Journal of Theoretical Biology 358 (2014) 11–24

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Dose-dependent model of caffeine effects on human vigilance during total sleep deprivation

Sridhar Ramakrishnan [a], Srinivas Laxminarayan [a], Nancy J. Wesensten [b], Gary H. Kamimori [b], Thomas J. Balkin [b], Jaques Reifman [a,*]

[a] DoD Biotechnology High Performance Computing Software Applications Institute, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Materiel Command, 504 Scott Street, Fort Detrick, MD 21702, USA
[b] Department of Behavioral Biology, Walter Reed Army Institute of Research, 503 Robert Grant Avenue, Silver Spring, MD 20910, USA

HIGHLIGHTS

- We modeled the dose-dependent effects of caffeine on human vigilance.
- The model predicted the effects of both single and repeated caffeine doses.
- We developed and validated the model using two laboratory studies.
- Individual-specific caffeine models outperformed population-average models.

ARTICLE INFO

Article history:
Received 6 December 2013
Received in revised form
11 March 2014
Accepted 8 May 2014
Available online 20 May 2014

Keywords:
Caffeine model
Dose dependency
Individualized model
Pharmacokinetic-pharmacodynamic model
Cross-study validation

ABSTRACT

Caffeine is the most widely consumed stimulant to counter sleep-loss effects. While the pharmacokinetics of caffeine in the body is well-understood, its alertness-restoring effects are still not well characterized. In fact, mathematical models capable of predicting the effects of varying doses of caffeine on objective measures of vigilance are not available. In this paper, we describe a phenomenological model of the dose-dependent effects of caffeine on psychomotor vigilance task (PVT) performance of sleep-deprived subjects. We used the two-process model of sleep regulation to quantify performance during sleep loss in the absence of caffeine and a dose-dependent multiplier factor derived from the Hill equation to model the effects of single and repeated caffeine doses. We developed and validated the model fits and predictions on PVT lapse (number of reaction times exceeding 500 ms) data from two separate laboratory studies. At the population-average level, the model captured the effects of a range of caffeine doses (50–300 mg), yielding up to a 90% improvement over the two-process model. Individual-specific caffeine models, on average, predicted the effects up to 23% better than population-average caffeine models. The proposed model serves as a useful tool for predicting the dose-dependent effects of caffeine on the PVT performance of sleep-deprived subjects and, therefore, can be used for determining caffeine doses that optimize the timing and duration of peak performance.

Published by Elsevier Ltd.

1. Introduction

Caffeine is the most widely used stimulant drug in both occupational and non-occupational settings. Results from numerous laboratory and field studies have shown that caffeine maintains (Kamimori et al., 2005) or restores (Penetar et al., 1993) neurobehavioral performance in sleep-deprived individuals, with minimal side effects (Bonnet et al., 2005; Brice and Smith, 2002). In the majority of these studies, caffeine has been administered as a single bolus dose of 600 mg (Wesensten et al., 2002; Wesensten et al., 2005) or as smaller, repeated doses of 50, 100, 200, or 300 mg (Kamimori et al., 2005; LaJambe et al., 2005). In these dose ranges, increasing caffeine intake progressively enhances its stimulant effects.

The pharmacokinetics (PK) of caffeine and its dose-dependent metabolism in humans have been well characterized (Bonati et al., 1982; Denaro et al., 1990), and its mechanism of action (antagonism of adenosine receptors) is also well-understood (Bertorelli et al., 1996). However, the pharmacodynamic (PD) effects of caffeine on neurobehavioral performance under sleep loss conditions are not well characterized. A limited number of studies (Wesensten et al., 2002, 2005; Killgore et al., 2008; Kamimori et al., 2005;

* Corresponding author. Tel.: +301 619 7915; fax: +301 619 1983.
E-mail address: jaques.reifman.civ@mail.mil (J. Reifman).

http://dx.doi.org/10.1016/j.jtbi.2014.05.017
0022-5193/Published by Elsevier Ltd.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
MAY 2014		**00-00-2014 to 00-00-2014**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Dose-dependent model of caffeine effects on human vigilance during total sleep deprivation	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **U.S. Army Medical Research and Materiel Command,DoD Biotechnology High Performance Computing Software Applications Institute,Telemedicine and Advanced Technology Research Center,Fort Detrick,MD,21702**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Caffeine isthemostwidelyconsumedstimulanttocountersleep-losseffects.Whilethepharmacoki- netics ofcaffeineinthebodyiswell-understood,itsalertness-restoringeffectsarestillnotwell characterized.Infact,mathematicalmodelscapableofpredictingtheeffectsofvaryingdosesofcaffeine on objectivemeasuresofvigilancearenotavailable.Inthispaper,wedescribeaphenomenological model ofthedose-dependenteffectsofcaffeineonpsychomotorvigilancetask(PVT)performanceof sleep-deprivedsubjects.Weusedthetwo-processmodelofsleepregulationtoquantifyperformance during sleeplossintheabsenceofcaffeineandadose-dependentmultiplierfactorderivedfromtheHill equationtomodeltheeffectsofsingleandrepeatedcaffeinedoses.Wedevelopedandvalidatedthe model fits andpredictionsonPVTlapse(numberofreactiontimesexceeding500ms)datafromtwo separatelaboratorystudies.Atthepopulation-averagelevel,themodelcapturedtheeffectsofarangeof caffeine doses(50?300mg),yieldinguptoa90%improvementoverthetwo-processmodel.Individual- specific caffeinemodels,onaverage,predictedtheeffectsupto23%betterthanpopulation-average caffeine models.Theproposedmodelservesasausefultoolforpredictingthedose-dependenteffectsof caffeine onthePVTperformanceofsleep-deprivedsubjectsand,therefore,canbeusedfordetermining caffeine dosesthatoptimizethetimingandddurationofpeakperformance.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**14**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

LaJambe et al., 2005; Penetar et al., 1993) have assessed the effects of caffeine on objective measures of performance during total sleep deprivation (TSD), but none under the more realistic chronic sleep-restriction condition. Further, the TSD studies differed widely in terms of (1) caffeine dose used, (2) frequency of dosing, (3) timing of dose across the sleep-loss period, and (4) neurobehavioral outcome metric utilized, making it difficult to characterize the caffeine effects. Although the TSD studies provide a basic understanding of the PD effects of caffeine, their utility could be enhanced by the use of mathematical models that could describe and predict such effects. In fact, mathematical models could be used to quantify the dosage and timing of caffeine intake so as to *safely* achieve performance peaks at the desired time of day.

Only two studies have been published that focus on modeling the neurobehavioral performance-enhancing effects of caffeine in humans, especially under acute sleep-loss conditions. In a seminal work, Puckeridge et al. (2011) proposed a 21-parameter model of caffeine's effects on sleep-wake dynamics, with five of the 21 parameters representing caffeine effects. While such a large number of parameters often provide the necessary degrees of freedom for the model to fully capture and fit the variability in the data, it also presents an inherent practical limitation, particularly if the goal is to develop individual-specific models, where the model parameters need to be customized (from limited data) to a particular individual. In addition, their caffeine model assumes a dose-independent PK elimination rate, which contradicts the well-established dose-dependent metabolism of caffeine that results in lower PK elimination rates at higher doses and is particularly prevalent under TSD scenarios (Denaro et al., 1990; Kamimori et al., 1995; Kaplan et al., 1997). Finally, in their work, the effects of caffeine were validated only on subjective sleepiness scores, which may not reflect objective cognitive performance measures (Van Dongen et al., 2003).

Recently, we proposed a parsimonious eight-parameter biomathematical model of the alertness-restoring effects of caffeine under TSD conditions (Ramakrishnan et al., 2013). Although the model was able to capture the effects of both single and repeated caffeine doses and was validated on objective measures of performance from two different studies, it was not a dose-dependent model as it did not provide a means to predict the effects of different caffeine doses.

In this work, we attempt to overcome this limitation by proposing a biomathematical model that quantifies caffeine's neurobehavioral effects as a function of dose under both single and repeated dosing scenarios, while accounting for the dose-dependent metabolism of caffeine in the body. This provides the needed capability to predict the effects of different caffeine doses using a single model. We developed and validated the proposed model, at both population-average and individualized levels, on objective measures of performance collected from two different TSD laboratory studies. Specifically, we developed a population-average model using data from subjects in one study and predicted the effects of a range of caffeine doses on psychomotor vigilance task (PVT) performance of subjects from a second study, and vice versa. In addition, we showed that the individual-specific model predictions were, on average, 23% better than those of the population-average model.

Because baseline measures of performance (i.e., first $\sim$20 h) generally vary from study to study, they need to be normalized to allow for proper inter-study comparisons. In addition, order-of-visit effects have been observed in crossover design studies involving repeated measures (Fayers and King, 2008; Senn, 1988), and require appropriate data processing to eliminate these effects before analysis of the data. Here, in addition to the proposed model, we also developed methods to normalize performance data and eliminate both within- and between-study baseline imbalances to facilitate model development and cross validation using data from different studies.

2. Methods

2.1. Study data

We used PVT data from two studies. The PVT is a simple (one-choice) reaction-time task in which subjects press a button in response to a visual stimulus that is presented on a random interval (2–10 s) schedule over a 10-min period, resulting in $\sim$100 stimulus-response pairs (Dinges and Powell, 1985; Dorrian et al., 2005). For modeling purposes, we calculated the number of response times exceeding 500 ms (the conventional threshold for a lapse) to quantify performance impairment. More lapses indicate greater neurobehavioral performance impairment.

In the first study (*study A*), we used PVT data obtained from a controlled laboratory experiment involving 48 healthy young adults who were kept awake for 29 consecutive hours (Kamimori et al., 2005; Syed et al., 2005). The 48 subjects were randomly assigned to one of the four dose groups (placebo, 50, 100, or 200 mg, $n=12$ subjects/group) and were administered the corresponding dose of Stay Alert® (Amurol Confectioners, Yorkville, IL) caffeinated chewing gum at the beginning of each of three 2-h test blocks after 20, 22, and 24 h of sleep loss (corresponding to 0300, 0500, and 0700 h, respectively, on *day 2*). All subjects completed 10-min PVTs starting at 0800 h on *day 1* and ending at 1200 h on *day 2*, for a total of 29 PVT sessions, including nine sessions before caffeine administration, six sessions during each of the three subsequent 2-h test blocks, and two additional tests after the third 2-h test block.

The data from the second study (*study B*) were collected as part of a randomized Latin Square crossover experiment across four laboratory sessions, each separated by at least 1 mo (washout period), in which 16 healthy young adults were kept awake for 27 consecutive hours (LaJambe et al., 2005). During each of the four laboratory sessions, subjects were administered placebo, 100, 200, or 300 mg of Stay Alert® caffeinated chewing gum three times (the same dose of caffeine was administered in each of the three times) after 20, 22, and 24 h of sleep loss (corresponding to 0300, 0500, and 0700 h, respectively, on *day 2*). Subjects completed 10-min PVTs starting at 0800 h on *day 1* and ending at 1000 h on *day 2*, for a total of 27 PVT sessions, including nine sessions before caffeine administration and six sessions after each of the three caffeine gum administrations.

All subjects in *study A* were habitually low to moderate caffeine users, with an average, self-reported daily caffeine consumption of <400 mg. However, subjects in *study B* were either habitually low (<100 mg/day, $n=8$) or habitually high (>400 mg/day, $n=8$) caffeine users; nevertheless, the differences in PVT performance between the habitually low and habitually high caffeine users were not statistically significant [Wilcoxon rank-sum test; $p>0.05$ (Zar, 1999)] for each of the four doses (placebo, 100, 200, and 300 mg). Consequently, we did not differentiate subjects based on their habitual caffeine usage in the ensuing analyses. All subjects in both studies reported a total sleep time of $\sim$6–9 h for the night preceding study participation. Both studies were approved by the Walter Reed Army Institute of Research Human Use Committee (Silver Spring, MD) and the United States (U.S.) Army Medical Research and Materiel Command Human Subjects Review Board (Ft. Detrick, MD), and written informed consent was obtained from all subjects prior to their participation.

2.2. Data screening and normalization

For *study A*, two subjects (one from placebo and one from 100 mg group) were excluded from analyses due to missing data, resulting in a sample size of 11 subjects for placebo and 100 mg groups. Three subjects from *study B* (crossover design) were

excluded due to missing data, resulting in a sample size of 13 subjects in this study.

Despite the 1-mo washout period between the repeated laboratory visits in *study B*, we observed an order-of-visit effect in the subjects' PVT data such that, for a given subject, PVT performance worsened across visits. Statistical comparisons [Wilcoxon paired, two-sided, signed-rank tests (Zar, 1999)] between pre-caffeine data from the first and last visits indicated that the performance during the last visit was significantly worse than that in the first visit ($p < 0.05$). To correct for this order-of-visit effect, we developed a generalized additive modeling (Hastie and Tibshirani, 1990) approach to quantify and eliminate this effect from each subject's PVT data. Next, to perform inter-study comparisons (across *studies A* and *B*) we normalized the performance data to eliminate baseline differences between the two studies. To this end, we applied an affine transformation (Hastie et al., 2001) to *study B* data. Details pertaining to the elimination of order-of-visit effect (within-study normalization) and elimination of baseline differences (between-study normalization) are provided in Appendix A.

2.3. Model of dose-dependent effects of caffeine on PVT lapses

To develop a dose-dependent model of the effects of caffeine on PVT lapses, we first determined the dependency of the parameters of the model proposed by Ramakrishnan et al. (2013) as a function of caffeine dose. In that model, we hypothesized that, after caffeine intake, the PVT performance estimate [$P_c(t)$] of a sleep-deprived individual at a discrete-time index t can be formulated as follows:

$$P_c(t) = P_0(t) \times g_{PD}(t), \tag{1}$$

where $P_0(t)$ represents the individual's performance without caffeine (referred to as caffeine-free performance) at time awake t and $g_{PD}(t)$ represents the caffeine effect factor, with $0 \le g_{PD}(t) \le 1$, where 1 corresponds to PD effects in the absence of caffeine, i.e., the most impaired performance, and 0 corresponds to the maximal PD effect on PVT performance, i.e., complete restoration with no impairment. Fig. 1 shows a schematic of the effects of caffeine dose D on performance $P_c(t,D)$ and on the caffeine effect factor $g_{PD}(t,D)$.

To characterize caffeine-free performance $P_0(t)$ (i.e., performance under sleep deprivation alone), we used the widely accepted two-process model of sleep regulation (Borbely, 1982), in which performance at time t is a function of the additive interaction of a process reflecting sleep debt (homeostatic Process S) and a process that reflects the circadian rhythm (Process C). Mathematically, in discrete-time notation, $P_0(t)$ can be expressed as follows (Achermann and Borbely, 1994; Rajaraman et al., 2008; Ramakrishnan et al., 2013):

$$P_0(t) = \alpha - \alpha S_0 \exp[-(t-1)\rho T_s] + \beta \sum_{i=1}^{5} a_i \sin\left\{\frac{2\pi}{\tau} i[(t-1)T_s + \phi]\right\}, \tag{2}$$

where α and β denote parameters that control the relative effect of processes S and C on performance, respectively, ρ represents the buildup rate of sleep pressure, T_s denotes the sampling period, S_0 represents the initial sleep pressure state (which depends on the prior sleep/wake history), τ denotes the fundamental time period of the circadian clock ($\sim$24 h), a_i (where $i=1, \ldots, 5$) represent the amplitudes of the five harmonics of Process C, and ϕ denotes the initial circadian phase. Here, we chose to keep the amplitudes of the five harmonics ($a_1=0.97$, $a_2=0.22$, $a_3=0.07$, $a_4=0.03$, and $a_5=0.001$) and the fundamental period ($\tau=24$ h) constant over time, thereby enforcing the shape of Process C to be identical among all individuals (Achermann and Borbely, 1992). The five parameters, α, ρ, β, S_0, and ϕ, were estimated from caffeine-free PVT performance measurements using the approach proposed by Rajaraman et al. (2009).

To model the caffeine effect factor (g_{PD}), we used the one-compartment PK model of caffeine (Bonati et al., 1982; Kamimori et al., 2002), related PK and PD through the Hill equation (Csajka and Verotta, 2006), and expressed g_{PD} of caffeine dose D, administered at time index t_0, as follows:

$$g_{PD}(t,D) = \{\ 1 + M_D \exp[-k_D T_s (t - t_0)]\}^{-1} \qquad \text{for } t \ge t_0, \tag{3}$$

where T_s denotes the sampling period and M_D and k_D represent the amplitude factor and elimination rate parameters of the caffeine model, respectively, which depend on caffeine dose D.

We used the following linear model to capture the effect of dose on M_D:

$$M_D = M_0 \times D, \tag{4}$$

where $M_0 = (F/V_d g_{PK_{50}})$ is the amplitude slope. Here, F and V_d denote the bioavailability of caffeine and volume of distribution in the body, respectively, and $g_{PK_{50}}$ represents the caffeine concentration at which g_{PD} attains half of its maximum effect (Ramakrishnan et al., 2013).

Prior studies on the PK and PD of caffeine under sleep deprivation scenarios have shown that the elimination rate of caffeine decreases with increasing dose (Denaro et al., 1990; Kamimori et al., 1995; Kaplan et al., 1997). We thus modeled the effect of dose on the elimination rate parameter k_D using the following exponential relationship:

$$k_D = k_0 \exp(-zD), \tag{5}$$

where k_0 and z denote the basal elimination rate and the decay constant, respectively. In what follows, we refer to M_0, k_0, and z as the dose-dependent caffeine model parameters.

The g_{PD} model in Eq. (3) does not consider the absorption of caffeine. This is a reasonable approximation for caffeine when ingested via coffee, tea, energy drinks, and most gum products, where the absorption rate is much faster (by a factor of greater than 15) than the elimination rate. For example, the Stay Alert® gum administered in *studies A* and *B* releases $\sim$85% of its caffeine dose within the first 5 min of gum chewing (Kamimori et al., 2002), and therefore has a significantly faster absorption rate than elimination rate.

2.4. Extension of the g_{PD} model for repeated doses

In our previous work (Ramakrishnan et al., 2013), we assumed that the elimination rate is independent of caffeine dose. This assumption allowed us to use linear superposition to extend the single-dose PK model for repeated doses. However, we cannot use this same principle here to compute g_{PD} in Eq. (3) because the elimination rate k_D in Eq. (5) decreases with cumulative increases in caffeine concentration from repeated doses. Therefore, for repeated doses, we modified Eq. (3) so that at the beginning of each dose the amplitude factor and the elimination rate were adjusted based on extant plasma caffeine concentration.

Accordingly, the PD effect after j doses of caffeine of strengths D_1, D_2, …, D_j administered at discrete-time indices t_1, t_2, …, t_j, respectively, can be expressed as follows:

$$g_{PD}(t,D_j) = \begin{cases} 1 & \text{for } t < t_1 \\ (1 + M_{Dj} \exp[-k_{Dj} T_s (t - t_j)])^{-1} & \text{for } t \ge t_j, j = 1, 2, \ldots, \end{cases} \tag{6}$$

where M_{Dj} and k_{Dj} denote the effective amplitude factor and elimination rate parameters, respectively, that depend on the caffeine concentration at time t_j. Using Eqs. (4) and (5), these

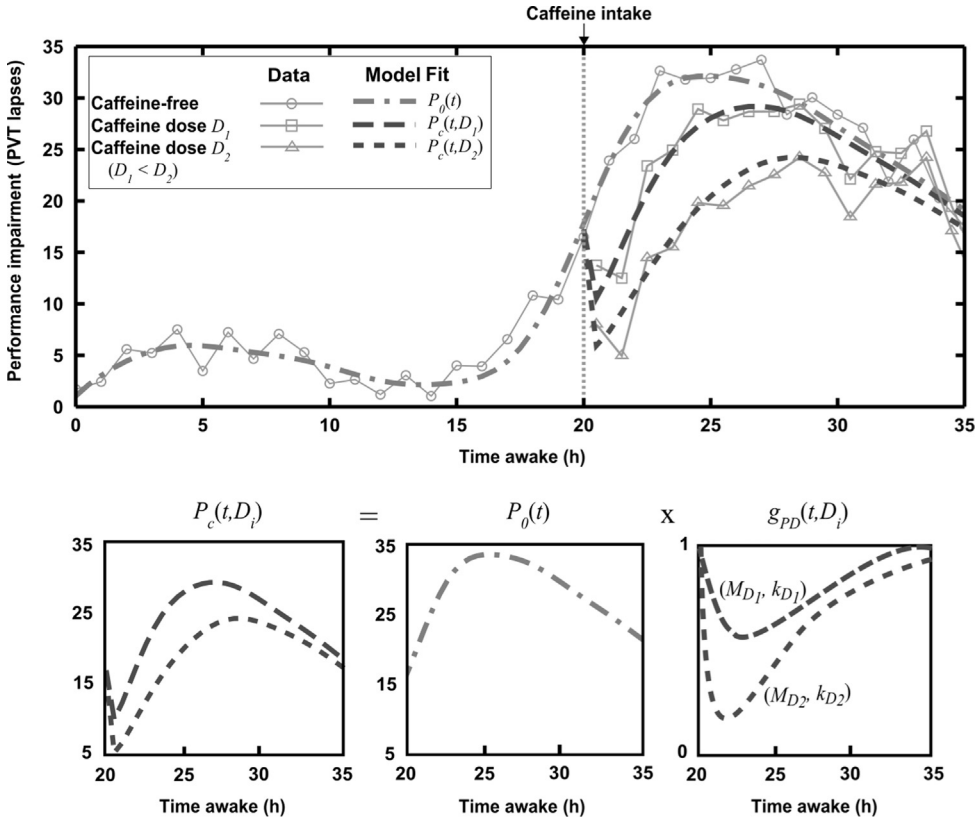

Fig. 1. Schematic showing the development of the dose-dependent caffeine model. Performance at time t following a dose D_i ($i=1$ or 2, for this example) of caffeine $P_c(t, D_i)$ is modeled as the product of (a) performance during sleep deprivation in the absence of caffeine (caffeine-free performance) $P_0(t)$, which is based on the two-process model of sleep regulation, and (b) a caffeine effect factor $g_{PD}(t, D_i)$, with $0 \le g_{PD}(t, D_i) \le 1$, which is governed by two dose-dependent parameters, M_{Di} and k_{Di}, representing the amplitude factor and elimination rate, respectively. PVT, psychomotor vigilance task.

parameters can be expressed as follows:

$$M_{Dj} = M_0 \times [D_j + E(t_j^-)] \quad \text{and} \quad k_{Dj} = k_0 \exp\{-z[D_j + E(t_j^-)]\}, \tag{7}$$

where $E(t_j^-)$ is the *equivalent* caffeine dose representing the caffeine concentration present at time t_j immediately prior to the administration of dose D_j. The expression for $E(t_j^-)$ follows from the standard one-compartment PK model:

$$E(t_j^-) = \begin{cases} 0 & \text{for} \quad j = 1 \\ [E(t_{j-1}^-) + D_{j-1}] \exp[-k_{Dj} T_s (t_j - t_{j-1})] & \text{for} \quad j = 2, 3, \ldots \end{cases} \tag{8}$$

The repeated-dose model in Eq. (6) reduces to Eq. (3) under single-dose conditions. However, the model in Eq. (6) assumes that: (1) each of the repeated caffeine doses are administered via the same formulation and (2) $g_{PK_{50}}$ of the Hill equation, which affects the amplitude slope M_0, remains constant with repeated doses.

2.5. Population-average models

To develop a dose-dependent population-average caffeine model for a study, we fitted the model described in Eqs. (1), (2), and (6) to PVT lapse data from that study. Specifically, we first obtained a population-average caffeine-free model $\overline{P}_0$ by fitting Eq. (2) to data from the placebo group of subjects. We then obtained the population-average caffeine model parameters (M_0, k_0, and z), needed to estimate the dose-dependent population-average $\overline{g}_{PD}$, by minimizing the combined sum of the squared errors between the caffeine model and the data from the different caffeine dose groups in the study. For example, in *study A*, we minimized the following objective function to obtain M_0, k_0, and z:

$$J(M_0, k_0, z) = \sum_{t=t_0}^{t_0+T-1} [\overline{P}_{cm}^{50}(t) - \overline{P}_c(t, 50)]^2 + [\overline{P}_{cm}^{100}(t) - \overline{P}_c(t, 100)]^2 + [\overline{P}_{cm}^{200}(t) - \overline{P}_c(t, 200)]^2, \tag{9}$$

where $\overline{P}_{cm}^{D}$ denotes the population-average data from the *D-mg* dose group, t_0 denotes the time index of the first caffeine dose administration, T represents the total number of PVT measurements taken after the first caffeine administration, and $\overline{P}_c(t, D)$ denotes the population-average performance model after repeated administrations of caffeine doses of D mg, which is given by

$$\overline{P}_c(t, D) = \overline{P}_0(t) \times \overline{g}_{PD}(t, D), \tag{10}$$

where $\overline{g}_{PD}(t, D)$ is computed from Eq. (6). Because the repeated doses could be of different strengths, we used a more general form of Eq. (10):

$$\overline{P}_c(t, D_j) = \overline{P}_0(t) \times \overline{g}_{PD}(t, D_j), \tag{11}$$

where D_j represents the caffeine strength at the j-th dose.

2.6. Individual-specific models

To develop individualized dose-dependent caffeine models for each subject i, we first obtained the caffeine-free component of the model P_0^i by fitting Eq. (2) to the i-th subject's performance data obtained under placebo administration. We then computed a population-average $\overline{g}_{PD}^i(t, D_j)$ using the approach described in the previous section. However, in computing the population-average $\overline{g}_{PD}^i$ for a study, we excluded performance data from the i-th subject. Accordingly, to predict performance of the i-th subject

at time t for the j-th caffeine dose D_j, we used

$$P_c^i(t, D_j) = P_0^i(t) \times \overline{g}_{PD}^i(t, D_j). \quad (12)$$

2.7. Goodness of fits

To assess the goodness of fits, we calculated the root mean squared error (RMSE) between the population-average model fits (and the individual-specific model predictions) and the performance data.

3. Results

We used population-average data from each study to obtain the dose-dependent caffeine model parameters (M_0, k_0, and z) and the corresponding population-average caffeine model fits. We then compared the fits and the cross-study predictions. Finally, we used *study B* data to construct individual-specific caffeine-free models P_0^i and individual-specific caffeine models P_c^i, and compared them with population-average models ($\overline{P}_0$ and $\overline{P}_c^i[=\overline{P}_0 \times \overline{g}_{PD}^i]$).

3.1. Population-average model fits and cross-study predictions

Using Eqs. (2), (6), and (11), we computed two sets of population-average model parameters by fitting these equations on population-average data from *studies A* and *B*. The caffeine-free model parameters (α, ρ, β, S_0, and ϕ) obtained from the two studies are listed in Table B.1 (see Appendix B). Table 1 lists the three caffeine model parameters (M_0, k_0, and z) and their corresponding 95% confidence intervals for each study. We observed that M_0 and z in study B were ~20% lower than their respective values in study A, while the basal elimination rate k_0 was ~40% higher. Fig. 2 illustrates the relationship of the amplitude factor M_D (*left panel*) and elimination rate k_D (*right panel*) as a function of caffeine dose in *studies A* and *B*. The higher elimination rates observed in *study B* compared to *study A* were primarily due to the differences in the population-average PVT performance of the placebo groups $\overline{P}_0$ in the two studies, where $\overline{P}_0$ in *study B* was almost 10 PVT lapses smaller than that in *study A* after the second and third placebo administrations. This resulted in larger $\overline{g}_{PD}(t) = \overline{P}_c(t)/\overline{P}_0(t)$ in *study B* (i.e., caffeine exerted a smaller effect on performance), which translated into higher elimination rates. In contrast, the effect of dose on the amplitude factor M_D in both studies was similar.

We assessed the caffeine model fits by calculating the RMSEs between the fits and the population-average performance data for each of the three caffeine dose groups in each study. To assess the ability of the models to predict the effects of different caffeine doses in a different study, we computed RMSEs between population-average data from one study and population-average model predictions, where the model was fitted on data from the other study. In these assessments, we computed RMSEs for (1) the first 2-h test block (i.e., after the first caffeine dose) and (2) the combined three 2-h test blocks (i.e., after three repeated caffeine doses). Fig. 3 shows population-average caffeine model fits and cross-study predictions of post-caffeine performance for the first 2-h test block (i.e., after the first caffeine dose only) in *studies A* and *B*. Fig. 3, *top*, shows the population-average PVT lapse data for *study A* in the three caffeine dose groups (50, 100, and 200 mg; one in each panel), population-average model fits on these data, and model predictions based on a model trained using the entire *study B* data. The figure also shows the population-average caffeine-free model $\overline{P}_0$, obtained by fitting $P_0(t)$ in Eq. (2) to the placebo data averaged across all subjects in the study. Similarly, Fig. 3, *bottom*, shows data for *study B*, model fits on these data, and model predictions based on a population-average model trained on the entire *study A* data. The caffeine models derived from both studies captured the dose-dependent effects of caffeine on performance and showed improved fits to the data compared to the caffeine-free models [i.e., $\overline{P}_0 = \overline{P}_c$ in Eq. (11) with $\overline{g}_{PD} = 1$]. We observed similar results for the repeated-dose conditions (see Fig. C.1 in Appendix C).

Table 2 lists RMSEs of the population-average model fits and model predictions for both single- (first dose) and repeated-dose conditions. It also lists the associated RMSEs of the caffeine-free models within parentheses, where the caffeine-free predictions correspond to cross-study model fits (i.e., the caffeine-free model fit for *study B* is used as a prediction for *study A*, and vice versa). The RMSEs of the caffeine models were smaller than their caffeine-free counterparts and this difference increased with larger caffeine doses. This suggests that the caffeine model is capturing the dose-dependent effects of caffeine. Across the two studies, the caffeine models showed 57–90% and 17–88% improvements over the caffeine-free models fits and predictions, respectively. RMSEs of the caffeine model predictions in both single- and repeated-dose conditions were only marginally higher than the corresponding fits, indicating good predictive capabilities.

3.2. Individual-specific model predictions

Using the repeated-dose caffeine model in Eq. (12), we also developed individual-specific caffeine models P_c^i to predict post-caffeine performance of each individual in *study B* (crossover design study). We compared these predictions with the corresponding individual-specific, caffeine-free model estimates P_0^i (in Eq. (12) used to compute P_c^i) and the population-average caffeine model predictions ($\overline{P}_c^i = \overline{P}_0 \times \overline{g}_{PD}^i$ in Eq. (11) based on *study B* data while excluding data from the i-th subject). Fig. 4 shows the model predictions after the first caffeine dose (0300 h) for three subjects, Subject #1 (*top*), Subject #2 (*middle*), and Subject #12 (*bottom*), who displayed different levels of sensitivity to caffeine, based on visual inspection of their performance after placebo and caffeine administration. For Subject #1 (high sensitivity to caffeine), both P_c^i and $\overline{P}_c^i$ predicted the dose-dependent effects of caffeine for each of the three caffeine doses better than P_0^i. For Subject #2 (medium sensitivity to caffeine), P_c^i provided more accurate prediction than P_0^i and $\overline{P}_c^i$ for the 100-mg dose condition but not for the 200- and 300-mg conditions. For the latter two caffeine doses, performance improved immediately after caffeine intake but dissipated quickly after ~30 min. For Subject #12 (low sensitivity to caffeine), P_c^i and $\overline{P}_c^i$ predictions were similar to each other and more accurate than the P_0^i estimates only for the 200-mg dose condition. For the 100- and 300-mg conditions, the subject appeared to be insensitive to caffeine. Model predictions for repeated caffeine doses for the same three subjects are shown in Appendix D (Figure D.1).

Table 3 shows a comparison of RMSEs of the individual-specific caffeine model predictions with those of the corresponding caffeine-free model estimates to assess the benefit of accounting for the effects of single caffeine doses of 100, 200, or 300 mg on performance in *study B* subjects. The table also lists the overall

Table 1
Caffeine model parameters obtained by fitting the model on *studies A* and *B* population-average psychomotor vigilance task lapse data after caffeine administration (≥20 h). Data from all three post-caffeine/placebo-administration test blocks (20–22 h, 22–24 h, and 24–26 h of sleep loss) were used to develop the models for each study. Also listed in parentheses are the 95% confidence intervals.

Study	M_0 (g^{-1}) (amplitude slope)	k_0 (h^{-1}) (basal elimination rate)	z (g^{-1}) (decay constant)
A	9.86 (6.96–13.98)	0.49 (0.28–0.88)	1.63 (0.59–4.46)
B	7.73 (5.57–10.71)	0.70 (0.41–1.19)	1.25 (0.50–3.12)

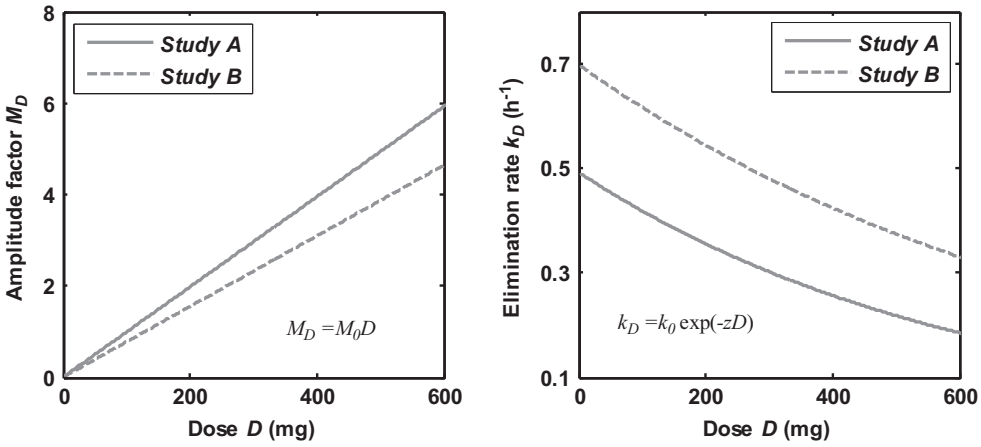

Fig. 2. Amplitude factor M_D [Eq. (4)] and elimination rate k_D [Eq. (5)] as a function of caffeine dose in *studies A* and *B*. We hypothesized linear and exponential relationships for M_D and k_D with caffeine dose, respectively, where the model parameters M_0, k_0, and z were obtained by optimizing the caffeine model on the combined data from all caffeine dose groups (see *Section* 2.5 under Methods).

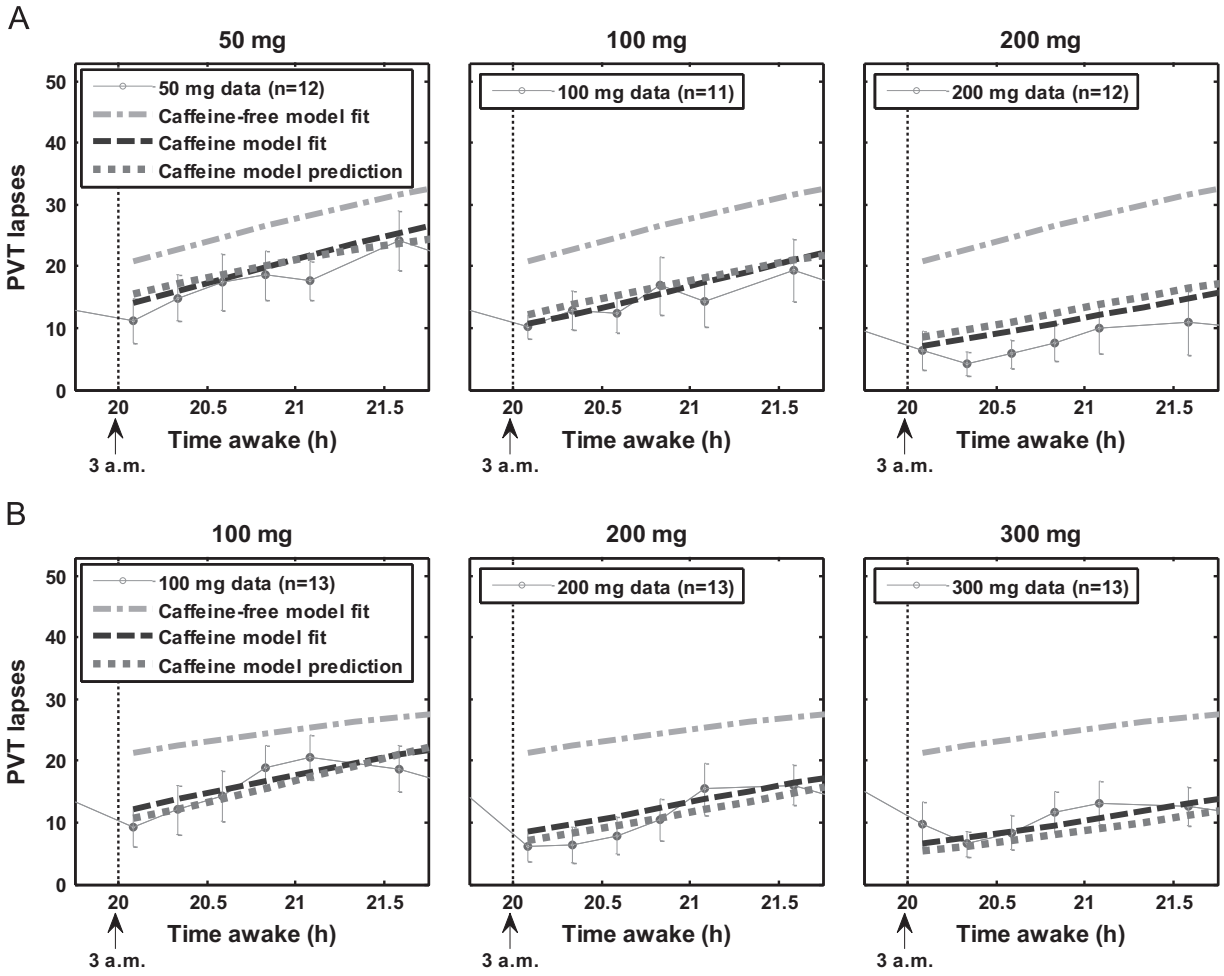

Fig. 3. Dose-dependent caffeine model fits and cross-study model predictions on population-average psychomotor vigilance task (PVT) lapse data (mean ± standard error) measured after the first caffeine dose administration (at 20 h of sleep loss – denoted by the thin dotted vertical line) in *studies A* and *B*. Within each study, the gray dashed-dotted lines represent the caffeine-free model obtained by fitting on PVT data from the placebo group (not shown). The thick dotted lines represent the cross-study caffeine model predictions based on a model obtained by fitting on PVT data from the other study.

mean RMSE over the three dosing conditions, for P_c^i and P_0^i, for each subject, and the average RMSEs across the 13 subjects. P_c^i performed better than P_0^i in at least nine of 13 subjects in each dosing condition, with average improvements ranging from 12% for low (100 mg) caffeine doses to 39% for high (300 mg) doses.

Fig. 5 shows a bar-chart comparison of the average ($n=13$) RMSEs for four different models for a single caffeine dose of 100, 200, and 300 mg in *study B*: (1) population-average caffeine-free model $\overline{P}_0$, (2) individual-specific caffeine-free model P_0^i, (3) population-average caffeine model $\overline{P}_c^i$, and (4) individual-specific

caffeine model P_c^i. Also shown are the overall mean RMSEs for each of the four model predictions across all three doses. Results of statistical comparisons [Wilcoxon paired, two-sided, signed-rank test (Zar, 1999)] between the average RMSEs are indicated by asterisks for the comparisons that showed statistical significance ($p < 0.1$). The caffeine models consistently yielded smaller average RMSEs than their respective caffeine-free counterparts for all dosing conditions, with the differences being significant for the higher doses (200 and 300 mg). The RMSEs for the individual-specific caffeine model predictions were smaller than those for the population-average model for each of the dosing conditions, but the improvement was statistically significant only for the 100-mg dose (~23% improvement). Except for this lowest dose, we observed a consistent trend in the errors ($\overline{P}_0 > P_0^i > \overline{P}_c^i > P_c^i$), with the differences being significant for the overall RMSEs. Equivalent comparisons for each of the three repeated caffeine doses showed a similar RMSE trend: $\overline{P}_0 > P_0^i > \overline{P}_c^i > P_c^i$ (Appendix E, Fig. E.1).

4. Discussion

Caffeine is an efficacious and widely used fatigue countermeasure. However, its dose-dependent effects on neurobehavioral performance have not been adequately characterized, limiting the development of quantitative mathematical models. If available, such models could serve as a tool to more accurately determine the timing and amount of caffeine doses that result in performance peaks at the desired times and that can safely prolong peak performance.

One of the most characteristic effects of sleep loss is degradation in vigilance, as measured by increased response time (Basner and Dinges, 2011). Also, of the various tasks that are purported to measure vigilance, PVT response speed has been shown to be the most sensitive to varying levels of sleep restriction (second only to speed of falling asleep on the Multiple Sleep Latency Test). PVT's high sensitivity is attributed to its insusceptibility to practice effects (Balkin et al., 2004). Furthermore, the PVT can be implemented on mobile platforms, such as personal digital assistants, that can be used in the operational environment.

Table 2
Root mean squared errors (RMSEs) of the caffeine model fits on population-average, post-caffeine psychomotor vigilance task (PVT) lapse data for *study A* and corresponding predictions using a model based on *study B* data, and vice versa. The RMSEs were computed for the first dose only (administered at 20 h of sleep loss; Fig. 3) and for all three repeated doses (administered at 20, 22, and 24 h of sleep loss; Figure C.1). Numbers within parentheses reflect RMSEs of the corresponding caffeine-free model fits and predictions. RMSE units are number of PVT lapses.

Dose (mg)	*Study A*		*Study B*	
	Fit	Prediction	Fit	Prediction
First dose only (at 20 h of sleep loss)				
50	2.25 (8.57)	2.62 (7.05)	–	
100	1.65 (11.56)	2.27 (9.81)	2.18 (8.78)	2.21 (10.37)
200	3.17 (18.41)	4.65 (16.59)	2.36 (13.81)	1.81 (15.44)
300	–		1.89 (13.79)	2.98 (15.64)
All three repeated doses (at 20, 22, and 24 h of sleep loss)				
50	3.12 (11.46)	4.59 (5.57)	–	
100	5.94 (13.83)	6.38 (7.84)	3.18 (9.65)	3.08 (16.86)
200	2.44 (24.01)	2.90 (16.91)	3.43 (14.52)	3.90 (21.58)
300	–		4.68 (15.28)	5.70 (22.51)

The mathematical model described here captures the dose-dependent effects of caffeine on PVT lapses during sleep loss. It builds on our previously developed fixed-dose caffeine model (Ramakrishnan et al., 2013), which estimates the effect of caffeine by multiplying the phenomenological two-process model of sleep regulation (used to characterize performance in the absence of caffeine) with a caffeine-effect factor (g_{PD}) that ranges from 0 (maximal caffeine effect) to 1 (no caffeine effect). The g_{PD} factor is based on the PK-PD sigmoidal relationship of caffeine derived via the Hill equation. It is described by two parameters: (1) an amplitude factor M_D, which describes the magnitude of caffeine effect, and (2) an elimination rate k_D, which describes the duration of caffeine effect.

Here, we incorporated dose-dependent effects into g_{PD} by modeling M_D as a linearly increasing function of caffeine dose and k_D as an exponentially decreasing function of dose. Accordingly, due to the sigmoidal shape of g_{PD}, the magnitude of caffeine effect $[=1/(1+M_D)]$ increases with larger doses but only up to a point (best PVT performance) beyond which the magnitude of the effect saturates. However, the duration of effect keeps increasing with dose, reflecting the saturable metabolic processes involved in the clearance of caffeine (Cheng et al., 1990; Denaro et al., 1990).

In addition, we extended this model to capture the effects of repeated caffeine doses by accounting for the effects of extant caffeine concentration on parameters M_D and k_D at the time of each subsequent caffeine dose administration. The resulting model, containing a total of eight parameters (five parameters to characterize performance in absence of caffeine and three parameters to represent dose-dependent caffeine effects), seems able to predict the effects of single and repeated caffeine doses ranging from 50 to 300 mg on PVT lapse performance of sleep-deprived individuals.

We assessed the accuracy of the proposed dose-dependent caffeine model in two ways: (1) by developing and cross-validating a population-average model using performance data from two separate repeated-dose studies (*A* and *B*) and (2) by developing and validating individual-specific prediction models based on performance data from *study B*. Prior to model development and validation, we normalized the data to eliminate order-of-visit effects observed in *study B* (crossover study) and to eliminate baseline differences in PVT data between *studies A* and *B* (techniques described in Appendix A). Although the modeling results presented here are limited to PVT lapse data, we anticipate that the model output can be scaled to represent other neurobehavioral performance statistics derived from PVT response time measurements.

The population-average, dose-dependent caffeine models obtained from *studies A* and *B* captured the dose-dependent effects of caffeine on performance under both single- and repeated-dosing regimens. The model fits and cross-study predictions were substantially better (up to 88% and 90% for the single- and repeated-dose scenarios, respectively) than the caffeine-free models (Table 2), with greater improvements typically observed in larger doses. Although the model derived from *study B* had a faster elimination rate than the model derived from *study A* (Fig. 2), the models' cross-study predictions were almost as good as their fits. However, for the 100-mg dose group in *study A* and the 200- and 300-mg dosing conditions in *study B*, we observed that after the third caffeine dose (after 24 h of sleep loss) the models predicted better performance (fewer lapses) than was actually obtained (Fig. C.1). In the datasets we utilized for modeling, time awake, time of dosing, and circadian phase were confounded: that is, repeated caffeine administrations occurred at the same times of day and at the same point within time awake (i.e., first dose at 20 h of wakefulness at 0300 h, second dose at 22 h of wakefulness at 0500 h, etc.). It may be that these confounds require redress in future model iterations.

To predict post-caffeine performance of the subjects in *study B*, we developed individual-specific caffeine models for every subject, where the caffeine-free component was individualized and the caffeine-effect multiplier g_{PD} was based on a population-average model using data

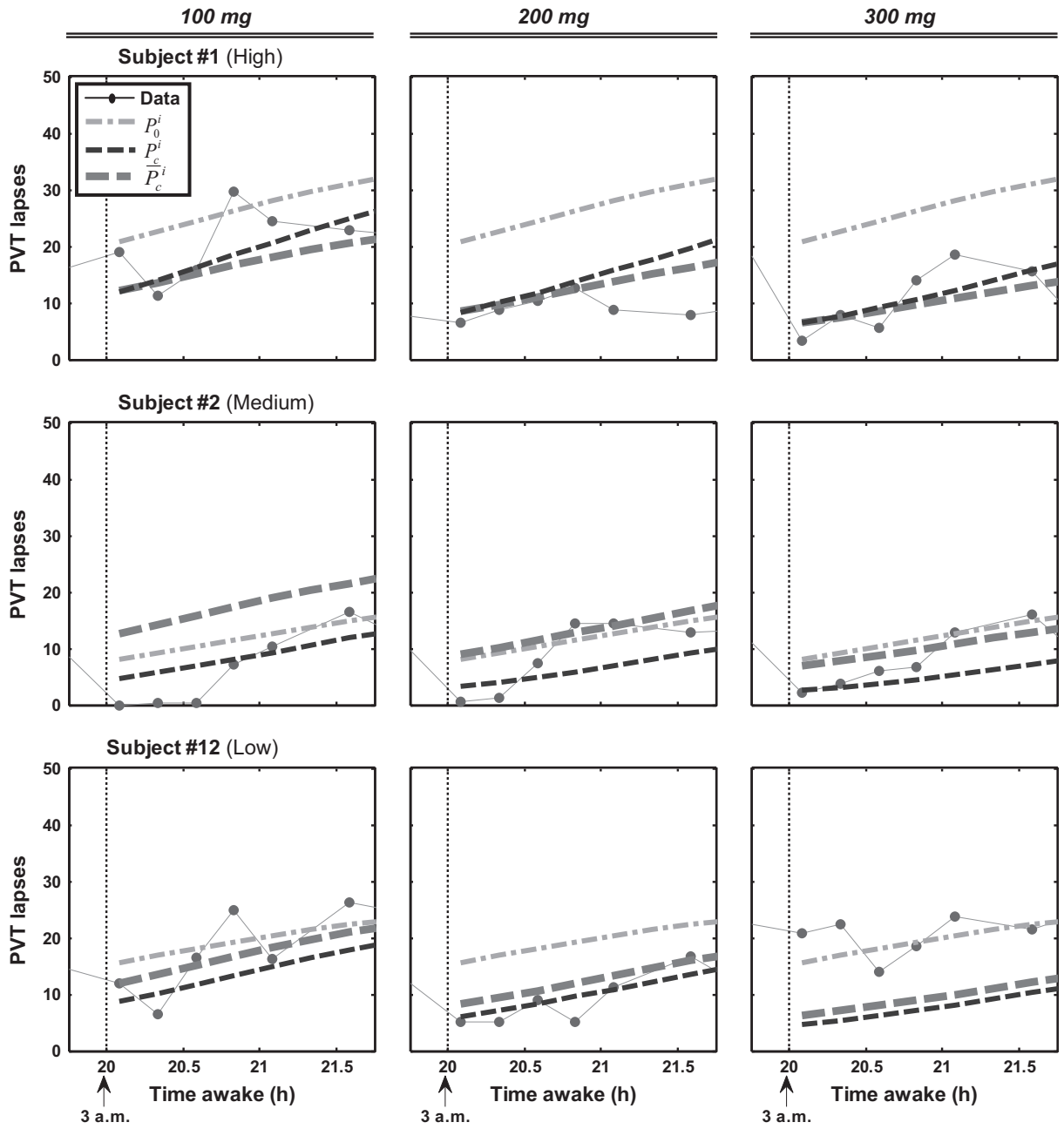

Fig. 4. Individual-specific (P_c^i) and population-average ($\bar{P}_c^i$) dose-dependent caffeine model predictions of three subjects' psychomotor vigilance task (PVT) lapse data after the first caffeine dose of 100 (*left*), 200 (*middle*), and 300 mg (*right*) administered at 20 h of sleep loss (denoted by thin dotted vertical line) in *study B*. Each subject's caffeine sensitivity (determined visually) is shown in parentheses. For each subject, the dashed-dotted lines represent the individual-specific caffeine-free model fit (P_0^i) on PVT lapse data obtained under placebo administration (not shown).

from all subjects except the subject to be predicted. We also compared these results with those obtained from three other models: population-average caffeine model, individual-specific caffeine-free model, and population-average caffeine-free model. The caffeine models yielded an average reduction in prediction error of 40% when compared with their caffeine-free counterparts for single doses of 200 or 300 mg; the reduction in prediction error was only ~13% for the 100-mg dose (Fig. 5). For repeated doses, the prediction errors were reduced by 5% for 100 mg, 30% for 200 mg, and 33% for 300 mg (Fig. E.1). Further, the individual-specific caffeine model was consistently better than the other models across all doses (Figs. 5 and Fig. E.1).

The proposed model has some limitations. One critical limitation of the individual-specific caffeine model is the requirement of individualized caffeine-free performance estimates. In the present work, we used an individual's performance measured after placebo intake to obtain the individualized caffeine-free performance estimates. However, in practice, such data are unlikely to be available. Alternatively, caffeine-free estimates could be obtained by developing an individual-specific model along the lines proposed by Ramakrishnan et al. (2013) using performance data before caffeine intake. This approach, however, requires the availability of sufficient performance data (~20 PVT data points) so that model parameters of the caffeine-free model can be customized to the individual (Rajaraman et al., 2008, 2009).

Another limitation of our proposed individual-specific caffeine model is that the caffeine effect component of the model, g_{PD}, is based on a population-average effect. Individualizing g_{PD} would require availability of performance data after caffeine intake for the specific subject we wish to predict, which may not be attainable in practical applications. Recently, Retey et al. (2006) observed that

Table 3
Root mean squared errors (RMSEs) of the dose-dependent, individual-specific caffeine model predictions (P_c^i) and individual-specific caffeine-free model estimates (P_0^i) of psychomotor vigilance task (PVT) lapse data for *study B* subjects ($n=13$) after the first caffeine administration only (at 20 h of sleep loss) as a function of dose. The overall mean RMSEs are collapsed over the three dosing conditions for each subject. RMSEs of the caffeine models that performed better than their corresponding caffeine-free models are in boldface. RMSE units are number of PVT lapses.

Subject	*Caffeine dose (at 20 h of sleep loss)*							
	100 mg		*200 mg*		*300 mg*		*Overall Mean*	
	P_c^i	P_0^i	P_c^i	P_0^i	P_c^i	P_0^i	P_c^i	P_0^i
1	**5.82**	7.05	**5.75**	16.77	**3.54**	15.12	**5.04**	12.98
2	**4.44**	6.73	5.18	4.90	**4.98**	4.24	**4.87**	5.29
3	12.95	4.50	**10.40**	25.70	**3.24**	17.67	**8.86**	15.96
4	11.84	7.92	**5.40**	16.30	**8.14**	20.78	**8.46**	15.00
5	**8.33**	14.02	**10.53**	18.10	**9.74**	24.15	**9.54**	18.76
6	**6.49**	17.87	**9.91**	14.35	19.45	10.48	**11.95**	14.23
7	21.22	9.53	21.77	5.59	15.67	8.43	19.55	7.85
8	**8.87**	13.70	12.75	10.22	**6.18**	7.27	**9.27**	10.40
9	**13.76**	18.43	**6.98**	14.81	**5.57**	16.28	**8.77**	16.51
10	**10.90**	17.74	**13.67**	25.21	**3.38**	18.07	**9.32**	20.34
11	**7.51**	9.68	**6.75**	13.12	**6.54**	14.87	**6.93**	12.56
12	6.50	5.56	**2.42**	10.40	13.61	3.85	7.51	6.60
13	**4.06**	6.89	**5.40**	11.32	**5.53**	13.00	**5.00**	10.40
Average	**9.44**	10.74	**8.99**	14.37	**8.12**	13.40	**8.85**	12.84

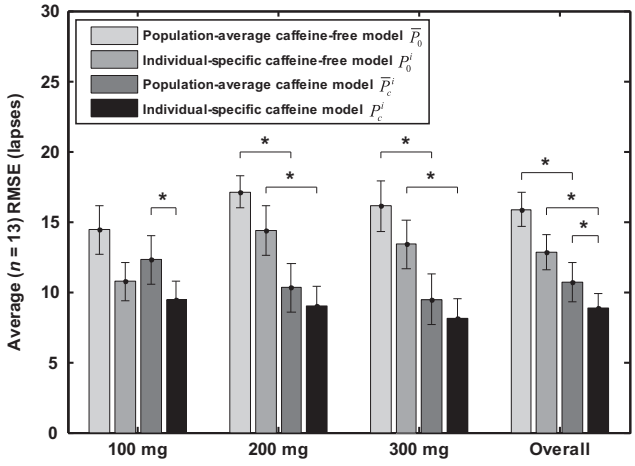

Fig. 5. Average ($n=13$) root mean squared errors (RMSEs) for four different models after the first caffeine dose (100, 200, or 300 mg) administered at 20 h of sleep loss in *study B*. Also shown are the overall mean RMSEs over the three doses. * indicates statistically significant differences ($p<0.1$) between $\overline{P}_0$ vs. $\overline{P}_c^i$, or P_0^i vs. P_c^i, or $\overline{P}_c^i$vs. P_c^i based on Wilcoxon paired, two-sided, signed rank tests. Error bars indicate standard errors.

performance in PVT tests was less impaired after prolonged wakefulness in self-rated caffeine-insensitive individuals than in caffeine-sensitive subjects, suggesting that the sensitivity of the adenosinergic system affects both an individual's resilience to sleep loss and sensitivity to caffeine, but in opposite directions. In other words, those subjects with the largest impairment from sleep loss showed the largest caffeine benefit (Landolt et al., 2012). This offers the possibility of using performance data to ascertain the sleep-loss phenotype of an individual, say, vulnerable, average, or resilient (Ramakrishnan et al., 2012), and then use this information to adjust the population-average g_{PD} to reflect the individual's caffeine sensitivity level.

In the present work, we modeled the effects of sleep loss and caffeine on response time (lapses). Thus, a potential limitation of the present model is that it is based on a simple (one-choice) reaction-time task for which accuracy (or number of errors) is not a relevant metric. Errors of commission can be approximated by "false starts" (responding in the absence of a stimulus). However, in the datasets used in the present work, errors of commission were low and did not differ among caffeine conditions. Errors of omission can be approximated by the number of lapses (responses exceeding 500 ms). This latter metric is indeed sensitive to both sleep loss and to caffeine. Furthermore, because sleep loss generally manifests as decreased response speed rather than increased error rates (Basner and Dinges, 2011), PVT response time serves as a suitable metric for modeling the effects of sleep loss and caffeine.

In both studies considered in the present work, volunteers were also administered the Stanford Sleepiness Scale (SSS). However, SSS scores did not differ statistically between placebo and caffeine conditions (Kamimori et al., 2005). Because our interest is in metrics of objective performance, we did not consider modeling SSS scores —and indeed the failure to find SSS differences between placebo and caffeine conditions supports observations from numerous studies which show that subjective assessments are not appropriate surrogates for objective performance (Van Dongen et al., 2003).

Caffeine's effects on higher-order complex cognitive capacities, such as planning, sequencing, decision making, and memory, have not been considered in this work. While relatively consistent effects have been found with regard to the enhancement of alertness following caffeine consumption, effects on executive functioning (that include a broad spectrum of higher-order cognitive abilities) appear to be mixed (Killgore et al., 2009; Klaassen et al., 2013; Wesensten et al., 2005), showing different degrees of effectiveness at restoring/sustaining performance depending on the particular task in question and the caffeine dose amount. However, the present modeling work is limited to characterizing effects on alertness assessed through PVT and cannot be extrapolated to performance on executive functioning. Further study is thus required to identify the most suitable executive function tasks for which caffeine elicits beneficial dose-dependent effects, so that future efforts may be focused on modeling the effects of sleep loss and caffeine on performance in these tasks as well.

To enhance the utility of the proposed model, we seek to incorporate additional capabilities. In particular, we are developing strategies to model the effects of caffeine on chronically sleep-restricted individuals. We intend to extend the unified model of sleep/wake dynamics recently developed by our group (Rajdev et al., 2013) by incorporating the effects of caffeine under both chronic sleep restriction and total sleep deprivation scenarios in a single model. This supports our long-term goal of incorporating these model components into an integrated computational tool that prescribes countermeasures (e.g., the timing of naps and timing and dosage of caffeine), to optimize an individual's neurobehavioral performance and thereby reduce the risk of sleep-loss and/or circadian-desynchrony-related errors and accidents.

While many challenges remain, the proposed model provides another step towards the development of a wearable computer-based system that considers an individual's sleep/wake history, current and recent-past performance, and caffeine consumption to predict future levels of PVT performance (Khitrov et al., 2014). In fact, with the widespread use of caffeine in both foods and drinks, the ability to predict dose responses of caffeine in a single model holds the key for establishing such a unique capability.

Funding

This work was sponsored by the Military Operational Medicine Research Area Directorate of the U.S. Army Medical Research and Materiel Command, Ft. Detrick, MD, and by the U.S. Department of Defense Medical Research and Development Program (Grant No. DMRDP_13200).

Disclosure statement

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. Department of Defense.

Author contributions

S.R., S.L., and J.R. conceived research; S.R. implemented the model; N.J.W., G.H.K., and T.J.B. provided data for modeling; S.R. wrote the paper, which was edited by N.J.W. and J.R.

Acknowledgments

The authors thank Dr. Wei Lu and Dr. David Thorsley for helpful discussions.

Appendix A

Within-study normalization

In *study B*, each subject's performance under four different dosing conditions (0, 100, 200, and 300 mg) was measured via a crossover design experiment, with a 1-mo washout period between the repeated laboratory visits. We used PVT data of the nine sessions prior to caffeine administration (referred to as baseline sessions) from each visit from all subjects to determine whether there was an order-of-visit effect on performance. Comparisons between baseline data for the first and last visits using Wilcoxon paired, two-sided, signed-rank tests (Zar, 1999) indicated that the performance in each of the nine baseline sessions during the last visit was significantly worse than those in the first visit ($p<0.05$). In addition, we observed that the mean baseline performance impairment (averaged across all subjects and nine sessions) increased linearly with visit number, with a slope (the order effect) of $\sim$4 lapses/visit. When we estimated the order effect by averaging the data across the subjects for each of the nine sessions, we observed a linearly increasing trend with time as well, i.e., the order effect progressively increased with hours awake. Hence, to correct for these time-varying order effects for each subject, we first estimated the effects of the subsequent visits relative to the first visit using the subject's data in the nine baseline sessions, and then subtracted the effect for each session of each visit after the first visit. To this end, we used a generalized additive modeling approach (Hastie and Tibshirani, 1990) to characterize the baseline performance $P_v^i(t)$ of subject i at discrete-time index t on visit v as follows:

$$P_v^i(t) = P_0(\theta^i, t) + (v-1)(\eta_i t + \gamma_i), \tag{A.1}$$

where $P_0(\theta^i,t)$ represents the subject's caffeine-free performance at time awake t on the first visit (which is modeled using the two-process model of sleep regulation as described in Section 2.3 under Methods), θ^i corresponds to the subject's two-process model parameters, and η_i and γ_i denote the order effect parameters capturing the linear time-varying effect of the order of visit on performance. Thus, for each subject, we first used baseline data from the first visit to fit $P_0(\theta^i,t)$, and then estimated the order effect parameters, η_i and γ_i, by fitting Eq. (A.1) to baseline data of the remaining three visits. Finally, we subtracted $(v-1)(\eta_i t+\gamma_i)$ from each subject's PVT data across all sessions for each visit after the first visit to eliminate the time-varying effect of the order of visit.

To assess the effectiveness of this normalization procedure, we separately computed the following two one-factor F-statistics using repeated-measures analysis of variance (Zar, 1999) both before and after normalization: (1) the effect of visit F_{visit} (= between-visit variance/within-visit variance) on baseline data and (2) the effect of dose F_{dose} (= between-dose group variance/within-dose group variance) on post-caffeine data. We observed that F_{visit} (3, 36) dramatically reduced from 10.94 ($p<0.05$) to 0.69 ($p>0.05$) due to normalization, reflecting the reduction of the order-of-visit effect. At the same time, F_{dose} (3, 36) increased from 7.39 ($p<0.05$) to 13.47 ($p<0.05$), suggesting that the normalization procedure improved the differentiability between the four different caffeine dose groups.

Between-study normalization

To perform cross-study validation of the caffeine models between *studies A* and *B*, we required that the baseline PVT data (from the nine sessions prior to caffeine administration) to be similar in the two studies. However, results from the Wilcoxon two-sided rank-sum test suggested that there was a significant ($p<0.05$) difference in the baseline data between the studies. Therefore, we applied an affine transformation (Hastie et al., 2001) to the data in *study B* to ensure that baseline differences between the two studies were eliminated, while not significantly affecting the inter-subject variability within *study B* data. Accordingly, we used baseline PVT data from the two studies and estimated the optimal affine transformation parameters (μ and c) by minimizing the following constrained objective function:

$$J(\mu,c) = \sum_{t=1}^{9} |mean\{f[P_m^B(t),\mu,c]\} - mean\,[P_m^A(t)]|^2 + |std\{f[P_m^B(t),\mu,c]\} - std[P_m^B(t)]|^2, \tag{A.2}$$

where P_m^A and P_m^B denote the measured baseline PVT data from all subjects in *studies A* and *B*, respectively, *std* denotes the standard deviation function characterizing the inter-subject variability in the data, and $f(x,\ \mu,\ c)=\mu x+c$ is the affine transformation function. Finally, we applied the optimal affine transformation ($\mu=0.77$, $c=-1.08$) to each subject's PVT data in *study B* across all sessions for each of the visits to eliminate the baseline differences between the two studies.

To assess the effectiveness of the above baseline normalization procedure, we performed Wilcoxon two-sided rank-sum tests to compare the baseline data in *study A* and the affine-transformed data in *study B*. The results suggested no statistically significant differences among the baseline data following the normalization.

Appendix B

Table B.1 shows the caffeine-free model parameters (α, ρ, β, S_0, and ϕ) obtained in *studies A* and *B*.

Table B.1
Caffeine-free model (sleep loss only) parameters obtained by fitting Eq. (2) on population-average psychomotor vigilance task lapse data of placebo subjects in *studies A* and *B*. The caffeine-free model is derived from the two-process model of sleep regulation.

Study	α (lapses)	ρ (h^{-1})	β (lapses)	S_0	ϕ (h)
A	32.12	0.09	14.71	1.63	3.66
B	27.64	0.06	10.18	1.28	4.72

Appendix C

Figure C.1 shows a comparison of the population-average dose-dependent caffeine model fits and predictions after three repeated caffeine doses in *studies A* and *B*.

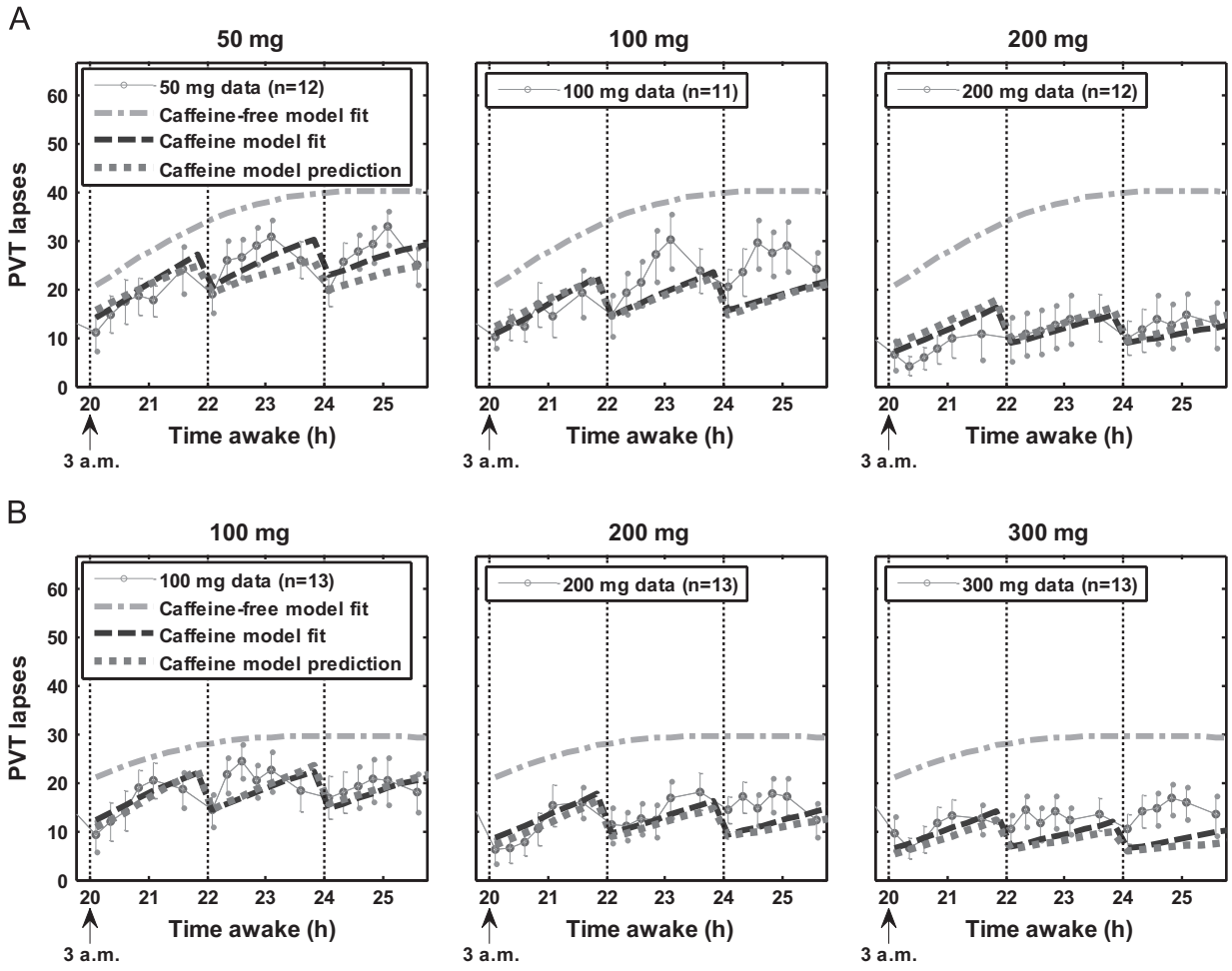

Fig. C.1. Dose-dependent caffeine model fits and cross-study model predictions on population-average psychomotor vigilance task (PVT) lapse data (mean ± standard error) measured after three repeated caffeine dose administrations (at 20, 22, and 24 h of sleep loss – denoted by the thin dotted vertical lines) in *studies A* and *B*. Within each study, the gray dashed-dotted lines represent the caffeine-free model obtained by fitting on PVT data from the placebo group (not shown).

Appendix D

Figure D.1 shows individual-specific, dose-dependent caffeine model performance predictions P_c^i in Eq. (12) after three repeated caffeine doses (the same dose of caffeine administered each of the three times) of 100, 200, and 300 mg for three subjects, Subject #1 (*top*), Subject #2 (*middle*), and Subject #12 (*bottom*), in *study B*. We also compared these predictions with individual-specific caffeine-free estimates P_0^i [in Eq. (12)] and population-average caffeine model predictions $\overline{P}_c^i$ [$= \overline{P}_0 \times \overline{g}_{PD}^i$ in Eq. (11) based on *study B* data while excluding data from the *i*-th subject].

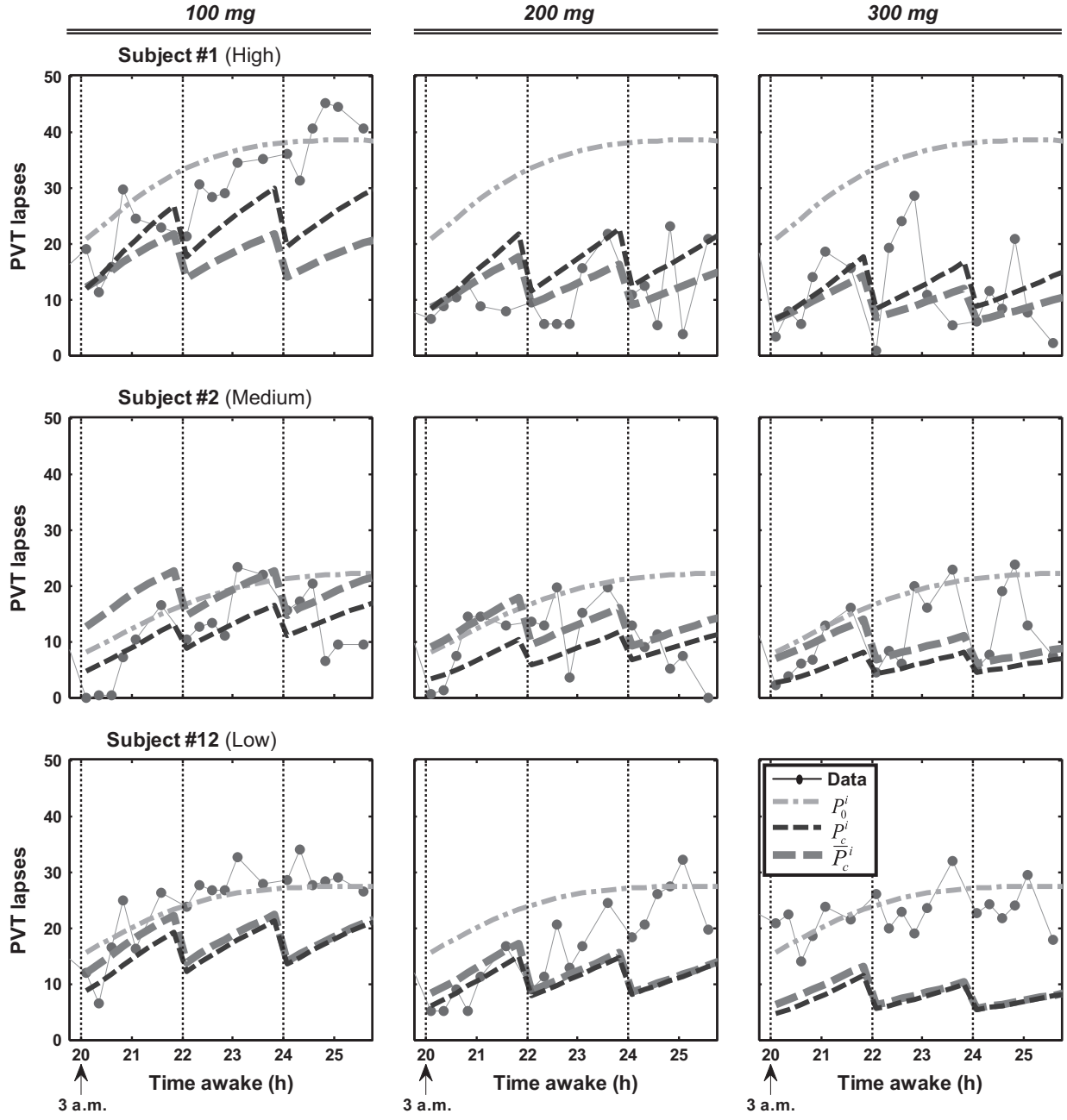

Fig. D.1. Individual-specific (P_c^i) and population-average ($\overline{P}_c^i$) dose-dependent caffeine model predictions of three subjects' psychomotor vigilance task (PVT) lapse data after three repeated caffeine doses of 100 (*left*), 200 (*center*), and 300 mg (*right*) administered at 20, 22, and 24 h of sleep loss (denoted by thin dotted vertical lines) in *study B*. Each subject's caffeine sensitivity (determined visually) is indicated within parentheses. For each subject, the dashed-dotted lines represent the individual-specific caffeine-free model fit (P_0^i) on PVT lapse data obtained under placebo administration (not shown).

Appendix E

Fig. E.1 shows a bar-chart comparison of the average ($n=13$) RMSEs of the four different prediction models for the three repeated caffeine doses (the same dose of caffeine administered each of the three times) of 100, 200, and 300 mg in *study B*. Also shown is the overall average RMSE for each of the four model predictions. The asterisks in Fig. E.1 indicate significant ($p<0.1$) differences based on Wilcoxon paired, two-sided, signed-rank test comparisons.

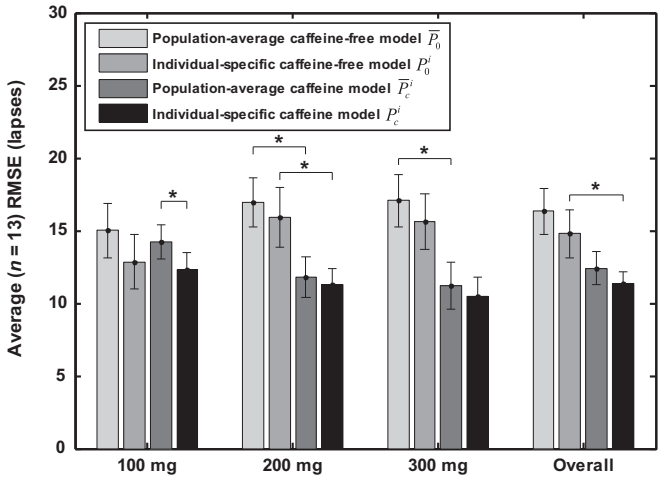

Fig. E.1. Average ($n=13$) root mean squared errors (RMSEs) for four different models after three repeated caffeine doses of either 100, 200, or 300 mg administered at 20, 22, and 24 h of sleep loss in *study B*. Also shown are the overall mean RMSEs over the three doses. * indicates statistically significant differences ($p<0.1$) between $\overline{P}_0$ vs. $\overline{P}_c^i$, or P_0^i vs. P_c^i, or $\overline{P}_c$vs. P_c^i based on Wilcoxon paired, two-sided, signed rank tests. Error bars indicate standard errors.

References

Achermann, P., Borbely, A.A., 1992. Combining different models of sleep regulation. J. Sleep Res. 1 (2), 144–147, jsr001002144.

Achermann, P., Borbely, A.A., 1994. Simulation of daytime vigilance by the additive interaction of a homeostatic and a circadian process. Biol. Cybern. 71 (2), 115–121.

Balkin, T.J., Bliese, P.D., Belenky, G., Sing, H., Thorne, D.R., Thomas, M., Wesensten, N.J., 2004. Comparative utility of instruments for monitoring sleepiness-related performance decrements in the operational environment. J. Sleep Res. 13 (3), 219–227 (http://dx.doi.org/10.1111/j.1365-2869.2004.00407.x).

Basner, M., Dinges, D.F., 2011. Maximizing sensitivity of the psychomotor vigilance test (PVT) to sleep loss. Sleep 34 (5), 581–591.

Bertorelli, R., Marina, N., Adami, M., Ongini, E., 1996. Effects of selective agonists and antagonists for A1 or A2A adenosine receptors on sleep-waking patterns in rats. Drug Develop Res. 37 (2), 65–72.

Bonati, M., Latini, R., Galletti, F., Young, J.F., Tognoni, G., Garattini, S., 1982. Caffeine disposition after oral doses. Clin. Pharmacol. Ther. 32 (1), 98–106 (http://dx.doi.org/0009-9236(82)90243-0).

Bonnet, M.H., Balkin, T.J., Dinges, D.F., Roehrs, T., Rogers, N.L., Wesensten, N.J., 2005. The use of stimulants to modify performance during sleep loss: a review by the sleep deprivation and Stimulant Task Force of the American Academy of Sleep Medicine. Sleep 28 (9), 1163–1187.

Borbely, A.A., 1982. A two process model of sleep regulation. Hum. Neurobiol. 1, 195–204.

Brice, C.F., Smith, A.P., 2002. Effects of caffeine on mood and performance: a study of realistic consumption. Psychopharmacology (Berl) 164 (2), 188–192 (http://dx.doi.org/10.1007/s00213-002-1175-2).

Cheng, W.S., Murphy, T.L., Smith, M.T., Cooksley, W.G., Halliday, J.W., Powell, L.W., 1990. Dose-dependent pharmacokinetics of caffeine in humans: relevance as a test of quantitative liver function. Clin. Pharmacol. Ther. 47 (4), 516–524.

Csajka, C., Verotta, D., 2006. Pharmacokinetic-pharmacodynamic modelling: history and perspectives. J. Pharmacokinet. Pharmacodyn. 33 (3), 227–279 (http://dx.doi.org/10.1007/s10928-005-9002-0).

Denaro, C.P., Brown, C.R., Wilson, M., Jacob 3rd, P., Benowitz, N.L., 1990. Dose-dependency of caffeine metabolism with repeated dosing. Clin. Pharmacol. Ther. 48 (3), 277–285.

Dinges, D.F., Powell, J.W., 1985. Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. Behav. Res. Methods Instr. Comput. 17 (6), 652–655.

Dorrian, J., Rogers, N.L., Dinges, D.F., 2005. Psychomotor vigilance performance: neurocognitive assay sensitive to sleep loss. In: Kushida, C.A. (Ed.), Sleep Deprivation. Clinical Issues, Pharmacology, and Sleep Loss Effects. Marcel Dekker, New York.

Fayers, P.M., King, M., 2008. A highly significant difference in baseline characteristics: the play of chance or evidence of a more selective game? Q. Life Res. 17 (9), 1121–1123 (http://dx.doi.org/10.1007/s11136-008-9390-x).

Hastie, T.J., Tibshirani, R.J., 1990. Generalized Additive Models. Chapman & Hall, London.

Hastie, T.J., Tibshirani, R.J., Friedman, J., 2001. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, New York.

Kamimori, G.H., Johnson, D., Thorne, D., Belenky, G., 2005. Multiple caffeine doses maintain vigilance during early morning operations. Aviat. Space Environ. Med. 76 (11), 1046–1050.

Kamimori, G.H., Karyekar, C.S., Otterstetter, R., Cox, D.S., Balkin, T.J., Belenky, G.L., Eddington, N.D., 2002. The rate of absorption and relative bioavailability of caffeine administered in chewing gum versus capsules to normal healthy volunteers. Int. J. Pharm. 234 (1-2), 159–167 (http://dx.doi.org/S0378517301009589).

Kamimori, G.H., Lugo, S.I., Penetar, D.M., Chamberlain, A.C., Brunhart, G.E., Brunhart, A.E., Eddington, N.D., 1995. Dose-dependent caffeine pharmacokinetics during severe sleep deprivation in humans. Int. J. Clin. Pharmacol. Ther. 33 (3), 182–186.

Kaplan, G.B., Greenblatt, D.J., Ehrenberg, B.L., Goddard, J.E., Cotreau, M.M., Harmatz, J.S., Shader, R.I., 1997. Dose-dependent pharmacokinetics and psychomotor effects of caffeine in humans. J. Clin. Pharmacol. 37 (8), 693–703.

Khitrov, M.Y., Laxminarayan, S., Thorsley, D., Ramakrishnan, S., Rajaraman, S., Wesensten, N.J., Reifman, J., 2014. PC-PVT: A platform for psychomotor vigilance task testing, analysis, and prediction. Behav. Res. Methods 46 (1), 140–147 (http://dx.doi.org/10.3758/s13428-013-0339-9).

Killgore, W.D., Kahn-Greene, E.T., Grugle, N.L., Killgore, D.B., Balkin, T.J., 2009. Sustaining executive functions during sleep deprivation: A comparison of caffeine, dextroamphetamine, and modafinil. Sleep 32 (2), 205–216.

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. J. Sleep Res. 17 (3), 309–321 (http://dx.doi.org/10.1111/j.1365-2869.2008.00654.x, http://dx.doi.org/JSR654 [pii]).

Klaassen, E.B., de Groot, R.H., Evers, E.A., Snel, J., Veerman, E.C., Ligtenberg, A.J., Veltman, D.J., 2013. The effect of caffeine on working memory load-related brain activation in middle-aged males. Neuropharmacology 64, 160–167 (http://dx.doi.org/10.1016/j.neuropharm.2012.06.026).

LaJambe, C.M., Kamimori, G.H., Belenky, G., Balkin, T.J., 2005. Caffeine effects on recovery sleep following 27 h total sleep deprivation. Aviat. Space Environ. Med. 76 (2), 108–113.

Landolt, H.P., Retey, J.V., Adam, M., 2012. Reduced neurobehavioral impairment from sleep deprivation in older adults: contribution of adenosinergic mechanisms. Front. Neurol. 3, 62 (http://dx.doi.org/10.3389/fneur.2012.00062).

Penetar, D., McCann, U., Thorne, D., Kamimori, G., Galinski, C., Sing, H., Belenky, G., 1993. Caffeine reversal of sleep deprivation effects on alertness and mood. Psychopharmacology (Berl) 112 (2-3), 359–365.

Puckeridge, M., Fulcher, B.D., Phillips, A.J., Robinson, P.A., 2011. Incorporation of caffeine into a quantitative model of fatigue and sleep. J. Theor. Biol. 273 (1), 44–54, S0022-5193(10)00673-9.

Rajaraman, S., Gribok, A.V., Wesensten, N.J., Balkin, T.J., Reifman, J., 2008. Individualized performance prediction of sleep-deprived individuals with the two-process model. J. Appl. Physiol. 104 (2), 459–468 (http://dx.doi.org/10.1152/japplphysiol.00877.2007).

Rajaraman, S., Gribok, A.V., Wesensten, N.J., Balkin, T.J., Reifman, J., 2009. An improved methodology for individualized performance prediction of sleep-deprived individuals with the two-process model. Sleep 32 (10), 1377–1392.

Rajdev, P., Thorsley, D., Rajaraman, S., Rupp, T.L., Wesensten, N.J., Balkin, T.J., Reifman, J., 2013. A unified mathematical model to quantify performance impairment for both chronic sleep restriction and total sleep deprivation. J. Theor. Biol. 331, 66–77 (http://dx.doi.org/10.1016/j.jtbi.2013.04.013).

Ramakrishnan, S., Laxminarayan, S., Thorsley, D., Wesensten, N. J., Balkin, T. J., & Reifman, J. (2012). Individualized performance prediction during total sleep deprivation: accounting for trait vulnerability to sleep loss. Conference Proceedings of IEEE Engineering in Medicine and Biology, 2012, 5574–5577.

Ramakrishnan, S., Rajaraman, S., Laxminarayan, S., Wesensten, N.J., Kamimori, G.H., Balkin, T.J., Reifman, J., 2013. A biomathematical model of the restoring effects of caffeine on cognitive performance during sleep deprivation. J. Theor. Biol. 319, 23–33.

Retey, J.V., Adam, M., Gottselig, J.M., Khatami, R., Durr, R., Achermann, P., Landolt, H. P., 2006. Adenosinergic mechanisms contribute to individual differences in sleep deprivation-induced changes in neurobehavioral function and brain rhythmic activity. J. Neurosci. 26 (41), 10472–10479 (http://dx.doi.org/10.1523/JNEUROSCI.1538-06.2006).

Senn, S.J., 1988. Cross-over trials, carry-over effects and the art of self-delusion. Stat. Med. 7 (10), 1099–1101.

Syed, S.A., Kamimori, G.H., Kelly, W., Eddington, N.D., 2005. Multiple dose pharmacokinetics of caffeine administered in chewing gum to normal healthy volunteers. Biopharm. Drug Dispos. 26, 403–409 (http://dx.doi.org/10.1002/bdd.469).

Van Dongen, H.P., Maislin, G., Mullington, J.M., Dinges, D.F., 2003. The cumulative cost of additional wakefulness: dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. Sleep 26 (2), 117–126.

Wesensten, N.J., Belenky, G., Kautz, M.A., Thorne, D.R., Reichardt, R.M., Balkin, T.J., 2002. Maintaining alertness and performance during sleep deprivation: modafinil versus caffeine. Psychopharmacology (Berl) 159 (3), 238–247 (http://dx.doi.org/10.1007/s002130100916).

Wesensten, N.J., Killgore, W.D. S., Balkin, T.J., 2005. Performance and alertness effects of caffeine, dextro amphetamine, and modafinil during sleep deprivation. J. Sleep Res. 14 (3), 255–266.

Zar, J.H., 1999. Biostatistical Analysis, fourth ed. Prentice Hall, Upper Saddle River, N.J.