Received: 10 January 2024 | Revised: 19 March 2024 | Accepted: 4 April 2024

DOI: 10.1111/jsr.14220

RESEARCH ARTICLE

Individualised prediction of resilience and vulnerability to sleep loss using EEG features

Manivannan Subramaniyan[1,2] | John D. Hughes[3] | Tracy J. Doty[3] | William D. S. Killgore[4] | Jaques Reifman[1]

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

[2]The Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc., Bethesda, Maryland, USA

[3]Behavioral Biology Branch, Center for Military Psychiatry and Neuroscience Research, Walter Reed Army Institute of Research, Silver Spring, Maryland, USA

[4]Department of Psychiatry, University of Arizona College of Medicine, Tucson, Arizona, USA

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 Development Command, ATTN: FCMR-TT, 504 Scott Street, Fort Detrick, MD 21702-5012, USA.
Email: jaques.reifman.civ@health.mil

Funding information
U.S. Army Medical Research and Development Command, Grant/Award Number: W81XWH20C0031

Summary

It is well established that individuals differ in their response to sleep loss. However, existing methods to predict an individual's sleep-loss phenotype are not scalable or involve effort-dependent neurobehavioural tests. To overcome these limitations, we sought to predict an individual's level of resilience or vulnerability to sleep loss using electroencephalographic (EEG) features obtained from routine night sleep. To this end, we retrospectively analysed five studies in which 96 healthy young adults (41 women) completed a laboratory baseline-sleep phase followed by a sleep-loss challenge. After classifying subjects into sleep-loss phenotypic groups, we extracted two EEG features from the first sleep cycle (median duration: 1.6 h), slow-wave activity (SWA) power and SWA rise rate, from four channels during the baseline nights. Using these data, we developed two sets of logistic regression classifiers (resilient versus not-resilient and vulnerable versus not-vulnerable) to predict the probability of sleep-loss resilience or vulnerability, respectively, and evaluated model performance using test datasets not used in model development. Consistently, the most predictive features came from the left cerebral hemisphere. For the resilient versus not-resilient classifiers, we obtained an average testing performance of 0.68 for the area under the receiver operating characteristic curve, 0.72 for accuracy, 0.50 for sensitivity, 0.84 for specificity, 0.61 for positive predictive value, and 3.59 for likelihood ratio. We obtained similar performance for the vulnerable versus not-vulnerable classifiers. These results indicate that logistic regression classifiers based on SWA power and SWA rise rate from routine night sleep can largely predict an individual's sleep-loss phenotype.

KEYWORDS
EEG, logistic regression, resilient, sleep loss, slow-wave activity, vulnerable

1 | INTRODUCTION

Sustained vigilance is required in military, aviation, public safety, and medical professions, where individuals often perform their duties under limited sleep due to work and social demands (Chattu et al., 2019; Good et al., 2020). However, vigilance and mental acuity are often impaired under sleep-loss conditions, potentially resulting in undesirable consequences, such as near misses, accidents, medical mistakes, or even loss of life (Erickson et al., 2017; Institute of Medicine, 2006). Ideally, tasks requiring sustained vigilance should be performed by individuals who can maintain mental acuity despite sleep loss. Indeed, it is well established that there is large

individual-to-individual variability in response to sleep loss, with some individuals being resilient, some vulnerable, and others in between these two extreme phenotypic responses (Rupp et al., 2012; Tkachenko & Dinges, 2018; Van Dongen et al., 2004). However, current methods to identify such phenotypic responses are not practical, often involving time-consuming, sleep-deprivation challenges performed in a laboratory setting (Chua et al., 2014, 2019; Cui et al., 2015; Rupp et al., 2012; St. Hilaire et al., 2019; Van Dongen et al., 2004; Yamazaki et al., 2021; Zhao et al., 2018). Here, we sought a more practical and scalable approach that does not require sleep-loss challenges. Such an approach to identify an individual's phenotypic response to sleep loss could help to improve occupational safety and productivity, where resilient individuals could be assigned to tasks requiring sustained vigilance and attention while vulnerable individuals could be offered sleep-loss countermeasures to mitigate safety concerns and fatigue-related performance impairments.

Previous studies to characterise an individual's phenotypic response to sleep loss without requiring a sleep-deprivation challenge found that, under well-rested awake conditions, resilient and vulnerable individuals differ in neurobehavioural task performance (Chua et al., 2014; Patanaik et al., 2014; Patanaik et al., 2015; Zhao et al., 2018), electroencephalographic (EEG) spectral power in the high-theta frequency band (Chua et al., 2014), heart rate and its variability (Chua et al., 2014), cardiovascular haemodynamic measures (Yamazaki et al., 2021), brain activation level (Caldwell et al., 2005; Chee et al., 2006; Mu et al., 2005), as well as brain structural (Cui et al., 2015; Rocklage et al., 2009) and functional connectivity (Yeo et al., 2015). Furthermore, other studies found that baseline metrics derived from neuroimaging (Caldwell et al., 2005; Chee et al., 2006; Cui et al., 2015; Rocklage et al., 2009; Zhao et al., 2018), psychomotor vigilance tests (PVTs) (Galli et al., 2022), or sleep (Subramaniyan et al., 2023) are linearly associated with vulnerability to sleep loss as indexed by neurobehavioural test performance. Finally, a few studies have used PVT metrics collected during well-rested baseline conditions (Chua et al., 2019; Patanaik et al., 2014; Patanaik et al., 2015; St. Hilaire et al., 2019) as well as metrics derived from structural (Xu et al., 2021) or functional (Yeo et al., 2015) brain imaging to develop predictive mathematical models of an individual's phenotypic response to sleep loss.

While these modelling studies have helped to identify some promising discriminatory features, they have methodological limitations. One limitation is that features derived from PVTs and used as inputs to the models (Chua et al., 2019; Patanaik et al., 2014; Patanaik et al., 2015; St. Hilaire et al., 2019) depend on an individual's level of effort (Brewer et al., 2017; Massar et al., 2016; Robison et al., 2021), making them less reliable as a predictive variable. In contrast, such a limitation is not present in features derived from brain imaging, where subjects do not need to perform cognitive tasks (Xu et al., 2021; Yeo et al., 2015). However, brain imaging is costly and not easily accessible or scalable. Another limitation is that some studies do not use an independent dataset, that is, data not used to train the models, to estimate model performance (Chua et al., 2019; Patanaik et al., 2014; Xu et al., 2021). As a result, the reported model performance is likely to be biased (Hastie et al., 2009; Varma & Simon, 2006), and the actual ability of these models to generalise and predict unseen data is unknown.

To overcome the above limitations, here we aimed to develop data-driven models using EEG features collected during routine night sleep, to predict the probability of a specific individual being either resilient or vulnerable to sleep loss, and to estimate the models' performance using independent datasets not used for model training. Previously, we found that, during routine night sleep, we could discriminate between groups of resilient individuals and groups of vulnerable individuals based on slow-wave activity (SWA) power during the first sleep cycle and SWA rise rate during the first 20 min of sleep (Subramaniyan et al., 2023). Here, we sought to investigate the ability of these features to discriminate between these phenotypic responses to sleep loss at the individual level. Although the ultimate use of these models does not require a sleep-deprivation challenge, training of such models required knowledge of the sleep-loss phenotype of individuals as measured by the PVT. Hence, we leveraged data previously collected from five sleep-deprivation studies to obtain EEG signals recorded during routine sleep on baseline nights before the start of the challenge and PVT data during the challenge. In total, we used data from 96 healthy young men and women, and built logistic regression classifiers following a nested cross-validation procedure that conservatively estimated the models' performance on novel unseen data.

2 | METHODS

2.1 | Study design

For developing classifiers to predict an individual's resilience or vulnerability to sleep loss, we used data from five previously published studies (Doty et al., 2017; Hansen et al., 2019; Reifman et al., 2019; Rupp et al., 2012; Vital-Lopez et al., 2023), involving 96 healthy young adults (41 women) between 18 and 39 years of age (Table 1). In Studies 1, 4, and 5, a total of three subjects (one in each study) were left handed while the remaining were right handed. Handedness information was not available for Studies 2 and 3. In all five studies, subjects slept in the laboratory for 1 to 7 baseline nights, from which we obtained polysomnography (PSG) data. Although the studies took place at three different sleep laboratories (Studies 1, 2, and 3 at the Walter Reed Army Institute of Research, Silver Spring, MD; Study 4 at the Washington State University, Spokane, WA; and Study 5 at the University of Arizona, Tucson, AZ), all study subjects experienced comparable sleep environments during PSG recordings. Specifically, all subjects slept in sound-attenuated rooms with the lights turned off and ambient temperature set to around 20–23°C, and had no access to personal electronic devices or the Internet. Following the baseline sleep, the subjects underwent total sleep deprivation (TSD) in Studies 2, 4, and 5, chronic sleep restriction (CSR) in Study 3 or both in Study 1, followed by a recovery phase. During scheduled wakefulness, the subjects performed PVTs every 1–3 h starting immediately after the

TABLE 1 Description of the five studies used to develop and validate the logistic regression classifiers for predicting an individual's resilience or vulnerability to sleep loss.

Study	Baseline nights, *n*	Baseline nights with PSG data, *n*	TIB during baseline nights, *h*	Sleep-loss protocol	Number of subjects (women)			Age, years [mean (1 SD)]
					R	*I*	*V*	
1. (Rupp et al., 2012)[a]	7	1	10	63 h TSD and 3 h TIB × 7 d	6 (3)	6 (3)	6 (3)	28.1 (4.8)
2. (Reifman et al., 2019)	1	1	8	62 h TSD	7 (2)	7 (4)	7 (1)	24.6 (4.6)
3. (Doty et al., 2017)	5	5	10	5 h TIB × 5 d	8 (5)	8 (2)	8 (3)	25.4 (3.6)
4. (Hansen et al., 2019)	3	3	10	48 h TSD	4 (1)	4 (2)	4 (3)	27.4 (6.9)
5. (Vital-Lopez et al., 2023)	1	1	8	62 h TSD	7 (2)	7 (2)	7 (5)	21.9 (4.4)
Overall					32 (13)	32 (13)	32 (15)	25.2 (5.1)

Abbreviations: I, intermediate subject group; PSG, polysomnography; R, resilient subject group; SD, standard deviation; TIB, time in bed; TSD, total sleep deprivation; V, vulnerable subject group.

[a]Cross-over study design with a gap of 2–4 weeks between the 63 h of TSD and the 3 h TIB challenges.

baseline-sleep phase, through the sleep-loss challenge, and until the end of the recovery phase.

2.2 | Subject classification

We classified the subjects into three groups based on the extent to which their PVT reaction times during sleep-loss periods changed relative to those during baseline wake periods, as described previously (Subramaniyan et al., 2023). Briefly, for each subject, we first normalised the PVT reaction times by dividing the mean reaction time during the sleep-loss period by that of the baseline period. For the crossover study (Rupp et al., 2012), consisting of TSD and CSR challenges, we separately normalised the reaction times for each subject for each of the two phases, resulting in two values that we averaged to obtain a single normalised reaction time per subject. Then, within each study, we rank-ordered the subjects by their average normalised reaction times and labelled the lower third as resilient, the upper third as vulnerable, and the middle third as "intermediate".

2.3 | Sleep EEG data and preprocessing

The EEG data of the five studies (Doty et al., 2017; Hansen et al., 2019; Reifman et al., 2019; Rupp et al., 2012; Vital-Lopez et al., 2023) had sampling rates ranging from 100 to 500 Hz. For studies with a sampling rate greater than 100 Hz, we down-sampled the EEG data to either 100 Hz (Doty et al., 2017; Hansen et al., 2019; Vital-Lopez et al., 2023) or 128 Hz (Reifman et al., 2019). Because one of the five studies did not record signals from the frontal EEG channels, we only analysed the central and occipital channels (C3, C4, O1, and O2 referenced to the contralateral mastoids, Figure 1a) common to all five studies. We high-pass filtered the data (6 dB cutoff frequency at 0.125 Hz), and scored sleep stages in 30 s epochs following the guidelines of the American Academy of Sleep Medicine (Silber et al., 2007). Finally, we segmented each channel's data into 5 s epochs and excluded epochs in which we detected electrical or physiological artefacts, as described previously (Subramaniyan et al., 2023).

2.4 | EEG features

As potential predictors of sleep-loss phenotype, from each EEG channel, we extracted the following two EEG features, which we have previously shown to discriminate between groups of sleep-loss resilient and vulnerable individuals (Subramaniyan et al., 2023): (1) SWA power, defined as the mean power spectral density in the frequency range of 0.2–4.0 Hz, and (2) SWA rise rate, defined as the rate at which SWA power changed during the first 20 min after sleep onset. By extracting these two features from the four EEG channels (C3, C4, O1, and O2), we obtained a total of eight EEG features to predict an individual's sleep-loss phenotype.

We computed the SWA power and SWA rise rate as described previously (Subramaniyan et al., 2023). Briefly, to compute the SWA power for a given EEG channel, we obtained power spectral density estimates for artefact-free epochs within the N2 and N3 stages of the first sleep cycle and averaged the power spectral density within the frequency band of 0.2–4.0 Hz for each epoch. Then, we averaged the SWA power across all epochs and log (base 10) transformed the average. To compute the SWA rise rate, we fitted a robust linear regression to the SWA power data for epochs within the N1, N2, or N3 stages of the first 20 min of sleep. To enhance the reliability of the estimated EEG features, we required that subjects had at least one baseline night with a minimum of 30 min (cumulative) of artefact-free

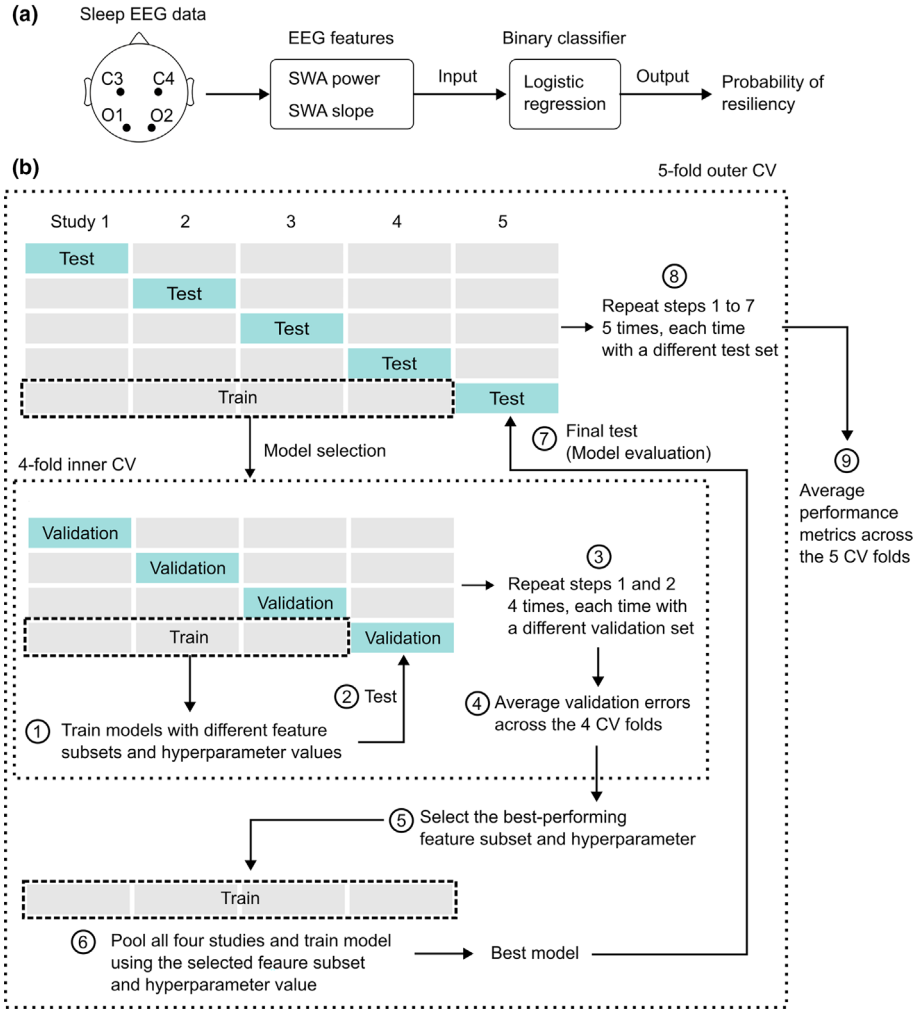

FIGURE 1 Model development. (a) We used slow-wave activity (SWA) power and SWA rise rate extracted from electroencephalographic (EEG) data (channels C3, C4, O1, and O2) collected during routine night sleep as inputs to a logistic regression model, to produce the probability of resilience to sleep loss as an output. (b) Five-fold nested cross-validation (CV) procedure for model selection and evaluation. For model selection, in the "inner CV", we trained models with different EEG feature subsets and hyperparameter values (Steps 1–3). We then used the feature subset and hyperparameter value of the model that showed the best validation performance based on validation error averaged across the four inner-CV folds (Steps 4 and 5), to fit a new model (the "best" model) to the pool of all four studies of the inner-CV (Step 6). Next, we tested (model evaluation) the best model associated with each outer-CV fold on the fifth study which was set aside for the final testing (Steps 7 and 8). Finally, to estimate the overall performance of the models, we averaged the results over the five outer-CV folds (Step 9).

N2 or N3 stage data in the first sleep cycle. Imposing this criterion resulted in the exclusion of one subject in Study 2 (Reifman et al., 2019) and two subjects in Study 5 (Vital-Lopez et al., 2023). For a few subjects (depending on the channel, 2–4 resilient, 3 intermediate, and 2–4 vulnerable from Study 1 (Rupp et al., 2012); 1 resilient from Study 3 (Doty et al., 2017); and 1 vulnerable from Study 4 (Hansen et al., 2019)), we only used data from one of the two nights because the other night did not meet the quality criterion. To reduce variability and minimise the "first-night" effect (Agnew Jr. et al., 1966), for subjects for whom we had PSG recordings from multiple nights of baseline sleep (Studies 3 and 4), we selected the first and last nights and averaged the EEG features from the two nights. For the crossover study (Rupp et al., 2012), for which we had one baseline night with PSG recordings prior to the TSD phase and another for the CSR phase, for each subject, we averaged the EEG features extracted from the baseline night of each of the two study phases. We corrected the EEG features for the effect of age (Sprecher et al., 2016) using a regression model, as described previously (Wang et al., 2020).

The raw EEG feature values of a given subject group (resilient, intermediate, or vulnerable) varied across the studies for a given EEG channel, likely due to recording-setup differences between studies, and across the channels within a given study due to differing recording locations on the scalp. Because of this variability, we could not directly associate the raw EEG feature values with the sleep-loss phenotypic categories (Subramaniyan et al., 2023). Therefore, to pool studies together, we performed a within-study *z*-scoring normalisation that brought feature values from different studies and channels into a common scale. To normalise the data, within a given study, for each EEG channel, we *z*-scored the feature values of each subject using the mean and standard deviation computed from the pool of all subjects of that study.

2.5 | Model for predicting resilience to sleep loss

Although we classified the subjects into resilient, intermediate, or vulnerable individuals, our main goal was to predict whether a given subject was resilient to sleep loss or not. Accordingly, we first developed a classifier with subjects labelled as resilient and not-resilient, where the not-resilient class consisted of both intermediate and vulnerable subjects pooled together. We then employed the following binary logistic regression classifier [Equation (1), Figure 1a]:

$$p(\text{resilient};X)=\frac{1}{1+e^{-(W^{T}X+b)}}, \quad (1)$$

where p represents the probability of a given subject being resilient, X denotes the vector of EEG features from a given subject, W denotes the vector of model coefficients (parameters) associated with the corresponding vector of EEG features, T denotes the matrix-transpose operation, and b represents the model's intercept parameter. We obtained the model parameters by minimising an objective function that consisted of a cross-entropy error term (Bishop, 2006) and a regularisation term (L2-norm penalty, for minimising overfitting) defined by a hyperparameter. The cross-entropy error measured how much the predicted probabilities deviated from their corresponding true values and the hyperparameter controlled the strength of regularisation. If two models differed in their parameter values or their number of input features, we treated them as different models. For simplicity, we referred to the cross-entropy error as "error". We developed the models using the open-source Python package Scikit-learn (Pedregosa et al., 2011).

2.6 | Model building

We built models using a five-fold nested cross-validation (CV) procedure (Figure 1b) consisting of an "inner CV", in which we performed model selection (training), and an "outer CV", in which we performed model evaluation (testing).

2.6.1 | Data preparation

To create the five folds of data, instead of pooling subjects from all five studies and randomly splitting them into five folds, we simply assigned data from each study to a separate data fold. We split the data this way because we used EEG features z-scored within each study, as z-scoring is not possible at the individual subject level. As such, our model was designed to be ultimately used to predict the sleep-loss phenotype of each individual of a group for which we collected EEG data during routine sleep.

2.6.2 | Model selection

For model selection (Steps 1–6, Figure 1b), we fitted different logistic regression models to training datasets and selected a single model based on the performances of the models on validation datasets. Specifically, for each of the five outer-CV folds, we performed a four-fold inner-CV (three studies for training and one study for validation; Steps 1–4, Figure 1b), which we used for model selection (i.e., the identification of the most informative feature-subset and optimal hyperparameter; Step 5, Figure 1b). When we selected the optimal hyperparameter or the most informative feature-subset (among those of different sizes) based on the validation error (cross-entropy error of the validation dataset; Step 5, Figure 1b), there were several models with validation errors that were close to the lowest validation error. Hence, among such models, we selected the least-complex model (i.e., the model fitted with the smallest number of features and the highest regularisation strength) as our "best-performing" model. If two models have similar validation errors, the less-complex model is likely to generalise better on future novel data.

To identify the most informative feature subset, we performed forward-stepwise feature selection (Hastie et al., 2009), an iterative procedure where we first assessed (validated) the performance of models fitted with a single EEG feature and then progressively tested models fitted with larger subsets of features while retaining the feature subsets associated with the best-performing models from the previous iterations. When fitting a model with a given feature subset, we also optimised the hyperparameter for that model. Next, we selected the model that had the best performance across all iterations. Finally, using the feature subset and hyperparameter value of the selected model, we fitted a new model to the training data set (four studies pooled together; Step 6, Figure 1b) of the given outer-CV fold and assessed the model performance on the test data set (fifth study) not used for model selection (Step 7, Figure 1b).

2.6.3 | Model evaluation and performance metrics

To assess the performance of a model in the outer-CV folds (Step 8, Figure 1b), we used six metrics: area under the receiver operating characteristic curve (AUC), accuracy, sensitivity, specificity, positive predictive value (PPV), and likelihood ratio (LR, true positive rate divided by false positive rate) as defined previously (Webb & Sidebotham, 2020). To compute each metric, we classified a subject as resilient if their predicted probability was greater than or equal to 0.50, and as not-resilient otherwise, and designated resilient individuals as the positive class. Finally, to obtain a measure of performance, we averaged each metric across the five outer-CV folds (Step 9, Figure 1b).

2.7 | Model for predicting vulnerability to sleep loss

To predict the probability of a given individual being vulnerable to sleep loss, we created a second set of classifiers following the same procedure as the one described above for the resilient versus not-resilient classifier. Namely, we labelled each subject as vulnerable or not-vulnerable, where the not-vulnerable class consisted of resilient and intermediate subjects. When computing performance metrics, we treated the vulnerable subjects as the positive class.

2.8 | Statistical analysis

For all performance metrics, we computed the mean and standard deviation values across the five outer-CV folds. To compute the confidence interval for the mean AUC value for the resilient versus not-resilient classification, or the vulnerable versus not-vulnerable classification, we aggregated the predicted class probabilities of the five outer-CV fold

TABLE 2 EEG feature subsets selected for the final resilient versus not-resilient classification models for each of the five outer cross-validation (CV) folds.

	EEG feature								
	SWA power				SWA rise rate				
CV test study	C3	C4	O1	O2	C3	C4	O1	O2	Feature subset size
1	yes	no	no	no	no	no	yes	no	2
2	yes	no	no	no	no	no	yes	no	2
3	yes	no	no	no	no	no	yes	no	2
4	yes	no	no	no	yes	no	yes	no	3
5	yes	no	yes	yes	yes	yes	yes	no	6
Feature inclusion frequency	5	0	1	1	2	1	5	0	

Note: Entries with a "yes" identify the most informative features selected to predict the sleep-loss phenotype of subjects of the corresponding outer-CV test study.
Abbreviations: EEG, electroencephalography; SWA, slow-wave activity.

test datasets and computed the 95% confidence interval using an analytic method described previously (Hanley & McNeil, 1982).

3 | RESULTS

3.1 | Classification of resilient versus not-resilient individuals

3.1.1 | Model selection

To select models with minimal complexity, we minimised overfitting and selected the most informative EEG feature subsets for each of the five outer-CV folds. Table 2 shows that the feature selection procedure selected the SWA power and SWA rise rate (from one or more channels) in each of the five outer-CV folds, suggesting that both of these SWA features provide discriminative information. Specifically, the procedure repeatedly selected the features C3-SWA power and O1-SWA rise rate in each of the five outer-CV folds, suggesting that these features are consistent indicators of resilience to sleep loss. In contrast, the procedure did not select features C4-SWA power or O2-SWA rise rate in any of the five outer-CV folds, indicating that these features are not discriminatory. The remaining four EEG features appeared mostly as part of the feature subset of only one of the outer-CV folds in the testing of Study 5. Taken together, these results suggest that, out of the eight evaluated EEG features, primarily two EEG features associated with channels C3 or O1 were the most informative ones for predicting the probability of an individual being resilient to sleep loss.

3.1.2 | Model evaluation

Table 3 shows the classifier performance values averaged across the five outer-CV folds for the training (Step 6, Figure 1b) and testing (Step 7, Figure 1b) procedures, where we used a 0.50 threshold for the binary classification (resilient versus not-resilient). The training and testing accuracies were comparable, suggesting that model overfitting was minimal. The testing AUC was 0.68 (Figure 2a) and its 95% confidence interval (0.56, 0.80) excluded 0.50, suggesting that the model performed better than a classifier that randomly assigned test subjects into either class. The testing accuracy was 0.72, which was higher than the random chance-level performance of 0.50. The model had a modest testing sensitivity (0.50) but high specificity (0.84). These results indicate that while the model correctly classified only half of the resilient subjects as such, it is less likely to classify an individual who is not resilient as resilient, which is desirable for mission-critical task assignments where it is important not to identify a more vulnerable individual as resilient. The testing likelihood ratio was 3.59, suggesting that the model was at least three times more likely to identify a resilient subject as resilient than a subject who was not resilient as resilient. The model's testing PPV was 0.61, indicating that out of all the testing subjects whom the model predicted to be resilient, 61% of them were truly resilient while 39% were falsely predicted to be resilient. For assigning individuals to critical tasks requiring a high level of alertness under limited sleep, it is essential to identify resilient individuals with more certainty. To this end, we computed the PPV for a range of classification thresholds and found that with a threshold of 0.59, the testing PPV increased to 0.81 (Table 3), resulting in an improved probability that an individual predicted to be resilient was truly resilient. Although at this alternative threshold the testing sensitivity decreased, the testing accuracy remained nearly the same, and the testing specificity increased. The likelihood ratio decreased, but it was based on only one of the five studies (see Table 3). Overall, these results suggest that the logistic regression classifier can be used to predict the probability of resilience to sleep loss.

3.2 | Classification of vulnerable versus not-vulnerable individuals

3.2.1 | Model selection

Table 4 shows the EEG feature subsets selected in the vulnerable versus not-vulnerable classification models for each of the five outer-CV

TABLE 3 Performance summary of the resilient versus not-resilient logistic regression classifiers, averaged over the five outer cross-validation (CV) folds.

	Performance [mean (1 SD)]			
	Threshold of 0.50		Threshold of 0.59	
Metric	Training	Testing	Training	Testing
AUC	0.72 (0.05)	0.68 (0.10)	0.72 (0.05)	0.68 (0.10)
Accuracy	0.71 (0.04)	0.72 (0.09)	0.73 (0.03)	0.71 (0.11)
Sensitivity	0.35 (0.12)	0.50 (0.20)	0.24 (0.12)	0.25 (0.17)
Specificity	0.90 (0.02)	0.84 (0.09)	0.98 (0.02)	0.95 (0.10)
Positive predictive value	0.63 (0.06)	0.61 (0.13)	0.91 (0.07)	0.81 (0.32)[a]
Likelihood ratio	3.40 (0.83)	3.59 (1.73)	11.12 (1.51)[b]	0.57 (–)[c]

Abbreviations: AUC, area under the receiver operating characteristic curve; SD, standard deviation.
[a]Positive predictive values (PPVs) were averaged across four outer-CV folds because PPV was undefined for one of the outer-CV folds (due to a division of zero by zero).
[b]Likelihood ratio (LR) values were infinity for two outer-CV folds (due to a division by zero).
[c]LR values were infinity for three outer-CV folds (due to a division by zero) and undefined for one outer-CV fold (due to a division of zero by zero).

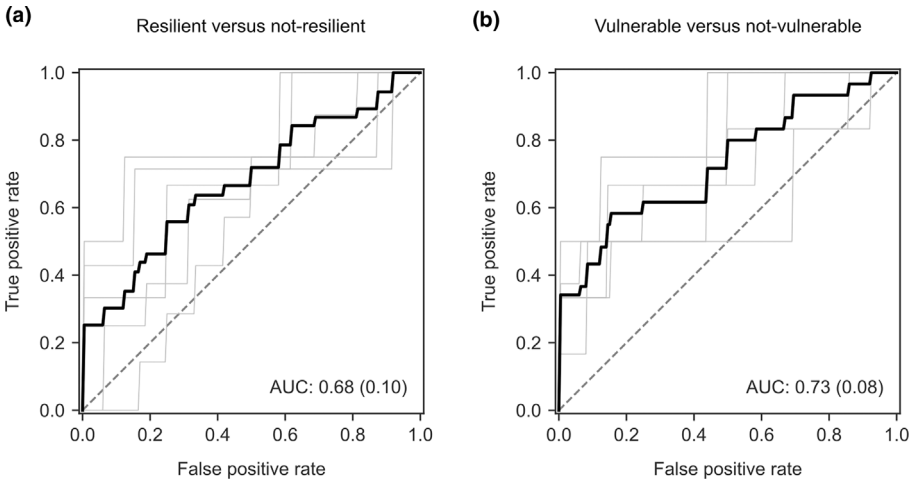

FIGURE 2 Model performance.
(a) Performance of the resilient versus not-resilient classifiers. The grey continuous traces indicate the receiver operating characteristic (ROC) curves obtained from the testing procedure in the five outer-CV folds, and the black continuous trace denotes the mean ROC curve. The dotted grey trace indicates the performance of a random classifier. AUC, area under the mean ROC curve. Value in parentheses denotes one standard deviation.
(b) Performance of the vulnerable versus not-vulnerable classifiers.

TABLE 4 EEG feature subsets selected for the final vulnerable versus not-vulnerable classification models for each of the five outer cross-validation (CV) folds.

	EEG feature								
	SWA power				SWA rise rate				
CV test study	C3	C4	O1	O2	C3	C4	O1	O2	Feature subset size
1	no	no	yes	yes	yes	no	no	no	3
2	no	yes	yes	no	yes	no	no	no	3
3	no	no	yes	no	yes	no	yes	no	3
4	no	no	yes	no	yes	no	yes	no	3
5	yes	no	no	no	no	no	yes	no	2
Feature inclusion frequency	1	1	4	1	4	0	3	0	

Note: Entries with a "yes" identify the most informative features selected to predict the sleep-loss phenotype of subjects of the corresponding outer-CV test study.
Abbreviations: EEG, electroencephalography; SWA, slow-wave activity.

folds. The results showed that the procedure selected SWA power and SWA rise rate in each of the five outer-CV folds, suggesting that both of these SWA features provide discriminative information. Specifically, the procedure repeatedly selected features O1-SWA power, C3-SWA rise rate, or O1-SWA rise rate at least three times, suggesting that these features are informative in predicting

TABLE 5 Performance summary of the vulnerable versus not-vulnerable logistic regression classifiers, averaged over the five outer cross-validation (CV) folds.

	Performance [mean (1 SD)]			
	Threshold of 0.50		Threshold of 0.69	
Metric	Training	Testing	Training	Testing
AUC	0.77 (0.01)	0.73 (0.08)	0.77 (0.01)	0.73 (0.08)
Accuracy	0.78 (0.02)	0.78 (0.04)	0.75 (0.02)	0.73 (0.03)
Sensitivity	0.44 (0.03)	0.41 (0.17)	0.24 (0.06)	0.15 (0.08)
Specificity	0.94 (0.01)	0.95 (0.06)	1.00 (0.00)	1.00 (0.00)
Positive predictive value	0.77 (0.05)	0.87 (0.16)	1.00 (0.00)	1.00 (0.00)[a]
Likelihood ratio	7.44 (2.07)	4.33 (0.33)[b]	[c]	[c]

Abbreviations: AUC, area under the receiver operating characteristic curve; SD, standard deviation.
[a]Positive predictive values (PPVs) were averaged across four outer-CV folds because PPV was undefined for one of the outer-CV folds (due to a division of zero by zero).
[b]Likelihood ratio (LR) values were infinity for three outer-CV folds (due to a division by zero).
[c]LR values were infinity for four outer-CV folds (due to a division by zero) and undefined for one outer-CV fold (due to a division of zero by zero).

vulnerability to sleep loss. In contrast, the procedure did not select features C4-SWA rise rate and O2-SWA rise rate in any of the CV folds, indicating that these features are not discriminatory. The remaining three features each appeared once in different CV folds, suggesting that these features are needed to account for study-to-study variations in how vulnerability to sleep loss is manifested in EEG signals. Taken together, these results suggest that, out of the eight evaluated EEG features, three features were the most informative ones for predicting vulnerability to sleep loss, and these features were associated almost exclusively with channels C3 or O1.

3.2.2 | Model evaluation

Table 5 shows the classifier performance values averaged across the five outer-CV folds for the training (Step 6, Figure 1b) and testing (Step 7, Figure 1b) procedures, where we used a 0.50 threshold for the binary classification (vulnerable versus not-vulnerable). The training and testing accuracies were comparable, indicating minimal model overfitting. The testing AUC (Figure 2b) was 0.72 with its 95% confidence interval (0.60, 0.83) excluding 0.50, which suggested that the model performed better than a classifier that randomly assigned test subjects into either class. The testing accuracy was 0.78, which was higher than the random chance-level performance of 0.50. The testing sensitivity was modest (0.41), whereas the specificity was high (0.95). These values suggest that although our model correctly classified fewer than half of the vulnerable subjects, it rarely classified a subject who was not vulnerable as vulnerable. The testing likelihood ratio was 4.33, indicating that the model was at least four times more likely to classify a vulnerable individual as vulnerable than an individual who was not vulnerable as vulnerable. (The likelihood ratio of 4.33 is an underestimate of the true value because in three studies it reached positive infinity; Table 5). The testing PPV was 0.87, suggesting a high probability of a subject classified as vulnerable being truly vulnerable. This PPV value can be further increased, if necessary. For example, this may be required when we need to identify vulnerable individuals with high certainty, either to exclude them from performing critical tasks or to offer them sleep-loss countermeasures. By investigating a range of classification thresholds, we found that at a threshold value of 0.69, the PPV increased to 1.0 (Table 5), indicating high certainty that an individual predicted to be vulnerable was truly vulnerable. At this higher threshold value, the testing sensitivity decreased substantially, however, the testing accuracy decreased only slightly, and the testing specificity increased (Table 5). We could not compute the likelihood ratio as its calculation involved division by zero in all five test studies (Table 5). Taken together, these results show that the logistic regression classifier can be used to predict vulnerability to sleep loss using routine night sleep EEG data.

4 | DISCUSSION

Using EEG signals collected during routine night sleep, we developed two sets of classifiers to predict the probability that an individual is resilient or vulnerable to sleep loss. Both of these predictions consistently required two or three features extracted from the central (C3) or occipital (O1) EEG channels located over the left cerebral hemisphere (Figure 1a). We estimated the performance of these models conservatively by using unique study data not involved in model selection as test sets, so that when our models are used in real-world applications their performance should be similar to those reported here. With the ability to place an individual on a continuous scale of resilience or vulnerability to sleep loss, these classifiers can help identify resilient individuals with the required sleep-loss phenotype for certain mission-critical tasks, and vulnerable individuals who may need additional sleep-loss countermeasures.

When building the classifiers to predict resilience or vulnerability to sleep loss, we used a five-fold nested cross-validation procedure. This resulted in five different models, each with its own EEG feature subset and model parameters. Hence, for future real-world

applications, an ensemble of these five models should be used as the "final" model, where the respective EEG feature subsets of a test subject are given to these models as inputs, and the model outputs are averaged across the five models to produce a single value of class probability. Although feature selection results indicated that the most informative features primarily originated from channels located on the left cerebral hemisphere (C3 and O1), the ensemble model for predicting resilience or vulnerability to sleep loss would also include features from channels on the right hemisphere (C4 and O2). Nevertheless, the overall result is encouraging and warrants further investigation into hemisphere asymmetry in predicting sleep-loss phenotype.

The performance results of our models were comparable to those reported in previous studies. Model accuracy (0.72 for the resilient versus not-resilient and 0.78 for the vulnerable versus not-vulnerable) was similar to that of PVT-based models (0.69–0.71) (Chua et al., 2019; Patanaik et al., 2014, 2015; St. Hilaire et al., 2019) and MRI-based models (0.60–0.85) (Xu et al., 2021; Yeo et al., 2015). Only a subset of previous modelling studies (Patanaik et al., 2014, 2015; Xu et al., 2021) reported the AUC values of their models. Nevertheless, the AUC values of our models (0.68 for the resilient versus not-resilient and 0.72 for the vulnerable versus not-vulnerable) were similar to those based on PVTs (0.74) (Patanaik et al., 2014, 2015), but not as high as that based on brain imaging (0.94) (Xu et al., 2021). However, some of these studies (Chua et al., 2019; Patanaik et al., 2014; Xu et al., 2021) did not use independent datasets for model assessment, which likely biased their performance results.

While our model-performance values were comparable to those of previous efforts, the overall performance was still moderate. In an attempt to improve the results, we fitted more complex models, such as support vector machines with both linear and non-linear kernels, however, the performance did not improve. Another potentially beneficial approach, which we could not investigate due to limitations in our data, is to evaluate a broader range of EEG channels located on other areas of the scalp. Specifically, given that resilient and vulnerable individuals likely differ in accumulated sleep pressure (Subramaniyan et al., 2023), one possibility is to include SWA features from frontal EEG channels, which are known to reflect sleep pressure more reliably than other EEG recording locations (Cajochen et al., 1999; Finelli et al., 2001; Munch et al., 2004).

An unexpected observation in our study was that the most informative EEG features were largely associated with EEG channels located over the left cerebral hemisphere, suggesting that the ability to discriminate between resilient and vulnerable individuals primarily resides in the left hemisphere of the brain. Interestingly, a few studies have shown that sleep pressure results in higher spectral power in the low frequency range (overlapping with the SWA frequency band) in the left hemisphere as compared with the right hemisphere (Achermann et al., 2001; Ferrara et al., 2002; Vyazovskiy et al., 2002). Given that the two features used in our study, SWA power and its rise rate, are markers of sleep pressure (Brunner et al., 1993; Dijk et al., 1990), these observations suggest that sleep-loss phenotypes differ in their accumulated sleep pressure under baseline conditions, and that this difference is more pronounced in the left hemisphere. Furthermore, our results are in line with those of brain imaging studies involving working memory tasks which find that, under rested wakefulness, task-related activation (Chee et al., 2006; Cui et al., 2015; Mu et al., 2005) and microarchitecture characteristics (Cui et al., 2015) measured in the left hemisphere are more discriminative of an individual's sleep-loss phenotype than those measured in the right hemisphere. However, our findings are intriguing because we obtained the same hemisphere asymmetry in discriminatory power based on sleep EEG features (i.e., without involving the subjects in specific tasks). We speculate that, during wakefulness, resilient and vulnerable individuals differ in the extent to which they engage their left hemispheres, resulting in a corresponding difference in the recuperative slow-wave activity during sleep, potentially explaining the higher discriminatory power of the left hemisphere. In addition, the right handedness that was predominant in our study population (94% in Studies 1, 4, and 5 combined; we did not have this information for Studies 2 and 3) could have led to the preferential engagement of the left hemisphere, further augmenting the observed asymmetry in the discriminative power. However, future studies involving a sufficient number of left-handed subjects are required to assess the contribution of handedness to sleep-loss phenotype discrimination.

5 | REAL-WORLD USE AND APPLICABILITY CONSIDERATIONS

A typical real-world use of our models would involve the screening of civilian and military personnel to determine whether they are suitable for shift work, which requires sustained vigilance and attention. In this context, the following points should be considered. First, EEG signals should be recorded for at least two nights from the central (C3 and C4) and occipital (O1 and O2) channels until the end of the first sleep cycle. Although the median duration of the first sleep cycle is only 1.6 h, a relatively longer duration (~4 h) would be advisable so as to include most individuals, unless sleep cycles are detected in real time. Second, the quality of the recordings should be such that all sleep stages (N1, N2, N3, and REM) can be determined and be free of artefacts for at least ~80% of the time. Third, EEG-measuring devices often differ in the absolute magnitude of the recorded signals. This device-to-device variability by itself does not pose a problem when using our models because the EEG features are *z*-scored across individuals (which discards absolute signal magnitude information) before being used as inputs to the models. However, when testing a group of individuals, it is essential to use the same type of EEG device for all individuals to avoid potential device-dependent characteristics that could affect the signal magnitude used to discriminate sleep-loss phenotypes. Finally, due to *z*-scoring of the EEG features, the models' predictions should be treated as a relative ranking of individuals within a group with regard to vulnerability or resilience to sleep loss, rather than an absolute sleep-loss phenotypic labelling. For example, the vulnerable versus not-vulnerable model will rank an individual more or less vulnerable depending on the EEG feature values of the remaining individuals in the group.

With the rapid progress in the development of consumer-level wireless EEG systems, collecting EEG data at home is becoming more practical and accessible (Chinoy et al., 2021; Wood et al., 2023). Therefore, an important question becomes whether the models developed here using data collected with laboratory-grade EEG devices would be different if we used data collected with at-home, consumer-level EEG devices. Laboratory EEG data could be different from at-home-collected EEG data because sleep in an unfamiliar laboratory environment is known to affect sleep architecture (Iber et al., 2004) and SWA power (Mayeli et al., 2022). However, such effects are removed by *z*-scoring of the EEG features. Moreover, the effects of a novel laboratory environment on SWA power (on which our EEG features are based) should be minimal, if any, because such effects are restricted to the frontal regions of the brain and are not observed in the central and occipital regions (Mayeli et al., 2022) from which we obtained our features. Therefore, at-home and laboratory EEG data should give rise to similar model predictions.

There are practical advantages to using sleep EEG to predict sleep-loss phenotypes as compared with brain imaging- or PVT-based methods. Brain-imaging methods (Xu et al., 2021; Yeo et al., 2015) are not practical due to their low accessibility and high cost. In contrast, PVT-based methods, which use one (Chua et al., 2019; St. Hilaire et al., 2019) or two (Patanaik et al., 2015) 10 min PVTs, offer the advantage of quick and simple data collection. However, PVTs are highly dependent on the individual's level of effort (Brewer et al., 2017; Massar et al., 2016; Robison et al., 2021), which makes them a less-desirable method for the assessment of sleep-loss phenotype. Furthermore, the predictive ability of PVT-based models changes depending on the time of day of test administration (Chua et al., 2019), necessitating the additional constraint of testing all individuals at the same time. While high-quality sleep EEG data collection is less practical than PVTs, it does not disrupt an individual's daytime activities and is not influenced by subjective factors. Moreover, the increasing availability of easy-to-use EEG recording systems will make EEG-based predictions practical, reliable, and accessible.

6 | LIMITATIONS

Our study has limitations. First, for the majority of the subjects, we combined baseline EEG features from two nights. Therefore, we do not know whether our results generalise to single-night measurements. Second, our datasets originated from studies that used different sleep-loss protocols (TSD and CSR), making it unclear if the sleep-loss phenotypic groups of the different studies are equivalent. However, it is well documented that sleep-loss phenotype is trait-like, irrespective of the sleep-loss challenge (Rupp et al., 2012). Third, given that we needed to perform within-study *z*-scoring, our model is limited to identifying the sleep-loss phenotype of an individual within a group. However, in most real-world settings, such as in selecting resilient individuals in the Armed Forces, individuals will often need to be selected from a group. Therefore, our models are applicable in those settings. Finally, we evaluated a limited number of EEG channels to build our classifiers. Hence, we do not know if other EEG channels are more discriminatory and therefore more suitable for building sleep-loss phenotype classifiers. Nevertheless, the EEG features from the channels we used had sufficient discriminatory information to classify individuals with reasonable accuracy, motivating further investigations to identify additional discriminative features.

7 | CONCLUSIONS

In this work, we developed two sets of classifiers based on EEG features obtained from routine night sleep and predicted the probability of an individual being resilient or vulnerable to sleep loss. We built these classifiers using a nested cross-validation procedure that allowed us to simultaneously minimise model overfitting and to obtain performance results that are representative of the models' performance in real-world usage. Our results indicate that the ability to discriminate phenotypic responses to sleep loss only requires data collection during the first sleep cycle, potentially simplifying the data collection process. In addition to being more practical and scalable, our model-building approach provides a framework to investigate additional features that could further improve the ability to predict an individual's sleep-loss phenotype.

AUTHOR CONTRIBUTIONS

Jaques Reifman: Conceptualization; methodology; writing – original draft; writing – review and editing; project administration. **Manivannan Subramaniyan:** Conceptualization; methodology; software; data curation; writing – original draft; writing – review and editing; formal analysis. **John D. Hughes:** Data curation; investigation; writing – review and editing. **Tracy J. Doty:** Data curation; investigation; writing – review and editing. **William D. S. Killgore:** Data curation; investigation; writing – review and editing.

ACKNOWLEDGEMENTS

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

CONFLICT OF INTEREST STATEMENT

This was not an industry-supported study. The authors declare no conflicts of interest. The opinions and assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the U.S. Army, the U.S. Department of Defense, or The Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc. This paper has been approved for public release with unlimited distribution. Material has been reviewed by the Walter Reed Army Institute of Research. There is no objection to its presentation and/or publication. The investigators have adhered to the policies for protection of human subjects as prescribed in AR 70–25.

DATA AVAILABILITY STATEMENT

The data used in this article cannot be shared publicly because the protocols and informed consent documents for these older, legacy studies do not support data sharing outside of entities defined in the original documentation.

ORCID

Tracy J. Doty https://orcid.org/0000-0001-9921-2457

REFERENCES

Achermann, P., Finelli, L. A., & Borbély, A. A. (2001). Unihemispheric enhancement of delta power in human frontal sleep EEG by prolonged wakefulness. *Brain Research*, *913*(2), 220–223.

Agnew, H. W., Jr., Webb, W. B., & Williams, R. L. (1966). The first night effect: An EEG study of sleep. *Psychophysiology*, *2*(3), 263–266.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

Brewer, G. A., Lau, K. K. H., Wingert, K. M., Ball, B. H., & Blais, C. (2017). Examining depletion theories under conditions of within-task transfer. *Journal of Experimental Psychology. General*, *146*(7), 988–1008.

Brunner, D. P., Dijk, D. J., & Borbély, A. A. (1993). Repeated partial sleep deprivation progressively changes in EEG during sleep and wakefulness. *Sleep*, *16*(2), 100–113.

Cajochen, C., Foy, R., & Dijk, D. J. (1999). Frontal predominance of a relative increase in sleep delta and theta EEG activity after sleep loss in humans. *Sleep Research Online*, *2*(3), 65–69.

Caldwell, J. A., Mu, Q., Smith, J. K., Mishory, A., Caldwell, J. L., Peters, G., Brown, D. L., & George, M. S. (2005). Are individual differences in fatigue vulnerability related to baseline differences in cortical activation? *Behavioral Neuroscience*, *119*(3), 694–707.

Chattu, V. K., Manzar, M. D., Kumary, S., Burman, D., Spence, D. W., & Pandi-Perumal, S. R. (2019). The global problem of insufficient sleep and its serious public health implications. *Healthcare (Basel)*, *7*(1), 1.

Chee, M. W. L., Chuah, L. Y. M., Venkatraman, V., Chan, W. Y., Philip, P., & Dinges, D. F. (2006). Functional imaging of working memory following normal sleep and after 24 and 35 h of sleep deprivation: Correlations of fronto-parietal activation with performance. *NeuroImage*, *31*(1), 419–428.

Chinoy, E. D., Cuellar, J. A., Huwa, K. E., Jameson, J. T., Watson, C. H., Bessman, S. C., Hirsch, D. A., Cooper, A. D., Drummond, S. P. A., & Markwald, R. R. (2021). Performance of seven consumer sleep-tracking devices compared with polysomnography. *Sleep*, *44*(5), zsaa291.

Chua, E. C., Yeo, S. C., Lee, I. T., Tan, L. C., Lau, P., Cai, S., Zhang, X., Puvanendran, K., & Gooley, J. J. (2014). Sustained attention performance during sleep deprivation associates with instability in behavior and physiologic measures at baseline. *Sleep*, *37*(1), 27–39.

Chua, E. C.-P., Sullivan, J. P., Duffy, J. F., Klerman, E. B., Lockley, S. W., Kristal, B. S., Czeisler, C. A., & Gooley, J. J. (2019). Classifying attentional vulnerability to total sleep deprivation using baseline features of psychomotor vigilance test performance. *Scientific Reports*, *9*(1), 12102.

Cui, J., Tkachenko, O., Gogel, H., Kipman, M., Preer, L. A., Weber, M., Divatia, S. C., Demers, L. A., Olson, E. A., Buchholz, J. L., Bark, J. S., Rosso, I. M., Rauch, S. L., & Killgore, W. D. S. (2015). Microstructure of frontoparietal connections predicts individual resistance to sleep deprivation. *NeuroImage*, *106*, 123–133.

Dijk, D. J., Brunner, D. P., & Borbély, A. A. (1990). Time course of EEG power density during long sleep in humans. *American Journal of Physiology. Regulatory, Integrative and Comparative Physiology*, *258*(3), R650–R661.

Doty, T. J., So, C. J., Bergman, E. M., Trach, S. K., Ratcliffe, R. H., Yarnell, A. M., Capaldi, V. F., II, Moon, J. E., Balkin, T. J., & Quartana, P. J. (2017). Limited efficacy of caffeine and recovery costs during and following 5 days of chronic sleep restriction. *Sleep*, *40*(12), zsx171.

Erickson, E. A., Stahlman, S., & McNellis, M. G. (2017). Insomnia and motor vehicle accident-related injuries, active component, U.S. armed forces, 2007-2016. *MSMR*, *24*(12), 2–11.

Ferrara, M., De Gennaro, L., Curcio, G., Cristiani, R., & Bertini, M. (2002). Interhemispheric asymmetry of human sleep EEG in response to selective slow-wave sleep deprivation. *Behavioral Neuroscience*, *116*(6), 976–981.

Finelli, L. A., Borbély, A. A., & Achermann, P. (2001). Functional topography of the human nonREM sleep electroencephalogram. *The European Journal of Neuroscience*, *13*(12), 2282–2290.

Galli, O., Jones, C. W., Larson, O., Basner, M., & Dinges, D. F. (2022). Predictors of interindividual differences in vulnerability to neurobehavioral consequences of chronic partial sleep restriction. *Sleep*, *45*(1), zsab278.

Good, C. H., Brager, A. J., Capaldi, V. F., & Mysliwiec, V. (2020). Sleep in the United States military. *Neuropsychopharmacology*, *45*(1), 176–191.

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, *143*(1), 29–36.

Hansen, D. A., Ramakrishnan, S., Satterfield, B. C., Wesensten, N. J., Layton, M. E., Reifman, J., & van Dongen, H. P. A. (2019). Randomized, double-blind, placebo-controlled, crossover study of the effects of repeated-dose caffeine on neurobehavioral performance during 48 h of total sleep deprivation. *Psychopharmacology*, *236*(4), 1313–1322.

Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. Springer.

Iber, C., Redline, S., Gilpin, A. M. K., Quan, S. F., Zhang, L., Gottlieb, D. J., Rapoport, D., Resnick, H. E., Sanders, M., & Smith, P. (2004). Polysomnography performed in the unattended home versus the attended laboratory setting – sleep heart health study methodology. *Sleep*, *27*(3), 536–540.

Institute of Medicine. (2006). Functional and economic impact of sleep loss and sleep-related disorders. In H. R. Colten & B. M. Altevogt (Eds.), *Sleep disorders and sleep deprivation: An unmet public health problem* (pp. 137–172). The National Academies Press.

Massar, S. A. A., Lim, J., Sasmita, K., & Chee, M. W. L. (2016). Rewards boost sustained attention through higher effort: A value-based decision making approach. *Biological Psychology*, *120*, 21–27.

Mayeli, A., Janssen, S. A., Sharma, K., & Ferrarelli, F. (2022). Examining first night effect on sleep parameters with hd-EEG in healthy individuals. *Brain Sciences*, *12*(2), 233.

Mu, Q., Mishory, A., Johnson, K. A., Nahas, Z., Kozel, F. A., Yamanaka, K., Bohning, D. E., & George, M. S. (2005). Decreased brain activation during a working memory task at rested baseline is associated with vulnerability to sleep deprivation. *Sleep*, *28*(4), 433–448.

Münch, M., Knoblauch, V., Blatter, K., Schröder, C., Schnitzler, C., Kräuchi, K., Wirz-Justice, A., & Cajochen, C. (2004). The frontal predominance in human EEG delta activity after sleep loss decreases with age. *The European Journal of Neuroscience*, *20*(5), 1402–1410.

Patanaik, A., Kwoh, C. K., Chua, E. C., Gooley, J. J., & Chee, M. W. (2015). Classifying vulnerability to sleep deprivation using baseline measures of psychomotor vigilance. *Sleep*, *38*(5), 723–734.

Patanaik, A., Zagorodnov, V., Kwoh, C. K., & Chee, M. W. (2014). Predicting vulnerability to sleep deprivation using diffusion model parameters. *Journal of Sleep Research*, *23*(5), 576–584.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Reifman, J., Ramakrishnan, S., Liu, J., Kapela, A., Doty, T. J., Balkin, T. J., Kumar, K., & Khitrov, M. Y. (2019). *2B-alert* app: A mobile application

for real-time individualized prediction of alertness. *Journal of Sleep Research*, *28*(2), e12725.

Robison, M. K., Unsworth, N., & Brewer, G. A. (2021). Examining the effects of goal-setting, feedback, and incentives on sustained attention. *Journal of Experimental Psychology. Human Perception and Performance*, *47*(6), 869–891.

Rocklage, M., Williams, V., Pacheco, J., & Schnyer, D. M. (2009). White matter differences predict cognitive vulnerability to sleep deprivation. *Sleep*, *32*(8), 1100–1103.

Rupp, T. L., Wesensten, N. J., & Balkin, T. J. (2012). Trait-like vulnerability to total and partial sleep loss. *Sleep*, *35*(8), 1163–1172.

Silber, M. H., Ancoli-Israel, S., Bonnet, M. H., Chokroverty, S., Grigg-Damberger, M. M., Hirshkowitz, M., Kapen, S., Keenan, S. A., Kryger, M. H., Penzel, T., Pressman, M. R., & Iber, C. (2007). The visual scoring of sleep in adults. *Journal of Clinical Sleep Medicine*, *3*(2), 121–131.

Sprecher, K. E., Riedner, B. A., Smith, R. F., Tononi, G., Davidson, R. J., & Benca, R. M. (2016). High resolution topography of age-related changes in non-rapid eye movement sleep electroencephalography. *PLoS One*, *11*(2), e0149770.

St. Hilaire, M. A., Kristal, B. S., Rahman, S. A., Sullivan, J. P., Quackenbush, J., Duffy, J. F., Barger, L. K., Gooley, J. J., Czeisler, C. A., & Lockley, S. W. (2019). Using a single daytime performance test to identify Most individuals at high-risk for performance impairment during extended wake. *Scientific Reports*, *9*(1), 16681.

Subramaniyan, M., Wang, C., Laxminarayan, S., Vital-Lopez, F. G., Hughes, J. D., Doty, T. J., & Reifman, J. (2023). Electroencephalographic markers from routine sleep discriminate individuals who are vulnerable or resilient to sleep loss. *Journal of Sleep Research*, e14060.

Tkachenko, O., & Dinges, D. F. (2018). Interindividual variability in neurobehavioral response to sleep loss: A comprehensive review. *Neuroscience and Biobehavioral Reviews*, *89*, 29–48.

Van Dongen, H. P. A., 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*(3), 423–433.

Varma, S., & Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics*, *7*(1), 91.

Vital-Lopez, F. G., Doty, T. J., Anlap, I., Killgore, W. D. S., & Reifman, J. (2023). *2B-alert* app 2.0: Personalized caffeine recommendations for optimal alertness. *Sleep*, *46*(7), zsad080.

Vyazovskiy, V. V., Borbély, A. A., & Tobler, I. (2002). Interhemispheric sleep EEG asymmetry in the rat is enhanced by sleep deprivation. *Journal of Neurophysiology*, *88*(5), 2280–2286.

Wang, C., Ramakrishnan, S., Laxminarayan, S., Dovzhenok, A., Cashmere, J. D., Germain, A., & Reifman, J. (2020). An attempt to identify reproducible high-density EEG markers of PTSD during sleep. *Sleep*, *43*(1), zsz207.

Webb, M. P. K., & Sidebotham, D. (2020). Bayes' formula: A powerful but counterintuitive tool for medical decision-making. *BJA Education*, *20*(6), 208–213.

Wood, E., Westphal, J. K., & Lerner, I. (2023). Re-evaluating two popular EEG-based mobile sleep-monitoring devices for home use. *Journal of Sleep Research*, *32*(5), e13824.

Xu, Y., Yu, P., Zheng, J., Wang, C., Hu, T., Yang, Q., Xu, Z., Guo, F., Tang, X., Ren, F., & Zhu, Y. (2021). Classifying vulnerability to sleep deprivation using resting-state functional MRI graph theory metrics. *Frontiers in Neuroscience*, *15*, 660365.

Yamazaki, E. M., Rosendahl-Garcia, K. M., Casale, C. E., MacMullen, L. E., Ecker, A. J., Kirkpatrick, J. N., & Goel, N. (2021). Left ventricular ejection time measured by echocardiography differentiates neurobehavioral resilience and vulnerability to sleep loss and stress. *Frontiers in Physiology*, *12*, 795321.

Yeo, B. T. T., Tandi, J., & Chee, M. W. L. (2015). Functional connectivity during rested wakefulness predicts vulnerability to sleep deprivation. *NeuroImage*, *111*, 147–158.

Zhao, R., Zhang, X., Zhu, Y., Fei, N., Sun, J., Liu, P., Yang, X., & Qin, W. (2018). Prediction of the effect of sleep deprivation on response inhibition via machine learning on structural magnetic resonance imaging data. *Frontiers in Human Neuroscience*, *12*, 276.

How to cite this article: Subramaniyan, M., Hughes, J. D., Doty, T. J., Killgore, W. D. S., & Reifman, J. (2024). Individualised prediction of resilience and vulnerability to sleep loss using EEG features. *Journal of Sleep Research*, *33*(6), e14220. https://doi.org/10.1111/jsr.14220