreme cases, tachycardia, arrhythmia, altered consciousness, and seizures [17, 18]. Interestingly, Doty et al. found that individuals who consume large amounts of caffeine during several days of restricted sleep may take longer to recover to baseline levels of alertness than those who do not consume caffeine [19]. Ideally, individuals should consume optimal caffeine doses (i.e. the correct amount at the right time) for it to be safe and effective as a sleep-loss countermeasure, while minimizing side effects.

Multiple studies have shown that caffeine can be safely and effectively used to sustain alertness during prolonged periods of restricted sleep [20–25]. However, it is not clear how to extrapolate caffeine schedules used as a countermeasure in one particular condition to a different condition, because the optimal caffeine dosage depends on the time and duration of the desired peak alertness periods (such as work hours), extent of sleep loss (which depends on sleep history), and amount of caffeine consumed in the recent past (i.e. the last 24 hours). In addition, we hypothesize that an individual's trait-like response to sleep loss [26, 27] is an important factor in determining the optimal dosage, with vulnerable individuals requiring more caffeine than those who are resilient to achieve the same level of alertness under the same sleep-loss condition.

To provide personalized caffeine countermeasures against sleep loss, over the years, our US Army group has developed (1) a mathematical model, the Unified Model of Performance (UMP), to predict the effects of sleep history, time of day, and caffeine consumption on alertness, as measured by the psychomotor vigilance test (PVT) [28–30], (2) a machine-learning algorithm that uses PVT data to automatically learn in real time an individual's trait-like response to sleep loss [31, 32], and (3) an optimization algorithm that uses UMP predictions to efficiently search through a large number of potential caffeine recommendations and automatically identify the one that maximizes alertness for the desired peak alertness period with the least possible amount of caffeine (i.e. the optimal dosage) [33]. The integration of these capabilities culminated in the *2B-Alert* app, a smartphone application for fatigue management that provides personalized alertness predictions and caffeine recommendations. To validate the first two capabilities, we recently conducted a prospective study where we use the *2B-Alert* app to learn a participant's trait-like response to sleep loss in a 62-hour total sleep deprivation (TSD) study under caffeine-free conditions [32]. For the first 36 hours of the challenge, 21 participants used the app to perform 12 PVTs, which the app used to automatically learn the trait-like response of each participant, customize the UMP, and provide personalized alertness predictions for the remainder of the challenge. Comparison of the personalized predictions with the measured PVT data for the last 26 hours of the challenge showed that the app is able to learn each participant's response to sleep loss and accurately predict their alertness (average error of 54 milliseconds) [32].

In this follow-up study, we prospectively validated the third capability of the *2B-Alert* app, to provide effective personalized caffeine recommendations tailored to the participant's trait-like response to sleep loss in a 62-hour TSD challenge. For the first 36 hours of the study, participants used the app to perform PVTs and, in the background, the app used these data to automatically and progressively learn the trait-like response to sleep loss of each participant. Then, at 36 hours into the challenge, the app provided personalized caffeine recommendations, with each participant consuming between 0 and 800 mg of caffeine, so as to sustain alertness at a desired target level during a 6-hour period on the second night of the TSD challenge. Subsequently, participants continued to perform PVTs during the remaining 26 hours of the challenge to assess the effectiveness of *2B-Alert*'s caffeine recommendations.

Materials and Methods

The *2B-Alert* app

Functionalities

Users interact with the current version of the *2B-Alert* app to perform three main tasks: (1) measure real-time alertness via the PVT, which the app also uses to personalize the UMP predictions. (2) predict alertness as a function of the user's sleep history, caffeine-consumption schedule, and time of day, and (3) recommend optimal caffeine dosages to mitigate alertness impairment during user-specified peak alertness periods. The app performs tasks 2 and 3 at both the group-average and individualized levels.

Inputs and outputs

The *2B-Alert* app requires as many as four user-provided inputs, depending on the selected tasks: (1) alertness measurements via PVT sessions (required for personalized predictions), (2) recent past (at least 5 days) and future sleep schedules (required), (3) recent past (at least 24 hours) and future caffeine-consumption schedules (required if caffeine is consumed), and (4) desired peak alertness periods (required for caffeine recommendations). Depending on the selected tasks, *2B-Alert* provides three main outputs: (1) alertness measurement, (2) alertness prediction, and (3) optimal caffeine recommendation.

Interface

Figure 1 shows screen-capture images of the main elements of the current *2B-Alert* app interface, including the main menu (A); sleep, caffeine, and peak alertness schedules (B); PVT stimulus (C); and prediction displays (D–F). On the main menu (Figure 1A), users can access the screens to provide inputs via "Sleep/Caffeine Schedule" and "PVT Session." Tapping on "Sleep/Caffeine Schedule" shows the schedule overview (Figure 1B), where users can view their inputs, i.e. sleep (blue bars) and caffeine (yellow dots) schedules, as well as their peak alertness periods (white dashed box). Users can view details as well as add, delete, or modify their inputs on the corresponding tabs on the top. Access to the remaining input (i.e. PVT data), required to obtain personalized predictions, is obtained by tapping on "PVT Session".

For the user-provided sleep and caffeine schedules, the app generates group-average and personalized alertness predictions, as determined by the mean response time (RT) in the PVT. (If users do not perform PVT sessions, then the personalized predictions are the same as the group-average predictions.) The predictions can be accessed by tapping "Status" or "Predict" on the main menu. On the "Status" display (Figure 1D), the app shows a dial with a needle indicating the corresponding personalized alertness level (mean RT of 248 milliseconds) at the corresponding time (November 11, 2022; 19:00). Users can drag the blue dot at the bottom of the screen horizontally to view predictions from the previous 48 hours or for the next 48 hours. On the "Prediction" display (Figure 1E), users can access a more detailed view of the alertness predictions and PVT data. On this screen, the app shows the predicted mean RT (*y*-axis; yellow solid line) and the PVT mean RT data (green dots) as a function of time (*x*-axis). Users can select to visualize "Group Prediction" or "Individualized

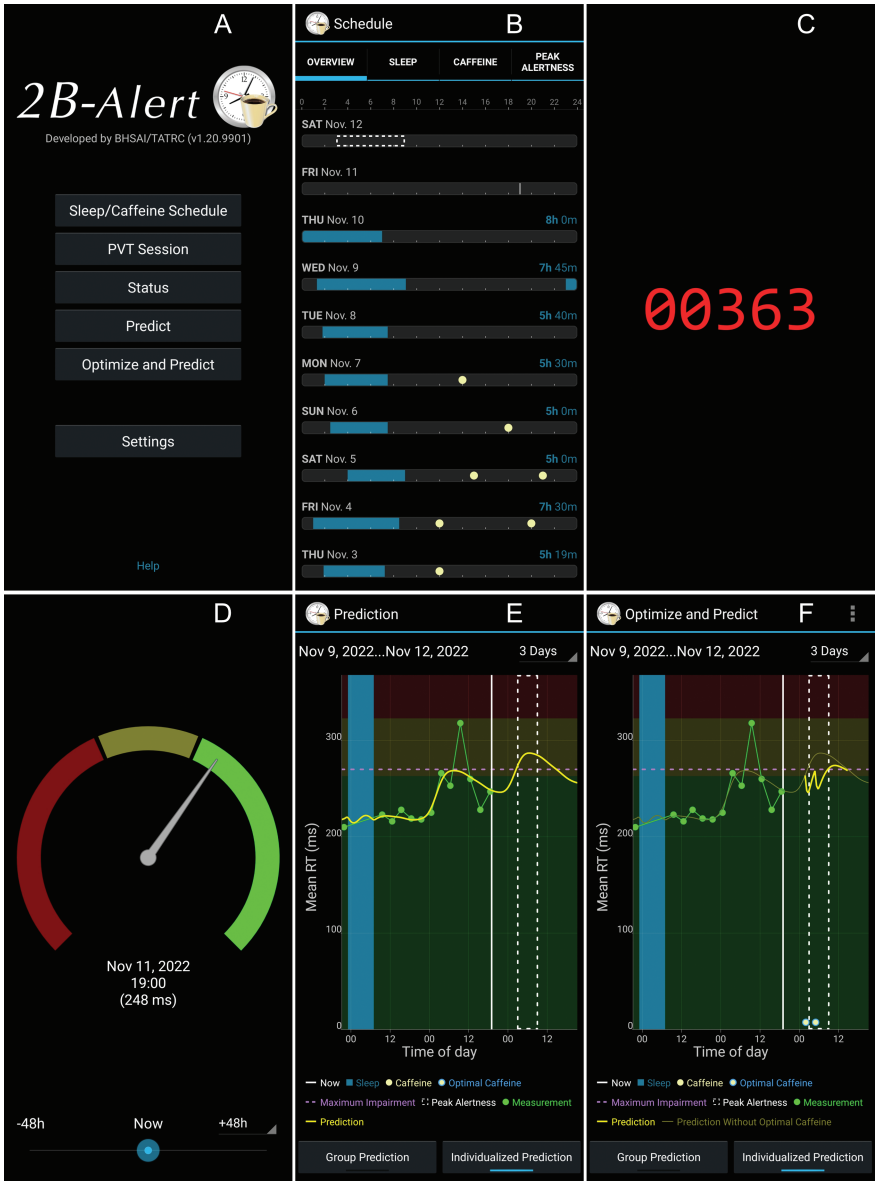


Figure 1. Screen capture of the key components of the *2B-Alert* app. (A) The main menu provides access to the app functionalities. (B) "Sleep/Caffeine Schedule" provides an overview of the user's schedule, showing sleep (blue bars) and caffeine (yellow dots) inputs, as well as the desired period of peak alertness (white dashed box). (C) "PVT Session" allows the user to perform a psychomotor vigilance test (PVT). The image shows a stimulus displayed during a PVT session. (D) "Status" displays alertness predictions through a dial-and-needle representation and (E) "Predict" displays alertness predictions (yellow line) for an individual, or a group of individuals, and the PVT results in terms of mean response time (RT) (green dots). (F) "Optimize and Predict" displays alertness predictions with the optimal caffeine recommendation (thick yellow line) and without caffeine (thin yellow line) and the timing of the caffeine doses (yellow dots with blue circles).

Prediction" (shown in Figure 1E). Users can also obtain personalized or group-average caffeine recommendations by tapping "Optimize and Predict" on the main menu, which takes the user to the screen in Figure 1F. Here, the user obtains caffeine recommendations in real time to optimize alertness during the selected peak alertness period (dashed white box in Figures 1B and 1F), which are saved in the caffeine schedule. Figure 1F also shows the predictions with the caffeine recommendation (thick yellow line) and without the caffeine recommendation (thin yellow line).

Tapping "Settings" on the main menu (Figure 1A) allows users to change the settings for the PVT sessions, including session duration (3, 5, or 10 minutes) and inter-stimulus interval (ISI; from 1 to 10 s, depending on the session duration), and caffeine optimization, including maximum acceptable alertness-impairment level for the peak alertness periods and maximum allowed amount of caffeine consumption in a 24-hour period.

Personalized unified model of performance

We started the development of the UMP [28–30] based on the seminal two-process model of Borbély [34], and extended it to account for the dampening effect of accumulated sleep debt on recovery sleep [28], individual differences in trait-like response to sleep loss [31, 32], and the stimulating effects of caffeine on alertness [30]. Briefly, the UMP predicts the effects of sleep history, time of day t, and caffeine dose c on alertness impairment (P), as determined by the mean RT in the PVT:

$$P(t,\theta) = P_0(t,\theta) \times g_{PD}(t,\theta,\ c) \quad (1)$$

$$P_0(t,\theta) = S(t,\theta) + \kappa C(t,\theta) \quad (2)$$

where P_0 denotes the alertness impairment without caffeine consumption, S represents the homeostatic process, κ denotes the amplitude of the circadian process C, g_{PD} represents the stimulating effect of caffeine, and θ represents 12 model parameters. We previously estimated θ for predicting the average alertness level for a group of healthy, young adults [29, 30]. However, for personalized predictions, we found that we only needed to adjust the value of five key, most sensitive parameters while the remaining parameters are set to the group-average value [31]. We refer the reader to Priezjev et al. [35] for a complete list of equations and parameter values for group-average model predictions.

To personalize the *2B-Alert* predictions, we individualized the five model parameters through a recurring process, during which the response to sleep loss of each individual is progressively learned using the results of the PVT sessions. The process starts with the assumption that the response of the individual is the same as that of the group average (i.e. the values for θ are those of the group-average model). Immediately after the user performs the first PVT, using a Bayesian learning algorithm [31, 32], the app automatically estimates, in real time, a new θ that reflects the individual's PVT data. This process repeats itself, where with each subsequent PVT the accuracy of the θ estimates increases, converging asymptotically to the "true" θ for the individual. A detailed description of the model's individualization algorithm can be found in Liu et al. [31].

We extensively validated the UMP predictions, both at the group-average and individual levels, under multiple sleep and caffeine-consumption conditions, including chronic sleep restriction (CSR; 3–5 hours of sleep per night), TSD (up to 88 hours), combinations of CSR and TSD, daytime sleep and sleep extension, as well as single and multiple caffeine doses ranging from 50 to 600 mg [29, 30, 35]. In particular, very recently, in a validation study using experimental data from 12 studies involving 301 unique participants, Priezjev et al. [35] showed that, for 244 participants (81%), three out of four personalized model predictions were indistinguishable from the PVT measurements. Of note, in these analyses, we used all reported PVT data for each participant, including obvious outliers.

Caffeine optimization

Previously, we developed an optimization algorithm that uses the predictive capabilities of the UMP to identify, in real time, caffeine recommendations to sustain alertness during sleep loss [33]. Briefly, the algorithm searches for caffeine solutions (i.e. time and amount) that mitigate alertness impairment to a desired target level (i.e. a maximum alertness-impairment threshold) for the desired peak alertness periods, using the least amount of caffeine possible. The algorithm uses the following constraints to obtain practical and safe solutions: (1) caffeine doses are restricted to 100, 200, or 300 mg, (2) dosing occurs on the hour, (3) the minimum time between doses is 2 hours, (4) the total amount of recommended caffeine in a 24-hour period does not exceed 800 mg, and (5) the caffeine concentration in the blood never exceeds the maximum level achieved by a single dose of 400 mg [36]. To obtain solutions in a matter of seconds, the algorithm iteratively generates new solutions by making "smart" changes to the current solution that are likely to reduce the alertness impairment while satisfying the imposed constraints. We refer the reader to Vital-Lopez et al. [33] for a detailed description of the optimization algorithm.

Prospective validation study

To assess the capability of the *2B-Alert* app to provide, in real time, personalized caffeine recommendations for each participant, we performed a prospective TSD study (Figure 2) at the Social, Cognitive, and Affective Neuroscience Lab at the University of Arizona (Tucson, AZ). We recruited healthy, young men and non-pregnant, non-lactating women and screened them to ensure they did not have sleep disorders or mental and physical health problems. Individuals self-reporting to consume more than 400 mg of caffeine per day were not eligible to participate. In addition, participants had to demonstrate compliance with the protocol throughout all phases of the study to be allowed to complete the study. Twenty-one participants (9 women), ranging in age from 18 to 36 years (mean = 21.9 years, standard deviation [SD] = 4.4 years), completed the study. The Institutional Review Board of the University of Arizona and the Human Research Protection Office (now known as the Office of Human Research Oversight) of the US Army Medical Research and Development Command (Fort Detrick, MD) approved the study, and each participant provided a written informed consent.

The study consisted of a home phase followed by an in-laboratory phase (Figure 2). At the beginning of the home phase, participants received a smartphone (Samsung Galaxy S20) with the *2B-Alert* app installed and were instructed to record their sleep and caffeine-consumption diaries in the app. To log their sleep, participants were instructed to enter the time they started to attempt to fall sleep each night and the time they woke up each morning. To help keep track of their caffeine consumption, participants were informed of products that contain caffeine and trained to use the app to log caffeine data. The app includes a menu with the caffeine content of popular caffeinated products and also provides the option for users to enter raw caffeine amounts. Participants were instructed to consume no more than 400 mg caffeine per day during the home phase.

Participants were instructed to perform five or six 5-minute PVTs (ISI: 1–4 seconds) per day, every 2–3 hours during waking hours. To increase compliance with the protocol, research staff contacted participants three times during the home phase (study days 2, 6, and 11) to check on their progress and remind them of data-collection requirements. In addition, the *2B-Alert* app automatically displayed daily notifications every 3 hours (during daytime) to remind participants to perform a PVT session and twice a day (early morning and late evening) to remind them to log their sleep and caffeine-consumption data.

Previously, in a similar study with a 62-hour TSD challenge, we showed that a 5-minute PVT with an ISI of 1–4 seconds is effective in capturing the effects of sleep loss and the circadian process on alertness [32]. Moreover, a validation study have shown that a similar configuration of a 5-minute PVT with an ISI of 1–5 seconds is comparably as sensitive to sleep loss as the standard 10-minute PVT with an ISI of 2–10 seconds [37], and other studies have also shown that this configuration is sensitive to sleep loss [20, 23, 25]. Although the app allows for different combinations of PVT session duration and ISI, in our study, we set the PVT duration to 5 minutes, the ISI to 1–4 second, and blocked other options on the app to ensure that each participant performed all PVT sessions with the same configuration. Participants also received a Philips Respironics Actiwatch-2 wrist actigraphy watch to corroborate the sleep schedule recorded in the app and were instructed to

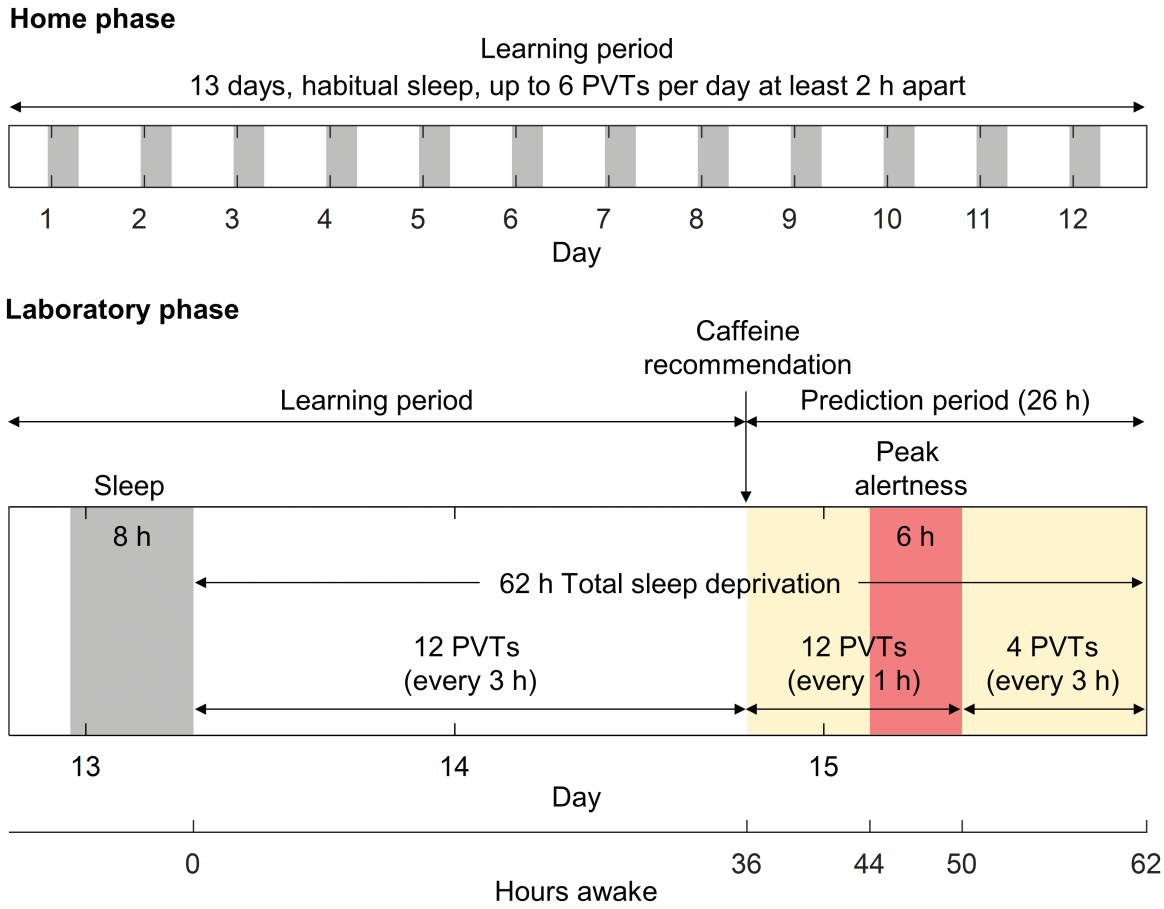


Figure 2. Schedule of the study to prospectively validate the *2B-Alert* app. The study consisted of a 12-day home phase followed by an in-laboratory phase. During the home phase, participants maintained their habitual sleep and caffeine-consumption schedules, and used the app to record these data. Participants also used the app to perform five or six 5-minute psychomotor vigilance tests (PVTs) per day, each at least 2 hours apart. On day 12, participants arrived at the sleep center to begin the laboratory phase, where they had an 8-hour sleep opportunity (from 23:00 to 07:00). Starting at 07:00 on day 13, participants underwent 62 hours of total sleep deprivation (TSD). For the first 36 hours, participants performed 5-minute PVTs every 3 hours. At the end of the learning period (at 36 hours of TSD), we used the app to obtain a personalized caffeine recommendation so that each participant would sustain the same maximum alertness-impairment level (270 milliseconds) during the 6-hour peak-alertness period (from 03:00 to 09:00 on day 15, red shade). Each participant consumed caffeine according to the app's personalized recommendation and continued to perform 5-minute PVTs. We used the personalized models obtained at the time of the caffeine recommendation (19:00 on day 14) to predict the alertness impairment for the prediction period (yellow shade) and compared the models' predictions with the subsequently collected PVT data, to assess the ability of the *2B-Alert* app to learn the trait-like response to sleep loss and provide personalized caffeine recommendations for each participant.

mark the start and end of each sleep episode by pressing a button in the watch. We used the Actiware 6.2.0 program to score the actigraphy data according to standard procedures (wake threshold set to medium) and obtained standard metrics of sleep, including total sleep time, for each sleep episode.

On day 12, participants who completed the home phase reported to the laboratory at 19:00 and had their sleep and caffeine dairies and PVT data inspected to assess compliance with the protocol. Non-compliant participants were discharged, and compliant participants continued with the in-laboratory phase. They went to bed at 23:00, woke up at 07:00 the next morning, and underwent 62 hours of TSD followed by one night of recovery sleep (12 hours). Participants performed a 5-minute PVT (ISI: 1–4 s) every 3 hours for the first 36 hours of TSD (12 PVT sessions), every 1 hour for the following 12 hours and every 3 hours for the rest of the study (16 PVT sessions), for a total of 28 PVTs during the in-laboratory phase.

The *2B-Alert* app used PVT data collected during the home phase and the first 36 hours of the TSD challenge ("the learning period") to personalize the model (Figure 2). At the end of this period, the app provided a caffeine recommendation for each participant to sustain performance at a maximum alertness-impairment threshold (mean RT of 270 milliseconds) for 6 hours beginning after 44 hours of TSD (i.e. a peak alertness period between 03:00 and 09:00 on day 15). We selected the same threshold for all participants, rather than a participant-specific threshold based on their baseline performance, because computer simulations during the design phase of this study indicated that a fixed threshold would result in a wider range of caffeine recommendations than a relative threshold. Moreover, the selected target threshold of 270 milliseconds is equivalent to the alertness impairment level of a blood alcohol concentration of ~0.06% [38–40], which is associated with a twofold increase in the risk of causing a traffic accident as compared with control drivers [41, 42].

For each participant, the app constrained its recommendation to no more than a total of 800 mg of caffeine distributed in multiple doses, with a maximum single dose of 300 mg. At each recommended time, participants consumed the corresponding dose by chewing caffeinated gum (Military Energy chewing gum, MarketRight, Inc., Plano, IL; 100 mg caffeine per piece) for 10

minutes. Chewing this caffeinated gum formulation for 5 minutes yields a normalized caffeine bioavailability of 90% [43], and chewing it for 10 minutes yields 96% of caffeine extraction [44]. Moreover, multiple studies have shown the effectiveness of this gum product in reducing alertness impairment caused by sleep loss [19, 23, 45].

The app collected PVT data and provided personalized predictions and caffeine recommendations in real time during the study, which were stored in the app and retrieved after the end of the study, as reported herein. However, to prevent participants from becoming predisposed to their alertness predictions and caffeine recommendations, we blinded them from the *2B-Alert* outputs.

Metrics to assess the *2B-Alert* personalized predictions and caffeine recommendations

We used two metrics to quantify *2B-Alert*'s ability to learn and predict the trait-like response to sleep loss of each participant. We computed the root mean square error (RMSE) between the model predictions P_n in equation 1 for the n-th PVT session, where $n = 1, 2, \ldots, N = 16$, and the mean RT PVT measurement y_n during the prediction period (i.e. the last 26 hours of the challenge; Figure 2):

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(P_n - y_n)^2}. \quad (3)$$

We also computed the percentage of the mean RT data that fell within the model's prediction intervals (PI_n):

$$PI_n = P_n \pm 2\sigma_{ws} \quad (4)$$

where σ_{ws} denotes the variance of the mean RT measurements upon repeated trials by the same participant under the same condition (i.e. a measure of within-subject variability), which we conservatively assumed to be 30 milliseconds, based on analysis of baseline PVT sessions of sleep-satiated participants [26, 46].

To quantify *2B-Alert*'s ability to provide personalized caffeine recommendations to sustain alertness impairment around the 270-milliseconds threshold during the 6-hour peak alertness period starting at 44 hours of the TSD challenge, we computed the percentage of the mean RT data that fell within the alertness interval Q around this threshold:

$$Q = 270 \pm 2\sigma_{ws} \quad (5)$$

where Q represents the interval in which we expect ~95% of the measurements to fall if the alertness-impairment level of the participant is 270 milliseconds and the variance of the mean RT data, corresponding to the within-subject variability, is conservatively estimated as $\sigma_{ws} = 30$ milliseconds. When the app recommended no caffeine, we counted all data that fell below the upper bound of Q (i.e. $270 + 2\sigma_{ws}$).

Results

Sleep, caffeine consumption, and PVT data during the home phase

Thirty participants enrolled in the study, however, only 21 (nine women) completed the TSD challenge. Based on the sleep diaries recorded in the app, participants slept an average of 7.5 hours (SD = 0.5 hours) per night during the home phase vs. 8.0 hours (SD = 0.6 hours) per night based on the actigraphy data. Also, based on the entries recorded in the app, participants consumed an average of 57 mg (SD = 58 mg) of caffeine per day, where three participants did not consume any caffeine and five participants consumed more than 100 mg per day. On average, participants performed 5.6 (SD = 0.3) PVT sessions per day during the 12-day home phase. Supplementary Figure S1 in Supplementary Material shows the PVT data, sleep schedule, and caffeine consumption recorded in the app for each participant who completed the 62-hour TSD challenge.

2B-Alert personalized predictions

First, we validated *2B-Alert*'s ability to learn each participant's response to sleep loss and predict alertness for the last 26 hours of the TSD challenge. To this end, we compared *2B-Alert*'s predictions of alertness impairment, obtained in real time using the models personalized with data collected during the home phase and the first 36 hours of the TSD challenge, with the PVT data collected during the prediction period (Figure 2). Figure 3 shows the personalized simulations (blue solid lines) for the entire 62-hour TSD challenge and the PVT data used by the app to learn each participant (blue symbols "+") and for prediction assessment (blue circles). We observed a wide range of responses to sleep loss, with the measured impairment of participant #1 remaining below the maximum alertness-impairment threshold (dotted black line) throughout the challenge and the measured impairment of participant #21 markedly exceeding the threshold, even during the first day of the challenge.

To quantify the accuracy of the predictions, we computed the RMSE between the model predictions and the mean RT data as well as the percentage of the data that fell within the PIs around the model predictions (equation 4). Table 1 shows these metrics for the learning and prediction periods. Overall, the app's alertness outputs accurately fitted the PVT data during the learning period, with an average RMSE of 24 milliseconds (SD = 8 milliseconds) and 97% (SD = 4%) of the data within falling the PIs. Similarly, we observed accurate personalized predictions for the last 26 hours of the challenge, which included the period after caffeine consumption, with an average RMSE of 50 milliseconds (SD = 24 milliseconds) and 80% (SD = 22%) of the data falling within the PIs for the entire prediction period. Except for participants #3 and #19, who had both a large RMSE (115 and 90 milliseconds, respectively) and a low percentage of measurements within the PIs (31% and 25%, respectively), and participant #4, whose alertness impairment was consistently underestimated during the peak alertness period, the *2B-Alert* app was able to learn the trait-like responses of 18 of the 21 participants (86%). Note that we used all the data collected during the study, including obvious outliers, for personalizing the model and assessing the predictions, as it would be done in a real-life application.

Personalized caffeine recommendations

At 36 hours into the TSD challenge, using the personalized models up to this point, *2B-Alert* provided, automatically and in real time, a personalized caffeine recommendation for each participant. The goal of the recommendations was for each participant to sustain performance at the same target alertness-impairment threshold of 270 milliseconds during the 6-hour peak alertness period beginning after 44 hours of the challenge (Figure 2), regardless of the participants' trait-like response to sleep loss. In accordance with the observed wide range of responses to sleep loss, the app provided recommendations that varied from no caffeine to a maximum of 800 mg. Examination of the mean RT data collected during the peak alertness period (as well as the entire prediction period) confirmed that, overall, the recommendations correctly reflected the participants' trait-like response to sleep loss (Figure 3), prescribing no caffeine to the true most resilient

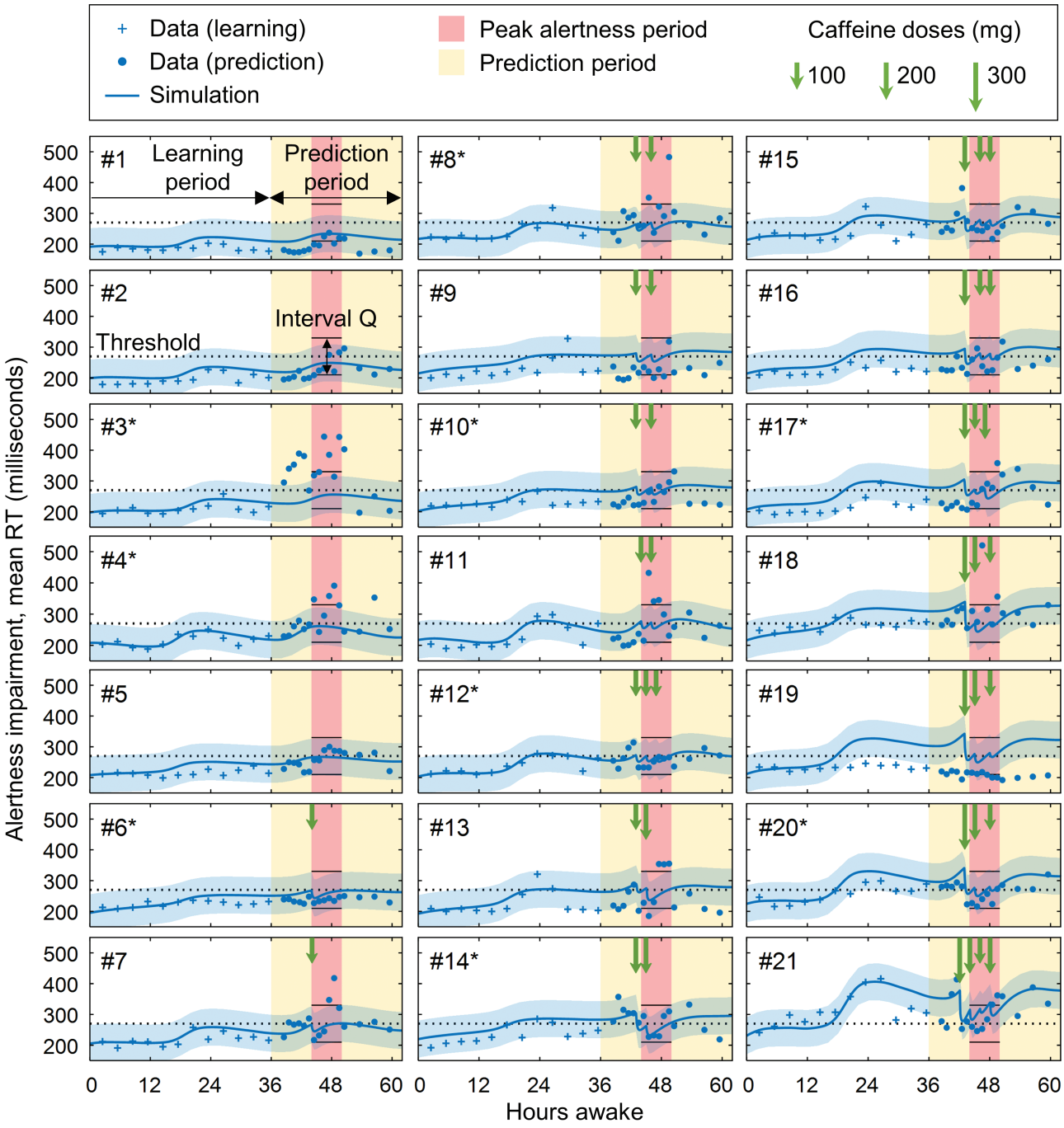


Figure 3. Personalized *2B-Alert* alertness-impairment predictions and caffeine recommendations. Using the personalized models obtained at the end of the "learning period" (36 hours into the 62-hour total sleep deprivation challenge), the *2B-Alert* app provided personalized caffeine recommendations (green arrows) so that each participant could sustain alertness impairment at the selected maximum alertness-impairment threshold of 270 milliseconds (dotted black lines) during the 6 hours of peak alertness (red shade). The blue "+" symbols correspond to the PVT data used to train the personalized models, and the blue circles correspond to the PVT data used to assess the *2B-Alert* predictions. The blue solid lines correspond to the personalized predictions at the end of the learning period. The blue shaded regions correspond to the 95% PIs (equation 4) and the horizontal black lines within the peak alertness period indicate the interval Q in which we expect ~95% of the measurements to fall if the alertness-impairment level of the participant is 270 milliseconds (equation 5). Women are indicated with an asterisk (*).

participants (#1, #2, and #5) and ≥600 mg to the true most vulnerable participants (#18, #20, and #21), using on average 281 mg (SD = 240 mg) of caffeine per participant (Table 1). Also, we observed no association between the participants' habitual caffeine consumption, recorded during the home phase of the study, and the total amount of caffeine recommended by *2B-Alert*. For example, the app recommended substantially more caffeine to participants #15 to #21 than to participants #1 to #7 [on average 557 mg (SD = 140 mg) vs. 29 mg (SD = 49 mg)], even though both groups consumed similar amounts of caffeine during the home phase [on average 60 mg (SD = 79 mg) vs. 54 mg (SD = 44 mg) per day]. Note that during the home phase, participants did not have much need to consume caffeine to improve performance as they slept on average 7.5 hours per night.

Table 1. Performance of the personalized models, assessment of alertness impairment during the peak alertness period, and total amount of caffeine recommended for each participant who underwent 62 hours of total sleep deprivation

Participant #	Personalized model accuracy				Optimal solution	
	RMSE (milliseconds)		Percentage of PVTs within 95% PI[†]		Percentage of PVTs within interval Q[‡] during peak alertness period	Total caffeine recommended (mg)
	Learning period	Prediction period	Learning period	Prediction period		
1	17	34	100	100	100	0
2	19	26	100	100	100	0
3*	20	115	99	31	50	0
4*	15	61	99	69	50	0
5	24	24	98	100	100	0
6*	18	23	100	100	100	100
7	32	48	95	88	67	100
8*	16	66	100	81	67	200
9	40	56	86	63	67	200
10*	21	40	99	94	100	200
11	15	60	100	81	50	200
12*	11	23	100	100	100	300
13	23	61	96	50	33	300
14*	32	36	93	88	100	400
15	21	34	99	94	100	400
16	32	43	94	88	100	400
17*	28	56	94	63	83	500
18	30	77	96	88	67	600
19	36	90	87	25	50	600
20*	23	31	100	100	100	600
21	27	49	98	81	67	800
Average (SD)	24 (8)	50 (24)	97 (4)	80 (22)	79 (23)	281 (240)

PVT, psychomotor vigilance test; RMSE, root mean square error; SD, standard deviation.
[†]PI: prediction interval around the model predictions (equation 4).
[‡]Q: interval around the maximum alertness impairment level (equation 5).
*Woman.

To quantify the extent to which the personalized caffeine recommendations yielded sustained alertness impairment around the target threshold level, we computed the percentage of mean RT data during the peak alertness period that fell within the alertness interval Q (equation 5). We observed that 16 of the 21 participants (76%) sustained performance at the target level during most of the peak alertness period, with ≥ 67% of the measurements falling within Q (Table 1, second to last column). Of the five participants who had ≤ 50% of the measurements within Q, the app failed to learn the response to sleep loss for three of them (#3, #4, and #19). For the other two participants (#11 and #13), the app underestimated the amount of caffeine needed for them to sustain the target alertness level. For these five participants, we did not observe a difference in their caffeine consumption during the home phase [on average 62 mg (SD = 74 mg) per day] with respect to the rest of the participants [on average 56 mg (SD = 55 mg) per day]. Overall, including all participants, 79% (SD = 23%) of the mean RT data fell within Q. That is, the app's personalized recommendations resulted in sustained alertness impairment at the desired target level in nearly 80% of the time.

Discussion

If consumed at the right time and in the right amount, caffeine can effectively and safely help mitigate the effects of sleep loss on alertness. However, currently, there are no tools that sleep-deprived individuals can use to obtain effective caffeine dosages tailored to their needs. To close this gap, we extended the *2B-Alert* app, a smartphone application with the unique capability to automatically learn an individual's trait-like response to sleep loss and make effective personalized caffeine recommendations specific to each individual's susceptibility to sleep loss. To assess the *2B-Alert* app, we conducted a prospective 62-hour TSD study in which we used the app to obtain, in real time, a personalized caffeine recommendation for each participant to sustain alertness at a desired target impairment threshold level during the second night of the challenge.

Previously, Reifman et al. [32] demonstrated that the *2B-Alert* app can automatically learn an individual's trait-like response to sleep loss and predict future alertness under caffeine-free conditions. Following a similar protocol, in the current study, we re-assessed the app's capability to provide personalized alertness

predictions after caffeine consumption. We confirmed that the prediction accuracy of the personalized models generated using PVT data collected up to the end of the learning period was consistent with our previous observations, with an average RMSE for the prediction period of 50 milliseconds in this study (Table 1) versus 54 milliseconds for the caffeine-free condition in the previous study [32]. The average RMSE for the present study was also smaller than the average personalized prediction RMSE of 67 milliseconds, we observed in a retrospective analysis of *2B-Alert* based on 12 different studies, including 22 different sleep and caffeine conditions [35]. We also observed consistent, slightly better results in terms of the percentage of measurements within the 95% PIs: 80% in this study (Table 1) vs. 75% for the caffeine-free condition [32] and 71% in the retrospective validation study [35]. Moreover, in both prospective studies, the accuracy of the models used to obtain the personalized predictions was comparable to the highest-possible accuracy, which was achieved when all PVT data at the end of the study were used to fit the models, with average RMSEs of 42 milliseconds (84% of measurements falling within the 95% PIs) in this study and 46 milliseconds (82%) in the previous caffeine-free study [32]. These results are consistent with our modeling assumption and previous experimental observations, which showed that the effect of caffeine on alertness is more closely dependent on an individual's impairment level and trait-like response to sleep loss than on their sensitivity to caffeine [47, 48].

After confirming that the *2B-Alert* app was able to learn the participant's trait-like response to sleep loss, we assessed the app's ability to provide effective personalized caffeine recommendations in real time to sustain alertness at a pre-defined target level and duration, regardless of the participant's vulnerability to sleep loss. Overall, the app provided suitable recommendations using the least-possible amount of caffeine, prescribing no caffeine for the most resilient participants and gradually increasing the total caffeine amount to ≥ 600 mg for those who were the most vulnerable (Table 1).

Notably, the recommendations allowed the participants to sustain alertness at the desired target threshold level (270 milliseconds) for the desired 6-hour duration nearly 80% of the time, with 76% of participants (16 out of 21) sustaining alertness ≥ 67% of the time. Although not perfect, these results suggest that the app can help mitigate the effects of sleep deprivation for three out of four participants, while suggesting the least amount of caffeine.

To provide some perspective on the benefits of the caffeine recommendations, we used the personalized models to compare the predicted alertness impairment during the 6-hour peak-alertness period with and without caffeine consumption (Supplementary Figure S2 in the Supplementary Material). Based on these predictions, caffeine consumption reduced alertness impairment from 14 milliseconds for participant #6 to 150 milliseconds for participant #21, with an average reduction of 55 milliseconds (SD = 38 milliseconds) over the 16 participants who consumed caffeine.

Examination of the performance of the personalized prediction models identified two situations where the *2B-Alert* app did not learn the participants' sleep-loss phenotype accurately. The first situation occurred when two participants (#3 and #4) initially displayed a resilient phenotype during the first night of TSD, when the app was learning their responses, and then subsequently revealed a vulnerable phenotype during the second night, when the app predicted they would show resilience (Figure 3). In this case, while the app customized the models to reflect the resilient phenotype displayed during the learning period (RMSEs of 20 milliseconds and 15 milliseconds for participants #3 and #4, respectively; Table 1), it could not anticipate that the participants had not fully realized their sleep-loss phenotype until the second night of sleep deprivation. Hence, for individuals who only reveal their phenotype to sleep loss after more stressful conditions, the app requires additional data to capture their response. For example, for participants #3 and #4, the RMSEs for the prediction period decreased from 115 to 74 milliseconds and from 61 to 45 milliseconds, respectively, when we used the entire set of PVT data to customize their models.

The second situation occurred when participants performed PVTs during the home phase with less diligence than during the laboratory phase. In this case, the consistently slow PVT results during the home phase caused the app to learn a response that seemed more vulnerable than it actually was. Then, when the PVT results improved during the laboratory phase, the app had to gradually re-learn the participant's actual response. However, because the app had diverged to a considerably different phenotype, the 12 PVTs collected during the laboratory learning period were not sufficient for the app to reverse course and capture the participant's true response to sleep loss. This was the case for participant #19, resulting in a RMSE of 90 milliseconds at the end of the learning period (Table 1). If the app had used only the 12 PVTs collected during the first 36 hours of TSD to personalize the model, then the RMSE for the prediction period would have been 37 milliseconds (Supplementary Table S1 in Supplementary Material), a reduction of 53 milliseconds. In sharp contrast, for the remaining 20 participants, the average absolute difference between the RMSEs of the predictions obtained with and without using the home-phase PVT data was only 6 milliseconds (SD = 5 milliseconds).

For participant #19, the mean RT during the home phase was on average 23 milliseconds slower than during the baseline laboratory phase (i.e. the first 16 hours of the challenge; Supplementary Table S1). However, this difference alone does not completely capture the participant's diligence in each phase, because other factors that affect PVT performance, e.g. time of day, sleep debt, and caffeine consumption, also varied between the two phases of the study. Interestingly, eight additional participants (#2, #5, #7, #9, #11, #14, #16, and #17) had similar or larger average differences between the home and baseline laboratory phases (Supplementary Table S1), however, the app was able to re-learn these participants with the first 12 PVTs of the laboratory phase because their home-phase data had led to smaller deviations from the participants' true phenotype than for participant #19. To quantify these deviations, we computed the RMSE between the predictions of the personalized models obtained at the end of the *home phase* and the PVT measurements for the prediction period. At the end of the home phase, the model for participant #19 yielded a RMSE of 138 milliseconds, compared to 42 to 86 milliseconds for these eight participants (Supplementary Table S1). These results highlight the need for quality PVT data for the app to properly learn an individual and make accurate personalized predictions, the ability of the app to tolerate some level of low-quality data, and the challenge in assessing PVT data quality in real time.

We also investigated whether *2B-Alert* was able to learn the participants' phenotype without using data collected during sleep deprivation. Specifically, we compared personalized models obtained using *only* the PVT data collected during the home phase with personalized models obtained at the end of the learning period, which included data collected during the home phase plus the first 36 hours of the sleep-deprivation challenge. For the prediction period, the former models yielded an average RMSE of

69 milliseconds (SD = 22 milliseconds) and an average percentage of PVT data within the 95% PIs of 50% (SD = 26%; Supplementary Table S1 in Supplementary Material), compared to an average RMSE of 50 milliseconds (SD = 24 milliseconds) and an average percentage of PVT data within the 95% PIs of 80% (SD = 22%) for the latter models (Table 1). These results highlight the importance of the information contained in the data collected during sleep-loss conditions to learn an individual's phenotype.

The effectiveness of caffeine as a sleep-loss countermeasure has been demonstrated in multiple controlled laboratory studies [20–25]. However, it is difficult to adapt the lessons learned in these studies to other conditions because of the large variations in sleep and work schedules between individuals and because of the large between-subject variability in the response to sleep loss. In addition, caffeine guidelines based on a "one-size-fits-all" approach, such as those of the US Army [49], only provide a partial solution. For example, for the current study's scenario of sustained operations (i.e. TSD), the US Army's guidelines recommend a total of 600 mg of caffeine for each participant, regardless of their actual requirements. In contrast, *2B-Alert* recommended no more than 300 mg (a 50% reduction) to 13 of the 21 participants, and 600 mg to only three participants (Table 1). Overall, the app's recommendations allowed participants to sustain the desired alertness 80% of the time while consuming 53% less caffeine (281 mg) compared to the US Army's guidelines.

This work has limitations. The core predictive model of the *2B-Alert* app, the UMP, predicts the time course of alertness as determined by the mean RT in the PVT. Thus, the personalized recommendations may not necessarily maximize the beneficial effects of caffeine for other cognitive functions. Moreover, the app does not account for an individual's sensitivity to caffeine or the development of caffeine tolerance. However, experimental data suggest that the most important determinant of the effect of caffeine on alertness is the individual's impairment level [47, 48], which is used in our model. Nonetheless, this limitation can be overcome by individualizing the parameters of the caffeine component of the model and updating them when caffeine is being consumed over prolonged periods. Another limitation is that the app does not have the ability to automatically determine in real time when it has fully learned the sleep-loss response of an individual. To address this limitation, we are currently investigating algorithms to add this important feature to the app. As discussed above, like any other data-driven model, our approach is also limited by the quality of the information available in the PVT data used to personalize the model, which depends on the individual's effort. This issue could be addressed by developing the capability to assess an individual's effort in performing PVTs during equivalent conditions (e.g. the same time of day, sleep debt, and caffeine consumption). Finally, the current model does not account for the potential disruptive effect of caffeine on sleep [50] for recommendations shortly before bedtime. This limitation can be alleviated by restricting caffeine intake for a number of hours prior to bedtime.

In summary, the *2B-Alert* app incorporates evidence-based predictive models developed and extensively validated over the last 15 years by our group at the US Army to provide personalized alertness predictions and personalized caffeine recommendations to mitigate the effects of sleep loss. For the first time, we demonstrate the ability of a smartphone app to automatically provide real-time, personalized caffeine recommendations for individuals under the same sleep-deprivation condition to achieve the same alertness level for a given time interval, regardless of their trait-like response to sleep loss. By offering predictions and recommendations tailored to the circumstances of each individual, the app can facilitate and enhance the self-management of fatigue, improving safety and productivity in both military and civilian settings.

Supplementary Material

Supplementary material is available at *SLEEP* online.

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. FGVL and JR receive royalties for the licensing of the *2B-Alert* technology to Integrated Safety Support. Non-financial 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. This paper has been approved for public release with unlimited distribution, and has been reviewed by the Walter Reed Army Institute of Research. There is no objection to its presentation and/or publication.

Author Contributions

FGVL, TJD, and JR designed the study. IA and WDSK performed the sleep study. FGVL performed the data analysis. FGVL and JR prepared 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. Medic G, *et al*. Short- and long-term health consequences of sleep disruption. *Nat Sci Sleep*. 2017;**9**:151–161. doi: 10.2147/NSS.S134864
2. Akerstedt T. Sleepiness as a consequence of shift work. *Sleep*. 1988;**11**(1):17–34. doi: 10.1093/sleep/11.1.17
3. Lim J, *et al*. A meta-analysis of the impact of short-term sleep deprivation on cognitive variables. *Psychol Bull*. 2010;**136**(3):375–389. doi: 10.1037/a0018883
4. Sallinen M, *et al*. Shift work, sleep, and sleepiness - differences between shift schedules and systems. *Scand J Work Environ Health*. 2010;**36**(2):121–133. doi: 10.5271/sjweh.2900
5. Wright KP, Jr, *et al*. Shift work and the assessment and management of shift work disorder (SWD). *Sleep Med Rev*. 2013;**17**(1):41–54. doi: 10.1016/j.smrv.2012.02.002

6. Hirshkowitz M, *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
7. Uehli K, *et al*. Sleep problems and work injuries: a systematic review and meta-analysis. *Sleep Med Rev*. 2014;**18**(1):61–73. doi: 10.1016/j.smrv.2013.01.004
8. Liu Y, *et al*. Prevalence of healthy sleep duration among adults — United States, 2014. *Morb Mortal Wkly Rep*. 2016;**65**:137–141. doi: 10.15585/mmwr.mm6506a1
9. Troxel WM, *et al*. *Sleep in the Military: Promoting Healthy Sleep Among U.S. Servicemembers*. Santa Monica, CA: RAND Corporation; 2015.
10. Hafner M, *et al*. *Why Sleep Matters — The Economic Costs of Insufficient Sleep: A Cross-Country Comparative Analysis*. Santa Monica, CA: RAND Corporation; 2016.
11. McCusker RR, *et al*. Caffeine content of energy drinks, carbonated sodas, and other beverages. *J Anal Toxicol*. 2006;**30**(2):112–114. doi: 10.1093/jat/30.2.112
12. Mitchell DC, *et al*. Beverage caffeine intakes in the U.S. *Food Chem Toxicol*. 2014;**63**:136–142.
13. Mahoney CR, *et al*. Intake of caffeine from all sources and reasons for use by college students. *Clin Nutr*. 2019;**38**(2):668–675. doi: 10.1016/j.clnu.2018.04.004
14. Roehrs T, *et al*. sleep and daytime sleepiness. *Sleep Med Rev*. 2008;**12**(2):153–162.
15. 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. *Morb Mortal Wkly Rep*. 2012;**61**(44):895–898.
16. Toblin RL, *et al*. 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.
17. Nawrot P, *et al*. Effects of caffeine on human health. *Food Addit Contam*. 2003;**20**(1):1–30. doi: 10.1080/0265203021000007840
18. Wikoff D, *et al*. Systematic review of the potential adverse effects of caffeine consumption in healthy adults, pregnant women, adolescents, and children. *Food Chem Toxicol*. 2017;**109**(Pt 1):585–648. doi:10.1016/j.fct.2017.04.002
19. Doty TJ, *et al*. Limited efficacy of caffeine and recovery costs during and following 5 days of chronic sleep restriction. *Sleep*. 2017;**40**(12). doi: 10.1093/sleep/zsx171
20. Killgore WD, *et al*. 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
21. McLellan TM, *et al*. Caffeine improves physical performance during 24 h of active wakefulness. *Aviat Space Environ Med*. 2004;**75**(8):666–672.
22. Wesensten NJ, *et al*. Performance and alertness effects of caffeine, dextroamphetamine, and modafinil during sleep deprivation. *J Sleep Res*. 2005;**14**(3):255–266. doi: 10.1111/j.1365-2869.2005.00468.x
23. McLellan TM, *et al*. Caffeine maintains vigilance and marksmanship in simulated urban operations with sleep deprivation. *Aviat Space Environ Med*. 2005;**76**(1):39–45.
24. Wesensten NJ, *et al*. Modafinil vs. caffeine: effects on fatigue during sleep deprivation. *Aviat Space Environ Med*. 2004;**75**(6):520–525.
25. Kamimori GH, *et al*. Caffeine improves reaction time, vigilance and logical reasoning during extended periods with restricted opportunities for sleep. *Psychopharmacology*. 2015;**232**(12):2031–2042. doi: 10.1007/s00213-014-3834-5
26. Rupp TL, *et al*. Trait-like vulnerability to total and partial sleep loss. *Sleep*. 2012;**35**(8):1163–1172. doi:10.5665/sleep.2010
27. Van Dongen HP, *et al*. Systematic interindividual differences in neurobehavioral impairment from sleep loss: evidence of trait-like differential vulnerability. *Sleep*. 2004;**27**(3):423–433.
28. Rajdev P, *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
29. Ramakrishnan S, *et al*. 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
30. Ramakrishnan S, *et al*. A unified model of performance for predicting the effects of sleep and caffeine. *Sleep*. 2016;**39**(10):1827–1841. doi: 10.5665/sleep.6164
31. Liu J, *et al*. Real-time individualization of the unified model of performance. *J Sleep Res*. 2017;**26**(6):820–831. doi: 10.1111/jsr.12535
32. Reifman 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
33. Vital-Lopez FG, *et al*. Caffeine dosing strategies to optimize alertness during sleep loss. *J Sleep Res*. 2018;**27**(5):e12711. doi: 10.1111/jsr.12711
34. Borbély AA. A two process model of sleep regulation. *Hum Neurobiol*. 1982;**1**(3):195–204.
35. Priezjev NV, *et al*. Assessment of the unified model of performance: accuracy of group-average and individualized alertness predictions. *J Sleep Res*. 2022;**32**:e13626. doi: 10.1111/jsr.13626
36. Institute of Medicine (U.S.) Committee on Military Nutrition Research. *Caffeine for the Sustainment of Mental Task Performance Formulations for Military Operations*. Washington, DC: The National Academies Press; 2001.
37. Thorne DR, *et al*. The Walter Reed palm-held psychomotor vigilance test. *Behav Res Methods*. 2005;**37**(1):111–118. doi: 10.3758/bf03206404
38. Dawson D, *et al*. Fatigue, alcohol and performance impairment. *Nature*. 1997;**388**(6639):235.
39. Williamson AM, *et al*. 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
40. Reifman J, *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
41. Borkenstein RF, *et al*. The role of the drinking driver in traffic accidents (The Grand Rapids Study). *Blutalkohol*. 1974;**11**(suppl 1):1–131.
42. Martin TL, *et al*. A review of alcohol-impaired driving: the role of blood alcohol concentration and complexity of the driving task. *J Forensic Sci*. 2013;**58**(5):1238–1250. doi:10.1111/1556-4029.12227
43. Kamimori GH, *et al*. The rate of absorption and relative bioavailability of caffeine administered in chewing gum versus capsules to normal healthy volunteers. *Int J Pharm*. 2002;**234**(1-2):159–167. doi: 10.1016/s0378-5173(01)00958-9
44. Morris C, *et al*. Caffeine release and absorption from caffeinated gums. *Food Funct*. 2019;**10**(4):1792–1796. doi: 10.1039/c9fo00431a
45. Killgore WDS, *et al*. Multiple caffeine doses maintain vigilance, attention, complex motor sequence expression, and manual dexterity during 77 hours of total sleep deprivation. *Neurobiol Sleep Circadian Rhythms*. 2020;**9**:100051. doi: 10.1016/j.nbscr.2020.100051

46. Khitrov MY, *et al*. PC-PVT: a platform for psychomotor vigilance task testing, analysis, and prediction. *Behav Res Methods*. 2014;**46**(1):140–147. doi: 10.3758/s13428-013-0339-9
47. Landolt HP, *et al*. Reduced neurobehavioral impairment from sleep deprivation in older adults: contribution of adenosinergic mechanisms. *Front Neurol*. 2012;**3**:62. doi:10.3389/fneur.2012.00062
48. Retey JV, *et al*. Adenosinergic mechanisms contribute to individual differences in sleep deprivation-induced changes in neurobehavioral function and brain rhythmic activity. *J Neurosci*. 2006;**26**(41):10472–10479. doi: 10.1523/JNEUROSCI.1538-06.2006
49. Army Techniques Publication 6-22.5. *A Leader's Guide to Soldier Health and Fitness*. Washington, DC: Headquarters, Department of the Army; 2016.
50. Drake C, *et al*. Caffeine effects on sleep taken 0, 3, or 6 hours before going to bed. *J Clin Sleep Med*. 2013;**9**(11):1195–1200. doi: 10.5664/jcsm.3170