REGULAR RESEARCH PAPER

Journal of Sleep Research ESRS

Caffeine dosing strategies to optimize alertness during sleep loss

Francisco G. Vital-Lopez[1] | Sridhar Ramakrishnan[1] | Tracy J. Doty[2] |
Thomas J. Balkin[2] | Jaques Reifman[1] iD

[1]Department of Defense Biotechnology
High Performance Computing Software
Applications Institute, Telemedicine and
Advanced Technology Research Center,
U.S. Army Medical Research and Materiel
Command, Fort Detrick, Maryland,

[2]Behavioral Biology Branch, Walter Reed
Army Institute of Research, Silver Spring,
Maryland

Correspondence
Jaques Reifman, Department of Defense
Biotechnology High Performance Computing
Software Applications Institute, Telemedicine
and Advanced Technology Research Center,
U.S. Army Medical Research and Materiel
Command, ATTN: MCMR-TT, 504 Scott
Street, Fort Detrick, MD 21702.
Email: jaques.reifman.civ@mail.mil

Funding information
Military Operational Medicine Program Area
Directorate of the U.S. Army Medical
Research and Materiel Command, Ft.
Detrick, MD, and by the U.S. Department of
Defense (DoD) Medical Research and
Development Program, Grant/Award
Number: DMRDP_13200

Summary

Sleep loss, which affects about one-third of the US population, can severely impair physical and neurobehavioural performance. Although caffeine, the most widely used stimulant in the world, can mitigate these effects, currently there are no tools to guide the timing and amount of caffeine consumption to optimize its benefits. In this work, we provide an optimization algorithm, suited for mobile computing platforms, to determine when and how much caffeine to consume, so as to safely maximize neurobehavioural performance at the desired time of the day, under any sleep-loss condition. The algorithm is based on our previously validated Unified Model of Performance, which predicts the effect of caffeine consumption on a psychomotor vigilance task. We assessed the algorithm by comparing the caffeine-dosing strategies (timing and amount) it identified with the dosing strategies used in four experimental studies, involving total and partial sleep loss. Through computer simulations, we showed that the algorithm yielded caffeine-dosing strategies that enhanced performance of the predicted psychomotor vigilance task by up to 64% while using the same total amount of caffeine as in the original studies. In addition, the algorithm identified strategies that resulted in equivalent performance to that in the experimental studies while reducing caffeine consumption by up to 65%. Our work provides the first quantitative caffeine optimization tool for designing effective strategies to maximize neurobehavioural performance and to avoid excessive caffeine consumption during any arbitrary sleep-loss condition.

KEYWORDS
alertness, fatigue, mathematical model, response time, vigilance

1 | INTRODUCTION

Sleep loss, which is a common stressor for both civilians and military personnel, can severely impair cognitive and physical performance, and thereby diminish productivity and compromise safety. Several studies have demonstrated that, when safely used, caffeine can help to sustain cognitive performance during prolonged periods of restricted sleep (Doty et al., 2017; Kamimori et al., 2015; Killgore et al., 2008; Mclellan, Bell, & Kamimori, 2004; Mclellan et al., 2005; Wesensten, Killgore, & Balkin, 2005). However, these investigations offer caffeine countermeasure guidance that is study-specific, and which cannot be readily adaptable to any arbitrary sleep-loss condition. Providing a foundation for addressing this need, our group has previously developed and validated a mathematical model, the

unified model of performance (UMP), which can predict the effects of sleep loss and caffeine, as a function of time of day, on objective measures of neurobehavioural performance (i.e. the psychomotor vigilance task, PVT) across a wide range of sleep–wake schedules and caffeine doses (Ramakrishnan et al., 2013, 2014, 2016). More recently, we have built upon the UMP to develop the open-access Web tool *2B-Alert* (Reifman et al., 2016), a decision aid to help users design sleep studies and work schedules, and the smartphone *2B-Alert* app (Reifman et al., 2017) for real-time, individualized performance prediction (Liu, Ramakrishnan, Laxminarayan, Balkin, & Reifman, 2017).

In this work, our goal was to develop a computational tool to provide, in real time, effective caffeine-dosing strategies for any arbitrary sleep-loss condition. Once incorporated into a mobile computing device, such a tool could provide customized caffeine-consumption guidance to, for example, sustain the attention of sleep-deprived military personnel. To this end, using the predictive ability of the UMP, we formulated an optimization problem to determine when and how much caffeine to consume, so as to safely maximize neurobehavioural performance at the desired time of the day for the desired duration. To solve this problem, we developed an efficient optimization algorithm that was able to find near-optimal solutions in real time. We assessed the optimization algorithm by comparing the effects of its predicted caffeine-dosing (timing and amount) strategies with those obtained in four experimental studies previously used to validate the UMP (Ramakrishnan et al., 2016). In particular, we obtained caffeine-dosing strategies that enhanced PVT performance while using the same total amount of caffeine as in the original studies, and strategies that yielded equivalent levels of performance as in the original studies while reducing caffeine consumption.

2 | METHODS

2.1 | The unified model of performance

The UMP has two components. The first, based on Borbély's two-process model (Borbely, 1982), describes performance as a function of the circadian cycle and a homeostatic process. The second is a pharmacokinetic and pharmacodynamic model, which estimates the caffeine level in the blood and predicts the duration and magnitude of the effect of caffeine intake on neurobehavioural performance. For a given sleep–wake schedule and caffeine consumption strategy, which constitute the inputs to the model, the UMP predicts the PVT mean response time (RT) for an 'average' individual. We refer the reader to Ramakrishnan et al. (2016) for detailed descriptions of the UMP, the parameter estimation process and model validation. We have provided the UMP equations and parameter values in Section I of the Supporting Information.

2.2 | Optimization problem

Our goal was to find a caffeine-dosing strategy that would minimize neurobehavioural performance impairment based on the PVT for a given sleep–wake schedule. To this end, we sought to minimize the objective function Z (Equation 1, Table 1), which considers both the area under the UMP-predicted PVT mean RT curve (AUC_C) that is above the baseline, and the worst performance (WP_C) (i.e. the difference between the peak of the mean RT curve and the baseline) (Figure 1). As the baseline mean RT, we used the highest predicted value of mean RT when an average individual has no sleep debt, wakes up at 07:00 and is awake for 16 hr. We normalized AUC_C and WP_C by the corresponding values for the predicted mean RT curve without caffeine consumption, AUC_{NC} and WP_{NC}, respectively. In addition, we included a penalty term in Z to limit the accumulation of caffeine in the blood [$C(t_i, D_i)$], which could result in unsafe consumption (Killgore et al., 2008). This term penalizes Z when the maximum level of caffeine in the blood is higher than the maximum level (C_{max}) achieved by a single 400-mg dose (Institute of Medicine, 2001). (Note that the value of C_{max} can be readily changed in the algorithm.) Hence, without considering the penalty term, Z varies from 0 (for a strategy that consistently maintains the mean RT below the baseline) to 100 (for a strategy that is no better than using no caffeine). Therefore, the smaller the value of Z, the better is the dosing strategy.

The optimization variables were the time (t_i) and caffeine amount (D_i) of dose i, with $i = 1, 2, \ldots, n$ (the number of doses used in the strategy). To obtain solutions in a practical computational time, we imposed the following constraints on the optimization variables (Table 1): (i) D_i was restricted to 100, 200 or 300 mg of caffeine (Equation 2); (ii) the dosing time was restricted to occur on the hour (e.g. 18:00, 22:00 and 24:00) (Equation 3); and (iii) the minimum time between doses was 2 hr (Equation 4). Equation 4 excluded strategies that prescribe doses too often, which could be too burdensome to follow in practice. We also included two additional constraints to obtain solutions with a desired total amount of caffeine (Equation 5) or number of doses (Equation 6). The objective function Z in Equation 1 is a nonlinear function of t_i and D_i, which in turn are discrete variables. Thus, the optimization problem is a mixed integer nonlinear problem (MINLP).

2.3 | Tabu search algorithm

Standard algorithms for solving MINLPs (i.e. branch-and-bound and simulated annealing) were not able to solve the optimization problem in Table 1 in a reasonable computational time (see Section II of the Supporting Information). To address this issue, we sought an approximate solution using the tabu search algorithm (Glover, 1986). The tabu search algorithm finds an optimal solution by evaluating the objective function Z at a number of test points in the neighbourhood of a current solution and moving to the next tested point with the best solution. The key feature of the algorithm is how to select the next testing points. Although the selection of a large set of testing points may produce better results, this comes at the cost of higher computational times.

We developed the following scheme to implement the algorithm (Figure 2a). First, for a given current solution, we divided the entire sleep–wake schedule into periods of time between doses. For each

TABLE 1 Optimization problem to find a caffeine-dosing strategy that minimizes neurobehavioural performance impairment

Objective function:

$$\min_{t_i,D_i} Z = 50\frac{AUC_C(t_i,D_i)}{AUC_{NC}} + 50\frac{WP_C(t_i,D_i)}{WP_{NC}} + 250\max\{C(t_i,D_i) - C_{max}, 0\} \qquad (1)$$

where Z denotes the objective function that we wish to minimize. The optimization variables t_i and D_i represent the time (in hours after the first wake-up time in the schedule) and the caffeine amount (in mg) of dose i, respectively, with $i = 1, 2, \ldots, n$ (the number of doses). AUC denotes the area under the PVT mean response time (RT) curve above the baseline, and WP the difference between the peak of the mean RT curve and the baseline. $C(t_i, D_i)$ denotes the level of caffeine in the blood. The subscripts C and NC denote caffeine and no caffeine, respectively. The last term penalizes the objective function when the maximum value of $C(t_i, D_i)$ is higher than the maximum caffeine level achieved by a single dose of 400 mg, denoted by C_{max}.

Constraints:
Allowed doses:

$$D_i \in \{100, 200, 300\} \qquad (2)$$

Allowed dosing times:

$$t_i \in \{1, 2, \ldots, t_f\} \qquad (3)$$

where t_f is the total number of hours in the sleep–wake schedule.

Time between doses:

$$t_{i+1} - t_i \geq 2, \text{ for } i \in \{1, 2, \ldots, n-1\} \qquad (4)$$

Total amount of caffeine:

$$\sum_i D_i \leq D_T \qquad (5)$$

where D_T denotes the maximum total caffeine intake. Note that, in this constraint, the inequality symbol can be replaced by the equality symbol.

Total number of doses:

$$n \leq n_T \qquad (6)$$

where n_T denotes the maximum number of doses. Note that, in this constraint, the inequality symbol can be replaced by the equality symbol.

period, we computed the value of Z using the UMP, and selected the period with the worst value (i.e. largest) (Figure 2b). Then, if the sum of the first and second terms in Equation 1 was larger than the third term, we evaluated eight test points to reduce the performance impairment for the period; otherwise, we evaluated four test points to reduce the caffeine level (Figure 2c). Then, we selected the point with the smallest overall Z (i.e. for the entire sleep–wake schedule) as the current solution (Figure 2D) and repeated the process. We marked all of the points that had been selected as the current solution as 'tabu' to avoid these points in subsequent iterations. The algorithm stopped when either a prespecified number of iterations was reached or no more feasible solutions were left. In this work, we selected the number of iterations to be 2000, as a compromise between the optimality of the solution and the computational time.

2.4 | Laboratory and field studies

Figure 3 schematically shows the sleep–wake schedules and the caffeine-dosing strategies for four studies used by Ramakrishnan et al. (2016) to validate the UMP. Here, we revisited these studies to demonstrate the benefit of the proposed algorithm. The studies

investigated the effect of caffeine on group-average performance during chronic sleep restriction (CSR; Study 1 [Doty et al., 2017;]), total sleep deprivation (TSD; Study 4 [Mclellan et al., 2004;]) or a combination of both (Studies 2 and 3 [Kamimori et al., 2015; Mclellan et al., 2005;]). For additional details on these studies, we refer the reader to Ramakrishnan et al. (2016).

3 | RESULTS

3.1 | Two optimization problems: enhancing neurobehavioural performance and reducing caffeine consumption

For each of the sleep–wake schedules used in Studies 1 to 4, we obtained dosing strategies that minimized performance impairment (Z) for a range of values of total amount of caffeine (i.e. we solved the MINLP in Table 1 for different values of D_T in Equation 5). Figure 4 shows the UMP-predicted performance impairment for both the optimal (blue circles) and the original dosing strategies (orange diamonds). From these results, we focused on two types of solutions. The first involved strategies for enhancing neurobehavioural performance using the same total amount of caffeine as in the original studies (Figure 4,
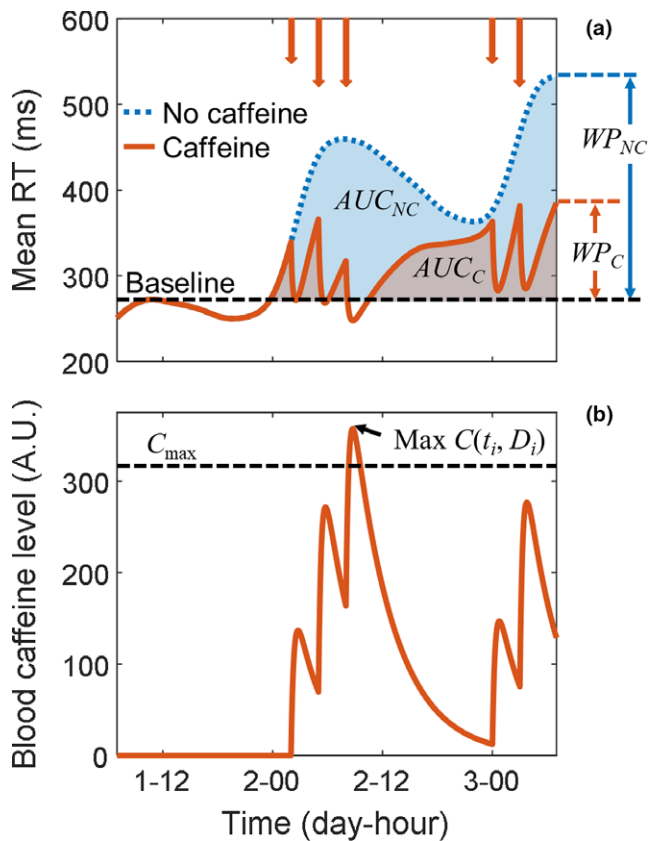
FIGURE 1 Quantities used to define the objective function for quantifying the benefits of different caffeine-dosing strategies. (a) The graph shows the psychomotor vigilance task mean reaction time (RT) predicted by the unified model of performance (UMP) for a 48-hr total sleep deprivation challenge, for the cases of no caffeine consumption (subscript NC) and an arbitrary caffeine-dosing strategy (subscript C). AUC denotes the area under the curve above the baseline, and WP the difference between the peak of the mean RT curve and the baseline. The orange arrows at the top of the plot indicate the time and amount of the caffeine doses for the arbitrary dosing strategy. The small and large arrows represent 200- and 300-mg doses, respectively. (b) UMP-predicted caffeine level in the blood for the same arbitrary strategy. C_{max} represents the maximum acceptable caffeine level in the blood for a single 400-mg dose, and Max $C(t_i, D_i)$ denotes the predicted maximum caffeine level achieved by the strategy over the entire period. The objective function Z (Table 1, Equation 1) attempts to simultaneously minimize the predicted area (AUC_C) and peak (WP_C), while constraining Max $C(t_i, D_i)$ below C_{max}. The predicted AUC_{NC} and WP_{NC} for the no-caffeine case are used to normalize Z

red arrows). The second type involved strategies that attempt to reduce caffeine consumption while achieving at least the same benefit as the original studies (i.e. a value of Z equal to or smaller than that of the original study) (Figure 4, green arrows).

3.2 | Enhancing neurobehavioural performance

Figure 5 shows the experimental mean RT data (orange dots with standard error bars) for each study (Doty et al., 2017; Kamimori et al., 2015; Mclellan et al., 2004, 2005), as well as the UMP-

predicted mean RT profiles for the original studies (dashed orange lines) and the optimal dosing strategies (blue lines) that attempted to enhance neurobehavioural performance, using the same total amount of caffeine as in the original studies (Figure 4, red arrows). Overall, the UMP satisfactorily predicted the mean RT for the original studies (i.e. the relative root mean squared error for the four studies ranged from 6% to 17%) (Ramakrishnan et al., 2016). Table 2 summarizes the changes in neurobehavioural performance for the optimal strategies. The optimal strategies using the same amount of caffeine showed substantially better performance compared with the original dosing strategies.

For Study 1, the major portion of the 64% improvement (Table 2, column 5) was a result of the reduction of the worst peak predicted daily for the original dosing strategy, which prescribed the same total amount of caffeine (400 mg) at the same time each day (Figure 5, orange arrows). In contrast, the optimal strategy prescribed more caffeine on later days (with the exception of the last day), owing to increasing sleep pressure, and allocated caffeine at the end of each of the first 4 days of CSR (Figure 5, blue arrows). Moreover, the optimal strategy did not prescribe caffeine early on the first day of CSR because performance impairment was mitigated by sleep banking (subjects spent 10 hr in bed on five previous nights) (Doty et al., 2017).

In Study 2, the optimal strategy improved the effect of caffeine by 48% (Table 2). In contrast to the original study, which prescribed the same total amount of caffeine during each of the three periods of wakefulness (Figure 5, orange arrows), the optimal strategy allocated more caffeine during longer periods of wakefulness (Figure 5, blue arrows; 900, 800 and 700 mg for the first, second and third periods, respectively). Also, we predicted that the original study prescribed the first dose earlier than needed in the second and third periods, resulting in large performance impairment at the end of each period. In the optimal strategy, the postponement of the first dose in the second and third periods reduced and balanced the performance impairment across the periods.

In Studies 3 and 4, the original countermeasures reduced performance impairment more than did the optimal strategies near the middle of the TSD challenge (i.e. the predicted mean RT for the original studies was below the mean RT of the optimal strategies; Figure 5). However, performance impairment was substantially greater for the original studies than for the optimal strategies during the last 6 hr of the TSD challenges. Overall, the optimal strategies improved performance by 16% and 41% compared with the original countermeasures for Studies 3 and 4, respectively (Table 2).

3.3 | Reducing caffeine consumption

Figure 6 shows the predicted mean RT profiles for the original studies (dashed orange line) and the optimal strategies (blue lines) that attempted to reduce caffeine consumption, while achieving at least the same benefit as the original studies (Figure 4, green arrows). Table 2 (columns 6 and 7) shows the changes in caffeine consumption for the optimal dosing strategies. In general, the optimal
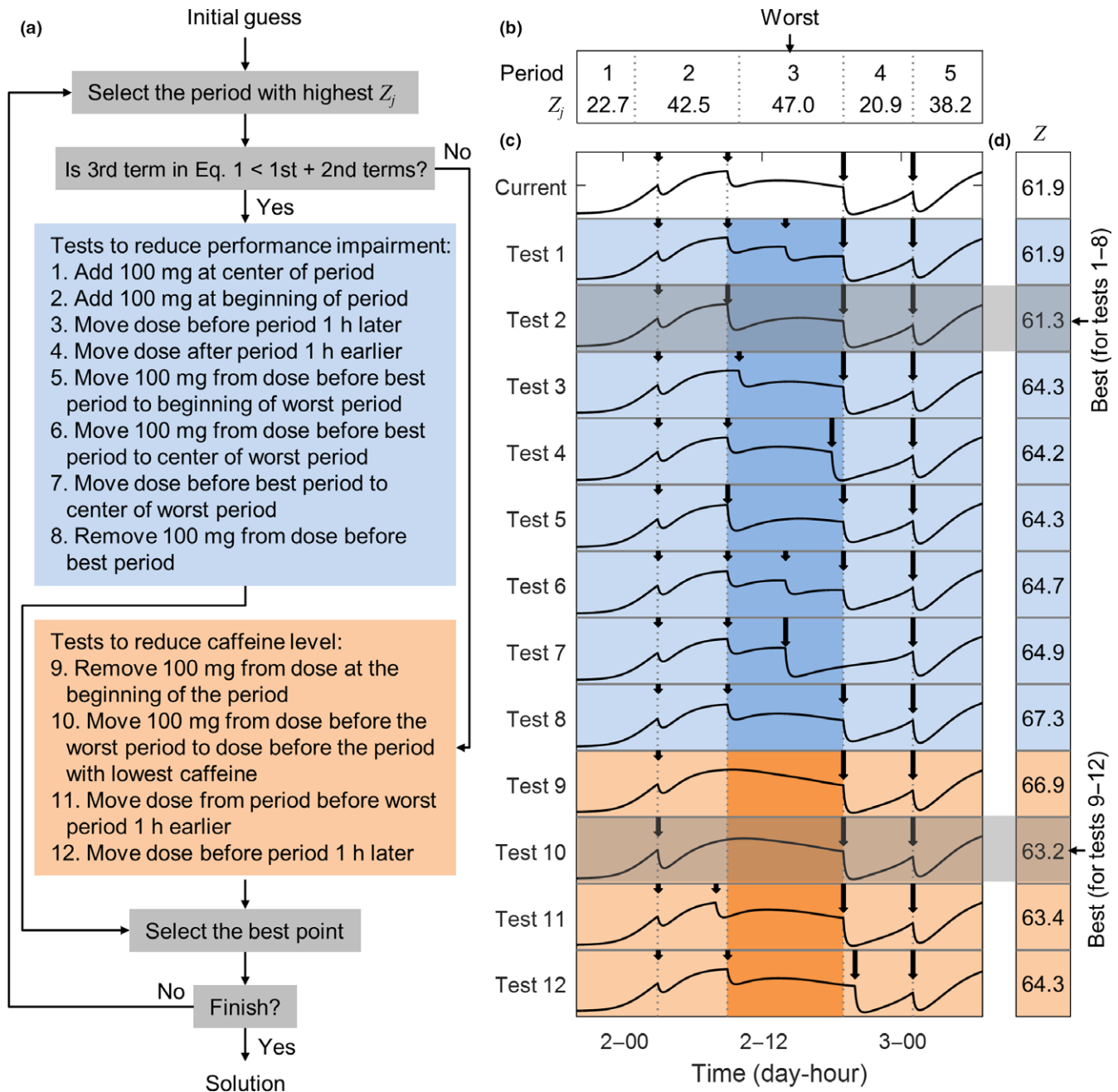
FIGURE 2 Tabu search algorithm. (a) Implementation of the tabu search algorithm. We used the unified model of performance (UMP) to compute the objective function value (Z_j) for each period j formed by the times between the doses of the current solution. (b) We used an arbitrary caffeine-dosing strategy as the current solution to illustrate the implementation of the algorithm. The diagram shows the periods between doses and the values of Z_j in the corresponding periods. Period 3 was the worst period ($Z_j = 47.0$). (c) Psychomotor vigilance task mean response time predicted by the UMP at the test points. The small, medium and large arrows at the top of each subplot represent 100-, 200- and 300-mg doses, respectively. The dotted lines indicate the dosing times for the current solution. We evaluated tests 1 to 8 (shaded in blue) when the third term in Equation 1 was smaller than the sum of the first and second terms (Table 1); otherwise, we evaluated tests 9 to 12 (shaded in orange). (d) Objective function values at the test points for the entire sleep–wake schedule. Highlighted in grey are the best tests for each of the two test sets

strategies required less caffeine consumption than the original studies.

For Study 1, the total amount of caffeine used in the optimal strategy was only 700 mg (Figure 6, blue arrows), which was 65% less than that used in the original study (Table 2, column 7), but still yielded a slightly better performance (Z: 61 versus 64) than the

original. In Study 2, the optimal strategy used 500 mg (21%) less caffeine than did the original study (Table 2, column 7), with the major caffeine savings occurring in the third wake period because it was the shortest period (Figure 6).

The optimal strategies for Studies 3 and 4 reduced the total amount of caffeine intake by 100 mg and 200 mg (17% and 33%),

6 of 10

Journal of
Sleep
Research

ESRS

VITAL-LOPEZ ET AL.
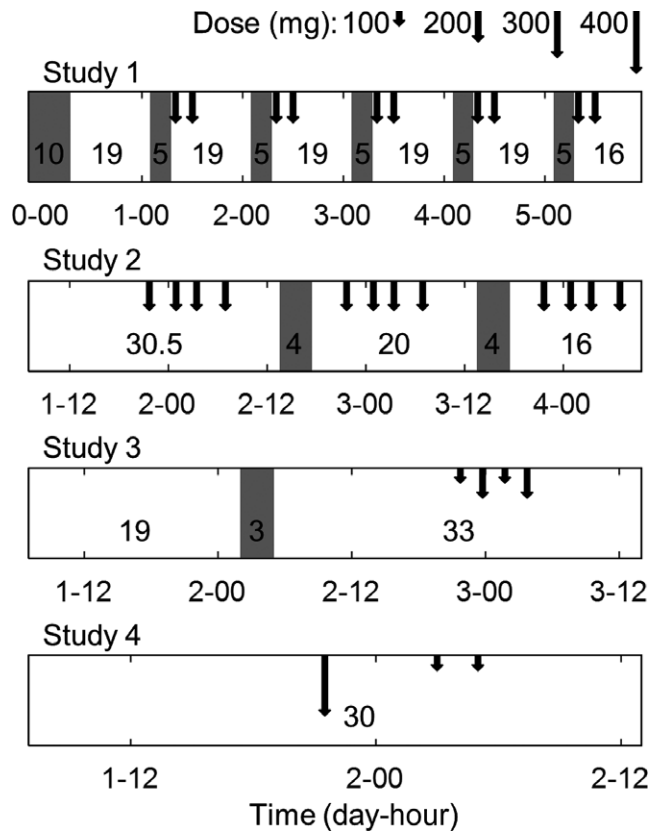
FIGURE 3 Schematic representations of caffeine dosing and sleep schedules for the four studies used to assess the benefits of the caffeine optimization algorithm. The grey and white areas represent time in bed and time awake, respectively, where the number in each area indicates the number of hours in that period. Each arrow at the top of each schedule indicates the time and amount of a dose

respectively. In these strategies, performance at the end of the period improved at the expense of performance at earlier times (Figure 6).
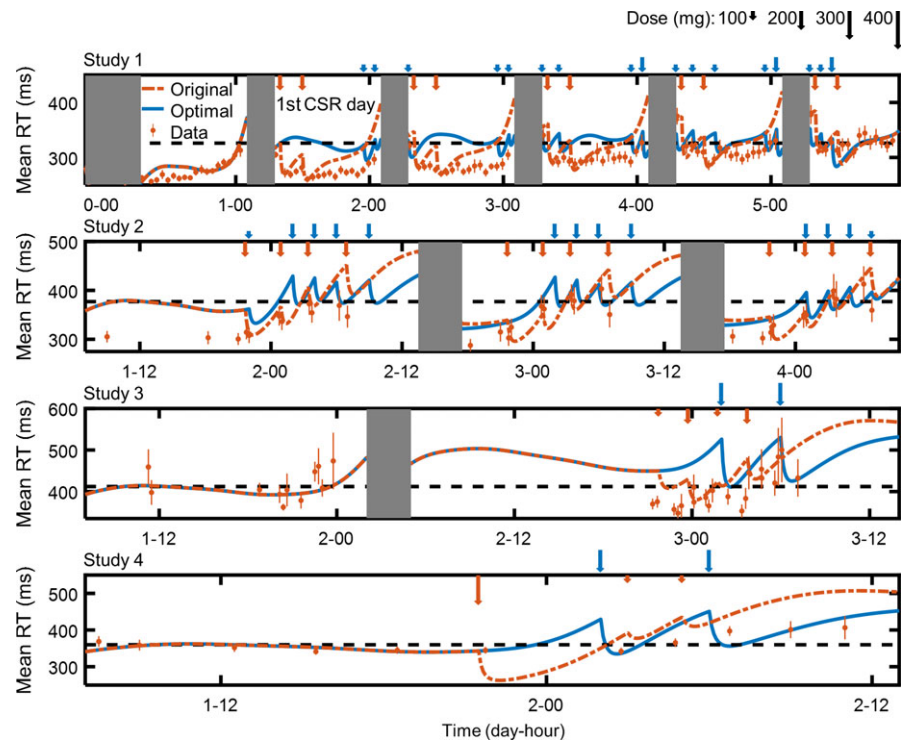
4 | DISCUSSION

Caffeine, if safely administered, is an effective countermeasure to mitigate impairment of alertness caused by sleep loss. This has been demonstrated in multiple laboratory and field studies for different sleep–wake schedules. However, to maximize its effectiveness, caffeine should be consumed at the right time and amount. Here, we developed an optimization algorithm to determine when and how much caffeine to consume so as to safely maximize alertness of a group of individuals for any situation. At the core of our algorithm is the UMP, a validated mathematical model that accurately predicts the effects of sleep–wake schedules and caffeine consumption (i.e. the model inputs) on neurobehavioural (PVT) performance. In conjunction with a new implementation of the tabu search algorithm, the UMP allowed for the identification of near-optimal caffeine-dosing strategies in a practical computational time (i.e. in seconds).
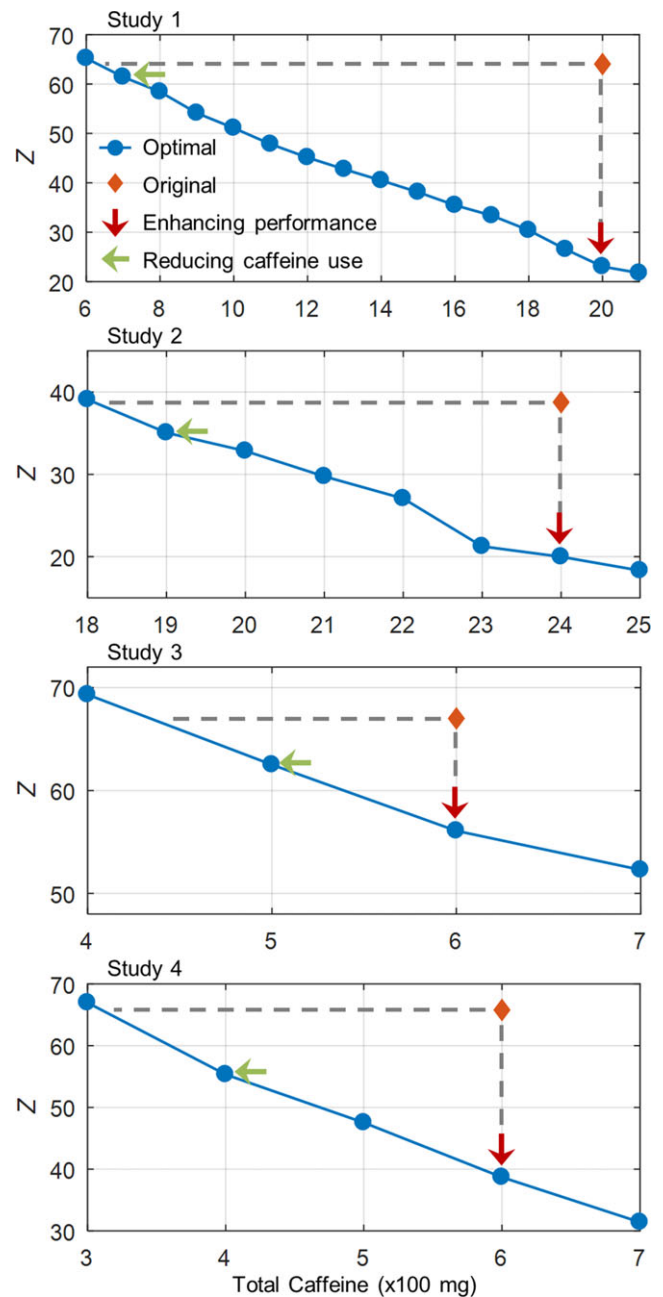
FIGURE 4 Objective function (Z) for optimal caffeine-dosing strategies and for the original study. The graphs show the value of Z for the optimal strategies using different amounts of total caffeine intake (blue circles) and for the original studies (orange diamonds). Red arrows indicate optimal strategies that enhanced neurobehavioural performance using the same total amount of caffeine as that in the original study. Green arrows indicate optimal strategies that reduced caffeine dosing, while achieving at least the same neurobehavioural performance as in the original studies

We used the algorithm to obtain optimal caffeine-dosing strategies for different sleep–wake schedules that included TSD, CSR and their combinations in four separate studies. For these studies, we found strategies that yielded up to 64% greater performance improvements than the original studies while using the same total

FIGURE 5 Optimal strategies to enhance neurobehavioural performance using the same total amount of caffeine as in the original studies. The dashed orange lines and the continuous blue lines represent the unified model of performance predictions of the psychomotor vigilance task (PVT) mean response time (RT) for the original study and optimal strategies, respectively. The orange dots and bars represent the experimental PVT mean RT data and one standard error, respectively. The orange and blue arrows at the top of each plot indicate the time and amount of the caffeine doses for the original study and optimal strategies, respectively. The horizontal dashed lines indicate the baseline and the grey vertical bars represent time in bed. CSR, chronic sleep restriction

TABLE 2 Effects of optimal dosing strategies to enhance neurobehavioural performance and reduce caffeine consumption

Study	Original		Enhanced performance[a]	Reduced caffeine[b]	
	Total caffeine, mg	Z	Z[c]	Z	Total caffeine, mg[d]
1	2,000	64	23 (64)	61	700 (65)
2	2,400	39	20 (48)	35	1,900 (21)
3	600	67	56 (16)	63	500 (17)
4	600	66	39 (41)	55	400 (33)

[a]The strategies used the same total amount of caffeine as did the original countermeasures.

[b]The strategies achieved benefits that were at least as good as, if not better than, those in the original studies.

[c]The numbers in parentheses indicate the percentage improvement in Z relative to the original studies.

[d]The numbers in parentheses indicate the percentage reduction in caffeine consumption relative to the original studies.

amount of caffeine (Table 2, columns 4 and 5). The results showed that the timing and amount of caffeine should be tailored to the particular situation to maximize its benefits (Figure 5). For example, in Study 1, the sleep–wake schedule was the same for 5 days of CSR (except the last day, when the wake period was shorter), gradually accumulating sleep debt that led to increasing performance impairment across the days of CSR. Accordingly, the optimal strategy allocated more caffeine to days with higher sleep pressure. This is in contrast to the original study, which repeated the same doses for each of the five CSR days (Figure 5, compare blue and orange arrows). Moreover, the timing of the doses in the original case (08:00 and 12:00) could not prevent the predicted performance

impairment for the last 2 hr prior to sleep of the first four CSR days. This impairment was mitigated in the optimal strategy by allocating two doses close to the end of each day.

Recently, Doty et al. (2017) found that consumption of high amounts of caffeine for several days of sleep restriction can impair the recovery of an individual. Hence, dosing strategies that reduce caffeine consumption can help to mitigate the negative effects of caffeine on recovery. Thus, we obtained dosing strategies that reduced caffeine consumption, while achieving at least the same level of neurobehavioural performance as the original studies. For the four studies, the optimal dosing strategies prescribed between 17% and 65% less caffeine than the original countermeasures.

One limitation of our optimization algorithm is that it does not guarantee the identification of global solutions. Nonetheless, as illustrated in Section II of the Supporting Information, the algorithm found strategies that were nearly as effective as those found by a standard optimization algorithm (i.e. simulated annealing), albeit faster by at least two orders of magnitude. This reduction in computational time enables practical use of the algorithm. Another limitation is that our algorithm does not account for the possibility that caffeine consumption before bedtime can reduce sleep quality in subjects with regular sleep–wake schedules (Drake, Roehrs, Shambroom, & Roth, 2013). Whether this has the same effect in sleep-deprived individuals remains unclear. Nonetheless, by including additional constraints to the optimization problem, we can avoid strategies that prescribe caffeine a number of hours before bedtime. For example, in Study 1, by restricting caffeine consumption in the last 6 hr of wakefulness, the optimal strategies were still able to reduce performance impairment by 26% using the same total amount of caffeine as the original countermeasure, and reduce caffeine consumption by
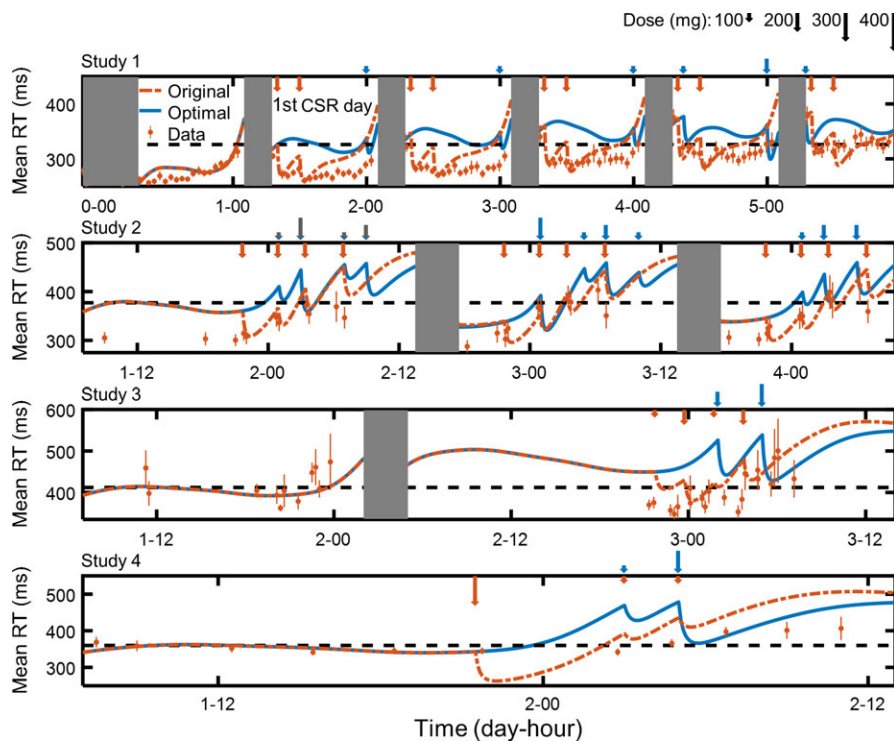
FIGURE 6 Optimal strategies to reduce caffeine consumption dosing while maintaining at least the same performance as in the original studies. The dashed orange lines and the continuous blue lines represent the unified model of performance predictions of the psychomotor vigilance task (PVT) mean response time (RT) for the original study and optimal strategies, respectively. The orange dots and bars represent the experimental PVT mean RT data and one standard error, respectively. The orange and blue arrows at the top of each plot indicate the time and amount of the caffeine doses for the original study and optimal strategies, respectively. The horizontal dashed lines indicate the baseline and the grey vertical bars represent time in bed. CSR, chronic sleep restriction

35% while achieving at least the same benefit as the original countermeasure.

It should also be noted that the UMP was developed to predict the effects of sleep loss and caffeine consumption on simple neurobehavioural tasks, such as the PVT. Because an individual's performance level in simple tasks may not reflect that individual's performance in other neurocognitive tasks (Rupp, Wesensten, & Balkin, 2012; Van Dongen, Baynard, Maislin, & Dinges, 2004), the computed caffeine strategies may be suboptimal for other tasks. Doty et al. (2017) also found that the benefits of caffeine decrease with accumulation of sleep debt (i.e. caffeine provides reduced benefits after 4 days of 5 hr of sleep per night). However, the UMP does not currently account for the effects of sleep debt on the benefits of caffeine. Consequently, our algorithm's strategies may potentially overestimate neurobehavioural performance or underestimate the amount of caffeine needed for long CSR scenarios. Moreover, the UMP does not consider individual differences in sensitivity, or the development of tolerance of caffeine, which could result in paradoxical effects. For example, the optimization algorithm could predict too much caffeine for a caffeine-sensitive individual, which could lead to extended sleep onset, reduced recovery sleep and increased caffeine consumption. In contrast, individuals with low sensitivity to caffeine may require considerably more caffeine than the amount prescribed by the optimal dosing strategy for an average individual.

To assess the sensitivity of the UMP predictions to variability in the model parameters, we carried out a sensitivity analysis by performing 10,000 simulations. In each simulation, we simultaneously selected different values for the 12 parameters in the model (i.e. by uniformly sampling from within two standard errors of the nominal values in Table S2), used those values in the model to predict the

PVT mean RT for the original caffeine strategy in Study 2, and computed the percentage of predictions after the first dose that fell within two standard errors of the experimental PVT mean RT (Kamimori et al., 2015). For the 10,000 simulations, this percentage was 66%. This means that, given the variability of the PVT data and the uncertainties in the model parameters, ~66% of the UMP predictions were statistically indistinguishable from the experimental data. This was only slightly less than the percentage when using the nominal parameter values, which was 75%. This result suggests that, although not perfect, the UMP is robust to uncertainties in the model parameters. Nonetheless, ultimately, to assess the effectiveness of caffeine-dosing strategies proposed by the optimization algorithm, we will need to carry out prospective experimental validation studies, where we compare and contrast different strategies.

Another limitation is that the parameters of the UMP were estimated to capture a 'group-average' response to sleep loss and caffeine consumption. However, there may be considerable individual variability in both the response to sleep loss (Van Dongen et al., 2004) and the restorative effects of caffeine (Ramakrishnan et al., 2014). Variation in the effect of caffeine is, in part, a result of genetic polymorphisms in the genes coding for the main caffeine-metabolizing enzyme, P-450, and the main caffeine targets, adenosine receptors A_1 and A_{2A} (Yang, Palmer, & De Wit, 2010). To assess how well a group-average model captures individual differences, we computed the root mean squared error (RMSE) between the model predictions and the measured mean RT data from each subject after caffeine consumption. For example, using the original caffeine strategy for the 10 subjects in Study 2, the average RMSE was 56 ms (range, 31 to 96 ms). In contrast, the RMSE between the group-average model predictions and the group-average data was 33 ms,

suggesting that our group-average model captures the mean alertness of the group better than it does that of each individual in the group.

To further assess the suitability of using the group-average model predictions for different subjects, we estimated how long the prediction error remained within a given threshold of the nominal parameter set predictions in the 10,000 simulations used for sensitivity analysis. For this purpose, we assumed that the 10,000 random parameter sets represented 10,000 individual subjects and computed the prediction error as the absolute difference between the mean RT predicted with the nominal parameter set and each of the random parameter sets. Then, for each simulation, we determined the time (after the first caffeine dose) for which the prediction error exceeded 25% of the mean RT predicted using the nominal parameters. For 53% of the cases, the predicted mean RT remained within 25% of the nominal mean RT predictions for the entire time (i.e. for the 59.8 hr from the first caffeine dose until the end of Study 2). For the remaining 47% of the cases, the average time to exceed 25% error was 13.6 hr (range, 5 min to 59.6 hr). In other words, for about half of the cases (representing 'average-like' subjects) the prediction error remained relatively small throughout the duration of the study, whereas for the other half (representing subjects highly vulnerable or resilient to sleep loss and/or highly sensitive to or tolerant of caffeine), on average, the error considerably increased after 13.6 hr. This result suggests that a group-average model cannot always be used to obtain optimal caffeine strategies at the individual level. Such inter-subject variability could be addressed in the future by coupling the caffeine optimization algorithm with an individualized prediction model (Liu et al., 2017) to provide tailored, subject-specific interventions.

In summary, we developed an optimization algorithm for designing safe and effective caffeine countermeasure strategies to mitigate performance impairment for arbitrary sleep-loss conditions. The unique capability of the proposed algorithm is that it combines a validated mathematical model with optimization methods to determine when and how much caffeine to consume to achieve peak performance at the most needed times. As a next step, we plan to incorporate this algorithm into the open access *2B-Alert* Web tool (Reifman et al., 2016), allowing for optimized caffeine prescription in the design of sleep studies and work schedules.

CONFLICT OF INTEREST

This was not an industry-supported study. The authors have indicated no financial 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 or of the U.S. DoD. This paper has been approved for public release with unlimited distribution.

AUTHOR CONTRIUTIONS

SR, JR and TJB conceived the research. FGV and SR developed the algorithms and performed the computations. TJB and TJD provided data for modelling and simulation. FGV and JR wrote the paper.

ORCID

Jaques Reifman http://orcid.org/0000-0001-7292-2029

REFERENCES

Borbely, A. A. (1982). A two process model of sleep regulation. *Human Neurobiology*, 1, 195–204.

Doty, T. J., So, C. J., Bergman, E. M., Trach, S. K., Ratcliffe, R. H., Yarnell, A. M., . . . Quartana, P. J. (2017). Limited efficacy of caffeine and recovery costs during and following 5 days of chronic sleep restriction: Caffeine and Sleep Restriction. *Sleep*, 40, zsx171.

Drake, C., Roehrs, T., Shambroom, J., & Roth, T. (2013). Caffeine effects on sleep taken 0, 3, or 6 hours before going to bed. *Journal of Clinical Sleep Medicine*, 9, 1195–1200.

Glover, F. (1986). Future paths for integer programming and links to artificial intelligence. *Computers & Operations Research*, 13, 533–549. https://doi.org/10.1016/0305-0548(86)90048-1

Institute of Medicine (2001). *Caffeine for the Sustainment of Mental Task Performance: Formulations for Military Operations*. Washington, DC: The National Academies Press.

Kamimori, G. H., Mclellan, T. M., Tate, C. M., Voss, D. M., Niro, P., & Lieberman, H. R. (2015). Caffeine improves reaction time, vigilance and logical reasoning during extended periods with restricted opportunities for sleep. *Psychopharmacology (Berl.)*, 232, 2031–2042. https://doi.org/10.1007/s00213-014-3834-5

Killgore, W. D., Rupp, T. L., Grugle, N. L., Reichardt, R. M., Lipizzi, E. L., & Balkin, T. J. (2008). Effects of dextroamphetamine, caffeine and modafinil on psychomotor vigilance test performance after 44 h of continuous wakefulness. *Journal of Sleep Research*, 17, 309–321. https://doi.org/10.1111/j.1365-2869.2008.00654.x

Liu, J., Ramakrishnan, S., Laxminarayan, S., Balkin, T. J., & Reifman, J. (2017). Real-time individualization of the unified model of performance. *Journal of Sleep Research*, 26, 820–831. https://doi.org/10.1111/jsr.12535

Mclellan, T. M., Bell, D. G., & Kamimori, G. H. (2004). Caffeine improves physical performance during 24 h of active wakefulness. *Aviation, Space and Environmental Medicine*, 75, 666–672.

Mclellan, T. M., Kamimori, G. H., Bell, D. G., Smith, I. F., Johnson, D., & Belenky, G. (2005). Caffeine maintains vigilance and marksmanship in simulated urban operations with sleep deprivation. *Aviation, Space and Environmental Medicine*, 76, 39–45.

Ramakrishnan, S., Laxminarayan, S., Wesensten, N. J., Kamimori, G. H., Balkin, T. J., & Reifman, J. (2014). Dose-dependent model of caffeine effects on human vigilance during total sleep deprivation. *Journal of Theoretical Biology*, 358, 11–24. https://doi.org/10.1016/j.jtbi.2014.05.017

Ramakrishnan, S., Rajaraman, S., Laxminarayan, S., Wesensten, NJ, Kamimori, GH, Balkin, TJ, Reifman, J. (2013). A biomathematical model of the restoring effects of caffeine on cognitive performance during

sleep deprivation. *Journal of Theoretical Biology*, *319*, 23–33. https://doi.org/10.1016/j.jtbi.2012.11.015

Ramakrishnan, S., Wesensten, N. J., Kamimori, G. H., Moon, J. E., Balkin, T. J., & Reifman, J. (2016). A unified model of performance for predicting the effects of sleep and caffeine. *Sleep*, *39*, 1827–1841. https://doi.org/10.5665/sleep.6164

Reifman, J., Kumar, K., Wesensten, N. J., Tountas, N. A., Balkin, T. J., & Ramakrishnan, S. (2016). 2B-Alert web: An open-access tool for predicting the effects of sleep/wake schedules and caffeine consumption on neurobehavioral performance. *Sleep*, *39*, 2157–2159. https://doi.org/10.5665/sleep.6318

Reifman, J., Liu, J., Khitrov, M., Kumar, K, Laxminarayan, S, Tountas, N, . . . Ramakrishnan, S. (2017). 2B-Alert app: Real-time learning and individualized prediction of cognitive performance. *Journal of Science and Medicine in Sport*, *20*, S51. https://doi.org/10.1016/j.jsams.2017.09.161

Rupp, T. L., Wesensten, N. J., & Balkin, T. J. (2012). Trait-like vulnerability to total and partial sleep loss. *Sleep*, *35*, 1163–1172. https://doi.org/10.5665/sleep.2010

Van Dongen, H. P., Baynard, M. D., Maislin, G., & Dinges, D. F. (2004). Systematic interindividual differences in neurobehavioral impairment from sleep loss: Evidence of trait-like differential vulnerability. *Sleep*, *27*, 423–433.

Wesensten, N. J., Killgore, W. D., & Balkin, T. J. (2005). Performance and alertness effects of caffeine, dextroamphetamine, and modafinil during sleep deprivation. *Journal of Sleep Research*, *14*, 255–266. https://doi.org/10.1111/j.1365-2869.2005.00468.x

Yang, A., Palmer, A. A., & De Wit, H. (2010). Genetics of caffeine consumption and responses to caffeine. *Psychopharmacology (Berl.)*, *211*, 245–257. https://doi.org/10.1007/s00213-010-1900-1

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.