

Review

# Biomarker-Informed Machine Learning Model of Cognitive Fatigue from a Heart Rate Response Perspective

Kar Fye Alvin Lee <sup>1,\*</sup>, Woon-Seng Gan <sup>1</sup>  and Georgios Christopoulos <sup>2</sup> 

<sup>1</sup> Smart Nation Translational Laboratory, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, Singapore; ewsgan@ntu.edu.sg

<sup>2</sup> Decision, Environmental and Organizational Neuroscience Lab (DeonLab), Nanyang Business School, Nanyang Technological University, Singapore 639798, Singapore; CGeorgios@ntu.edu.sg

\* Correspondence: alvin.lee@ntu.edu.sg

**Abstract:** Cognitive fatigue is a psychological state characterised by feelings of tiredness and impaired cognitive functioning arising from high cognitive demands. This paper examines the recent research progress on the assessment of cognitive fatigue and provides informed recommendations for future research. Traditionally, cognitive fatigue is introspectively assessed through self-report or objectively inferred from a decline in behavioural performance. However, more recently, researchers have attempted to explore the biological underpinnings of cognitive fatigue to understand and measure this phenomenon. In particular, there is evidence indicating that the imbalance between sympathetic and parasympathetic nervous activity appears to be a physiological correlate of cognitive fatigue. This imbalance has been indexed through various heart rate variability indices that have also been proposed as putative biomarkers of cognitive fatigue. Moreover, in contrast to traditional inferential methods, there is also a growing research interest in using data-driven approaches to assessing cognitive fatigue. The ubiquity of wearables with the capability to collect large amounts of physiological data appears to be a major facilitator in the growth of data-driven research in this area. Preliminary findings indicate that such large datasets can be used to accurately predict cognitive fatigue through various machine learning approaches. Overall, the potential of combining domain-specific knowledge gained from biomarker research with machine learning approaches should be further explored to build more robust predictive models of cognitive fatigue.

**Keywords:** cognitive fatigue; heart rate variability; biomarker; machine learning



**Citation:** Lee, K.F.A.; Gan, W.-S.; Christopoulos, G. Biomarker-Informed Machine Learning Model of Cognitive Fatigue from a Heart Rate Response Perspective. *Sensors* **2021**, *21*, 3843. <https://doi.org/10.3390/s21113843>

Academic Editor: Marco Di Rienzo

Received: 14 April 2021

Accepted: 21 May 2021

Published: 2 June 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Cognitive fatigue is a psychological state characterised by the subjective feelings of tiredness, insufficient energy, difficulty with concentration, and impaired ability to think [1,2]. This psychological state arises from cognitive “overloading” due to extended periods of sustained performance or cognitively demanding activities [1,3–7]. Attending meetings and report writing during work, as well as childcaring and meal prepping while at home, are examples of such activities. It should be noted that cognitive fatigue is not to be conflated with fatigue that arises from prolonged muscle activity (i.e., physical fatigue), emotional exhaustion (i.e., chronic fatigue) [8], sleep deprivation [9], or boredom [3]. Notably, cognitive fatigue has been demonstrated to have negative effects on executive functions, such as working memory, judgement, and attention [10–13]. Not surprisingly, cognitive fatigue increases the risk of accidents and errors in various mission-critical situations [14,15]. Hence, the ability to accurately assess and monitor cognitive fatigue levels during such situations is imperative in mitigating and minimising the risk of undesirable negative outcomes from occurring. While various methods have been developed to assess cognitive fatigue, these methods have their unique limitations and associated research gaps.

The present paper aims to examine the current research status on the assessment of cognitive fatigue and provide suggestions for prospective researchers regarding the measurement and analysis of cognitive fatigue, with emphasis on biomarker research and machine learning approaches. To this end, we searched for relevant research over the last three decades using various combinations of key terms, such as cognitive fatigue, mental fatigue, self-report, cognitive task, biomarker, heart rate variability (HRV), and machine learning via Onesearch and Google Scholar. Approximately 150 relevant published studies and reviews were identified and qualitatively assessed.

The rest of the paper is organised as follows. First, we describe some traditional psychological assessments of cognitive fatigue. Thereafter, biomarker-based measures of cognitive fatigue are delineated. In particular, we evaluate the imbalance between sympathetic and parasympathetic nervous activity as a potential physiological correlate of cognitive fatigue, and how this can be indexed through HRV. Next, we present in detail recent data-driven approaches in predicting cognitive fatigue through machine learning. Last, we highlight some issues that should be considered when building models of cognitive fatigue. In sum, there are potential benefits in combining knowledge gained from biomarker research with data-driven approaches to build better predictive models of cognitive fatigue.

## 2. Traditional Psychological Assessments of Cognitive Fatigue

Within the psychological literature, several subjective assessments of cognitive fatigue have been developed and validated, such as the Mental Fatigue Scale [16], the Chalder Fatigue Scale [17], and the Fatigue State Questionnaire [18]. These self-assessments of cognitive fatigue are generally questionnaires that employ Likert scale ratings (e.g., “How tired does your mind feel right now? 1. Not at all 2. A little 3. Moderately 4. Very 5. Extremely”) [18]. The scores are then calculated to provide a general index of cognitive fatigue. Notably, such self-report methods presume that respondents have some level of insight or introspection into their cognitive states [19]. Indeed, while useful in gaining introspective knowledge of one’s psychological state, these self-report measures of cognitive fatigue require individuals to be self-aware of their fatigue levels [20]. Unfortunately, individuals often do not have an accurate judgment of their cognitive states [21]. Fatigued individuals have inconsistent self-awareness of their decline in performance [22]. Furthermore, the level of self-awareness of fatigue is moderated by varying working conditions [23]. People may not be able to appreciate how fatigued they are until it is actually “too late”, which can have devastating consequences in critical situations [24,25]. In addition, even brief questionnaires require disruption of current activities to allow time for assessment [26,27] and thus may not be suitable for use in situations that require continuous focus.

Alternatively, cognitive fatigue could also be objectively, but indirectly, inferred from a decrease in cognitive and behavioural performance over time [11,13,28–30]. These performance metrics, such as accuracy and reaction time, are typically measured in the context of computerised versions of cognitive tasks, such as the Stroop task [31] or the Simon task [32]. Previous research has demonstrated that these performance measures are associated with self-reported levels of cognitive fatigue [33,34]. For these performance measures to be used as indices of cognitive fatigue, it is assumed that the decrease in performance is a result of an individual’s impaired ability to maintain optimal task performance due to cognitive “overloading” [11]. However, these objective measures are often task-dependent, and thus the models based around these measures may have limited generalisability across different situations in predicting cognitive fatigue. For instance, Liu and colleagues [35] examining cognitive fatigue found that an increase in reaction time was observed only during an arithmetic task, whereas an increase in error rate was observed only during a switching task. Furthermore, in a low cognitive demanding task, such as a vigilance task, the authors found that neither reaction time nor accuracy could be used to index cognitive fatigue [35]. Interestingly, a recent study found that 16-min dual-tasks were more effective at inducing cognitive fatigue as compared to a 90-min single-task [36]. The same study

also found that participants in the dual-task conditions had lower accuracy as compared to participants in the single-task condition [36]. Future research should examine whether using multitasking measures provide more reliable assessments of cognitive fatigue as compared to single-task measures. It is also important to note that these measures are only able to detect cognitive fatigue after a considerable decline in behavioural performance, which can often be detected too late in critical situations [37].

Despite their usefulness in furthering our theoretical understanding of cognitive fatigue, it appears that both self-assessments and task performance have limitations, in terms of subjectivity, disruptiveness, timeliness, and generalisability that are unsuitable for application beyond research settings. Subjectivity, in this paper, refers to whether the assessment is dependent on an individual's self-evaluation and thus may be influenced by biases, such as socially desirable responses and lack of introspection. Disruptiveness refers to whether current activities must be stopped for a certain period of time for the administration of the assessment (e.g., when examining the effects of cognitive fatigue during driving, participants are required to stop driving for the assessment). Timeliness refers to whether the assessment is made in real-time or if there is a lead time between the point at which the assessment is made and the point at which the results are known. Finally, generalisability refers to whether the assessment can be extended to another individual, group, task, or situation. These four factors are important considerations when evaluating the suitability of applying these assessments in various settings. For instance, workplaces would prioritise subjectivity, disruptiveness, and timeliness over generalisability. A summary of these assessments is described in Table 1.

**Table 1.** Traditional cognitive fatigue assessment examples.

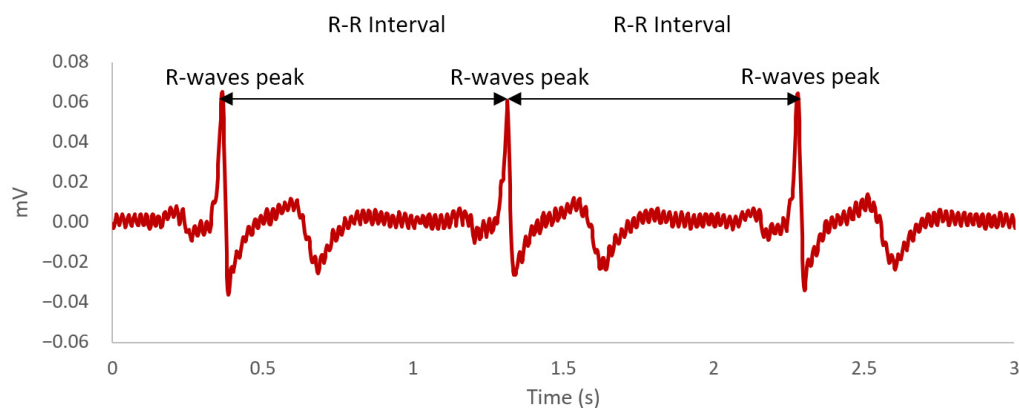
Method	Indicative Reference	Description	Subjectivity	Disruptiveness	Timeliness	Generalisability
<b>Self-report</b>						
Mental Fatigue Scale	[16]	15-items 7-point Likert Assessment of affective, cognitive, and sensory symptoms of fatigue	✓	✓	×	✓
Chalder Fatigue Scale	[17]	11-items 4-point Likert or Bimodal Assessment of physical and cognitive fatigue	✓	✓	×	✓
Fatigue State Questionnaire	[18]	4-items 5-point Likert Assessment of physical and cognitive fatigue	✓	✓	×	✓
<b>Behavioural Performance</b>						
Accuracy/Error Rates	[35,38–40]	Various cognitive Tasks (e.g., Simon task, Stroop task, Switch Task)	×	×	×	×
Reaction Time	[33,35,38,39]	Various Cognitive Tasks (e.g., Simon task, Stroop task, Arithmetic Task)	×	×	×	×
Reaction Time—Intraindividual Variability	[33,39]	Various Cognitive Tasks (e.g., Simon task, Stroop task)	×	×	×	×

Note. Subjectivity refers to whether the measure is dependent on one's self-evaluation of his/her level of cognitive fatigue. Disruptiveness refers to whether time is required to be set aside for assessment. Timeliness refers to whether the assessment of the current state of cognitive fatigue can be made in real-time. Generalisability refers to whether the assessments are comparable in other settings, such as different groups, population, task, or situation. Autonomic Nervous System Biomarkers of Cognitive Fatigue.

### 3. Autonomic Nervous System Biomarkers of Cognitive Fatigue

Over the last decade, there is a growing interest among researchers in identifying potential biological markers of cognitive fatigue [41]. A biological marker, more commonly referred to as biomarker, is defined as “a characteristic objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention” [42] (p. 91). Indeed, there is an accumulation of empirical evidence within the literature suggesting that the autonomic nervous system

is a physiological correlate of cognitive fatigue [43–51]. The autonomic nervous system as part of the peripheral nervous system consists of two components—sympathetic and parasympathetic [52]. The functional significance of the sympathetic nervous system is to prepare the body for physical demands by redirecting oxygen-rich blood to areas of the body where needed, whereas the parasympathetic nervous system is responsible for saving energy for future use as well as regulating bodily functions when the body is at rest [52]. Notably, the parasympathetic nervous system plays an inhibitory-disinhibitory role with the sympathetic nervous system, facilitating the returning of the body to calm states and mobilisation of energy, respectively [53]. Both the sympathetic and parasympathetic nervous systems are regulated by the preganglionic sympathetic and parasympathetic neurons in the central autonomic network [54]. These neurons are linked to the heart through the stellate ganglion and vagus nerve [54]. It is the interaction between sympathetic and parasympathetic neuronal outputs from the central autonomic network on the sinoatrial node of the heart that produces the phenomenon of complex variation in time intervals between heartbeats, which is more commonly known as HRV [55]. HRV is defined as the variation in R-R time intervals on the heartbeat waveform (see Figure 1). Each R-R time interval is measured as the time between each successive heartbeat, indicated by the R-wave peaks on the electrocardiogram [56].



**Figure 1.** An example of a heartbeat waveform across a 3-second time window. The R-wave peaks are the most positive deflection observed in the waveform. The R-R intervals are the time difference between each R-wave peaks.

Besides autonomic nervous regulation, the various components in the central autonomic network are also responsible for facilitating cognitive functions that are key for goal-oriented behaviour and behavioural adaptation [57,58]. Consistent with the overlap in neural structures, previous research has demonstrated a relationship between vagal tone and cognitive functioning, such as working memory and attention [59]. Extensive physiological research has provided supporting evidence for the validity of indexing different aspects of autonomic nervous activity through various HRV components (e.g., [60,61]). These HRV components are derived from various types of analysis, such as time-domain, frequency-domain, and non-linear analyses. Time-domain analysis, in particular, is a form of linear analysis that examines HRV across a specific period of time, whereas frequency-domain analysis (also another form of linear analysis) decomposes HRV signals into various frequency bands [62]. Time-domain analysis is generally more suitable for long-term recordings as it is less influenced by the instability of heart rate modulation, while frequency-domain analysis is more commonly used on short-term recordings due to easier physiological interpretation [61]. By contrast, non-linear analysis purports to quantify the dynamic nature of HRV and thus provides a more accurate representation of the complex interactions amongst various autonomic mechanisms underlying the cardiovascular system [60,62,63]. However, some would argue that it is more difficult to interpret as well as map fundamental autonomic mechanisms to non-linear components [64]. Overall, the

different types of analysis purport to assess the linear and non-linear components of the HRV. Some of the more common indices of HRV are presented in Table 2.

**Table 2.** Major Heart Rate Variability Indices.

Indices	References	Description	Functional Significance			
			Sympathetic Nervous System	Parasympathetic Nervous System	Overall Autonomic Nervous System	Others
<b>Time-Domain</b>						
AVNN	[60]	Average of all normalised R-R intervals				Equivalent to heart rate (inversed)
SDNN	[61]	Standard deviation of all normalised R-R intervals			✓	All cyclic components
RMSSD	[65]	Root mean square of the differences between each successive normalised R-R interval		✓		Respiratory activity
pNN50	[65]	Proportion of normalised R-R intervals that are more than 50 ms from preceding interval		✓		Respiratory activity
<b>Frequency-Domain</b>						
Very-low-frequency power	[66–68]	Spectral power within the very-low-frequency range of 0.003–0.04 Hz		✓		Thermoregulation; renin-angiotensin-aldosterone activity
Low-frequency power	[65,69–72]	Spectral power within the low-frequency range of 0.04–0.15 Hz	✓	✓		Baroreflex activity
High-frequency power	[65,70,71]	Spectral power within the high-frequency range of 0.15 to 0.4 Hz		✓		Respiratory activity
Low-frequency/High-frequency power	[70,73]	Ratio of low- to high-frequency spectral power	✓	✓		Sympathovagal balance
Total spectral power	[61]	Spectral power $\leq$ 0.4 Hz			✓	All cyclic components
<b>Non-Linear</b>						
SD1	[74–76]	Standard deviation—Poincaré plot (Perpendicular)		✓		Short term changes in heart rate variability
SD2	[75–77]	Standard deviation—Poincaré plot (Parallel)	✓	✓		Long term changes in heart rate variability
D <sub>2</sub>	[62,78]	Correlation dimension				Complexity of the system
CVI	[79]	$\log_{10}$ (longitudinal axis $\times$ transverse axis)—Lorenz plot		✓		
CSI	[79]	longitudinal axis/transverse axis—Lorenz plot	✓			

Note. The ticks represent the putative components of the autonomic nervous system each heart rate variability index is thought to reflect. CVI = Cardiac vagal index. CSI = Cardiac sympathetic index.

In particular, frequency-domain components have been extensively examined within the cognitive fatigue literature. This is likely due to the easier physiological interpretation of frequency-domain components and time constraints on experimental designs limiting the collection of long-term recordings. For instance, previous research has demonstrated lower high-frequency power after a 30-min 2-back task [48,49], 64-min Multi-attribute Task Battery [44], 90-min vigilance task [45], and 2-h set-shifting task [50] and simple arithmetic task [51]. A more recent study has also demonstrated the decrease in high-frequency power after an 8-hour fatigue-inducing task, which consisted of multiple sets of advanced trail making tests, kana pick-out tests, and mirror drawing tests [47]. Given that high-frequency power has been hypothesised to reflect parasympathetic nervous activity [65,70,71], it appears that decreased parasympathetic nervous activity is involved in cognitive fatigue.

However, previous research has also demonstrated higher low-frequency power after the aforementioned 2-back task [49], Multi-attribute Task Battery [44], vigilance task [45], shifting task [50], and simple arithmetic task [51]. A similar effect was also observed in a 4-hour driving task [43] and, albeit marginal, a 140-min visual tracking task [46]. Given that low-frequency power is sensitive to both sympathetic and parasympathetic nervous activity [65,69–72], these findings indicate that the sympathetic nervous activity is also involved in cognitive fatigue.

In contrast to low-frequency power, the ratio of low- to high-frequency power arguably serves as a more precise indicator of the balance and interaction between the sympathetic nervous system and parasympathetic nervous system [70,73]. An increase in the ratio of low- to high-frequency power after the vigilance task [45], set-shifting task [50], simple arithmetic task [51], and the 8-hour fatigue-inducing task [47] has also been found in previous studies. While inconsistent and marginal, a similar effect has also been observed in the 2-back task [48,49]. Furthermore, Tanaka and colleagues [49] also demonstrated that greater sympathetic nervous activity was associated with greater levels of self-reported fatigue, while lesser parasympathetic nervous activity was associated with greater levels of fatigue. In addition, the predominance of the sympathetic nervous activity in the overall autonomic nervous system was also positively associated with self-rated fatigue levels [47,49]. Overall, the pattern of results suggests a sympathovagal imbalance with a shift towards sympathetic predominance may be linked to cognitive fatigue in the general population. The sympathovagal imbalance, or rather balance, can be broadly conceptualised as the dynamic influence of the sympathetic and parasympathetic nervous activity on one's cardiac state [80]. Hence, it appears that there is empirical evidence satisfying the biomarker evaluation criteria of association [81,82], indicating that sympathovagal imbalance may be a putative physiological biomarker of cognitive fatigue.

#### 4. Digital Biomarkers of Cognitive Fatigue through Wearables and Machine Learning

With the advent of affordable mobile phones and wearables, the large amount of data collected from these devices can provide extensive information regarding the user [83–85], including working professionals [86]. Consequentially, using data-driven approaches, such as machine learning, to process these large datasets appear to be a promising avenue in predicting one's current psychological states [87]. Machine learning is a field within artificial intelligence broadly defined as an algorithmic approach that detects patterns through automation and optimisation, with minimal user input, to make predictions or decisions [88,89]. In practice, machine learning allows researchers to build computational models from large datasets that can learn, classify, predict, and improve through training [90–92].

When developing a machine learning model, a large dataset is typically divided into three subcategories—training, validation, and test datasets [93]. The training dataset is used for model fitting [93,94]. By contrast, the validation dataset is used to provide prediction error estimates of the fitted models during model selection [93]. In addition, the validation dataset is also used to make tuning adjustments to the parameters for further optimisation of the model [94]. Last, the test dataset is used only once after the training and validation phase to provide an unbiased assessment of the prediction error of the final model [93,94]. Ideally, a given dataset would be split into these three subsets for building, optimising, and evaluating a machine learning model.

Throughout the model building process, the models are evaluated with multiple performance metrics [95–100] (refer to Table 3). Accuracy is one of the key performance metrics of a robust machine learning model. In the context of a binary classification model, accuracy is calculated by the proportion of correct predictions divided by the total number of predictions. The correct predictions are the sum of true positive and true negative predictions (see Table 4). By contrast in non-classification models, such as regression models, accuracy can be calculated by mean absolute error, mean squared error, root mean squared error, or coefficient of determination ( $R^2$ ).

**Table 3.** Key Performance Metrics of Machine Learning Models.

Metrics	Formula	Description
<b>Classification Model [95,99,100]</b>		
Accuracy	$\frac{\text{True Positive} + \text{True Negative}}{\text{All Cases}}$	Overall ability of a model to make the correct classification
Precision	$\frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$	Ability of a classification model to make correct predictions within the positive class
Sensitivity (Recall)	$\frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$	Ability to correctly identify positive labels
Specificity	$\frac{\text{True Negative}}{\text{True Negative} + \text{False Positive}}$	Ability to correctly identify negative labels
F-score	$\frac{2 \times \text{Precision} \times \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}}$	Harmonic mean of sensitivity and precision
Area Under the Curve (AUC) of the Receiver Operating Characteristic Curve	$\frac{1}{2} (\text{Sensitivity} + \text{Specificity})$	Ability of a model to avoid misclassification
<b>Non-classification models [96–98]</b>		
MAE	$\frac{1}{n} \sum_{i=1}^n  \gamma_i - \hat{\gamma}_i $	Average of the absolute difference between observed values and predicted values
MSE	$\frac{1}{n} \sum_{i=1}^n (\gamma_i - \hat{\gamma}_i)^2$	Average of the squared difference between observed values and predicted values
RMSE	$\sqrt{\frac{1}{n} \sum_{i=1}^n (\gamma_i - \hat{\gamma}_i)^2}$	Standard deviation of the difference between observed values and predicted values
CoD/R <sup>2</sup>	$\frac{\sum_{i=1}^n (\hat{\gamma}_i - \gamma)^{-2}}{\sum_{i=1}^n (\gamma_i - \gamma)^2}$	Proportion of variance in the outcome variable explained by the predictor(s)

Note. MAE = Mean Absolute Error. MSE = Mean Squared Error. RMSE = Root Mean Squared Error.

**Table 4.** Confusion Matrix for Binary Classification Models.

		Actual Classification	
		Positive	Negative
Predicted Classification	Positive	True Positive	False Positive Type I Error
	Negative	False Negative Type II Error	True Negative

Note. The confusion matrix represents the four possible outcomes of a binary classification model.

Indeed, there is a growing interest in research to identify data-driven biomarkers [83,101]. More recently termed as digital biomarkers, these data-driven indices have unique advantages beyond traditional biomarkers, such as analysis at both the individual and population level, longitudinal and continuous measures, and passive monitoring [83]. More importantly, the emergence and increasing prevalence of wearables with the capability to measure physiological data allows for the further development of putative physiological-based digital biomarkers [101]. These wearables are capable of collecting physiological data, such as blood oxygen saturation, blood pressure, body temperature, electrodermal activity, and heart rate [102]. Not surprisingly, there have been preliminary successes in predicting both physical and mental health using wearable data both through traditional statistical modelling, and more recently, machine learning approaches [103–105].

In the context of cognitive fatigue, some researchers have recently attempted to predict fatigue levels by adopting the digital biomarker approach. For instance, a study by Al-Libawy and colleagues [106] using data collected from a wrist wearable compared two different machine learning methods (i.e., artificial neural network and support vector

machine) to predict cognitive fatigue. Six extracted physiological features were used (i.e., heart rate mean and standard deviation, wrist temperature mean and standard deviations, heart rate and wrist temperature power spectral density), which were chosen based on their influence on classification results. The artificial neural network and support vector machine models achieved 88.3% and 91.3% accuracy in classifying cognitive fatigue state, respectively, though the details of the test sample are unclear and might be inflated by resampling. Furthermore, the models achieved 94.7% and 97.2% accuracy in classifying alertness state, respectively. However, it should be noted that cognitive fatigue and alertness were not directly measured in this study but inferred from the ratio of low- to high-frequency power that was concurrently collected from an electrocardiograph.

Another study using reduced cognitive performance as an index of cognitive fatigue compared the support vector machine and random forest approaches with and without principal components analysis in predicting cognitive fatigue using various HRV features collected with a research-grade electrocardiograph [107]. The three-fold cross-validated random forest model achieved only 57.8% accuracy and, in combination with principal components analysis (leave-one-out cross-validated), improved to 63.9% accuracy. By contrast, they found that the three-fold cross-validated support vector machine model achieved 60% accuracy in predicting cognitive performance. Furthermore, the addition of principal components analysis increased accuracy to 84.4% with a precision of 92.6%, a recall of 73.3%, and an *f*-score of 81.8%. Notably, some of the selected features for the support vector machine model included not only the ratio of low- to high-frequency power, but also time-domain components, such as the number of R-R intervals, the average of all normalised R-R intervals (AVNN), and the standard deviation of all normalised R-R intervals (SDNN), and non-linear components, such as cardiac vagal index (CVI) and cardiac sympathetic index (CSI). These features were measured and averaged from the fifth to eighth trial during the onset of cognitive fatigue. Trial differences between baseline and onset of cognitive fatigue for these features were also included as additional features amongst others. Tsunoda and colleagues [107] highlighted that the use of principal components analysis increased prediction accuracy as the dimensional reduction technique reduced measurement noise. By analysing this model at the individual level, the researchers found that cognitive performance was more accurately predicted in participants with (1) greater number of R-R intervals, (2) larger trial difference in AVNN, (3) larger trial difference in CVI, and (4) larger, but negative, trial difference in the number of R-R intervals. This indicates that cognitive performance can be predicted with higher accuracy in participants with a certain type of physiological profile.

More recently, a study examined various machine learning approaches (i.e., support vector machine, K-nearest neighbour, naive Bayes, and logistic regression) in predicting cognitive fatigue, using data collected from a portable electrocardiogram patch [108]. Using a random forest approach, three time-domain features were selected based on their contribution to prediction accuracy—as indicated by the mean decrease accuracy and mean decrease Gini values: AVNN, the root mean square of the differences between each successive normalised R-R interval (RMSSD), and the proportion of normalised R-R intervals that are more than 50 ms from preceding interval (pNN50). In addition, three frequency-domain features were selected, namely very-low-frequency power, low-frequency power, and total spectral power. After comparing the various machine learning approaches with the combination of up to six of the aforementioned HRV features, this study demonstrated that the K-nearest neighbour model ( $k = 3$ ) with AVNN, low-frequency power, and total spectral power features achieved the highest five-fold cross-validated accuracy of 75.5% in predicting self-reported cognitive fatigue [108], as measured by the Chalder Fatigue Scale [17].

Overall, these studies indicate the feasibility of using machine learning in processing physiological data to monitor cognitive fatigue with moderate to high accuracy rates. It should be highlighted that the studies conducted by Huang and colleagues [108], as well as Tsunoda et al. [107], used electrocardiogram-derived HRV. The biomarker research



described in the previous section also predominantly used electrocardiogram to measure HRV [43–45,48–51]. The electrocardiogram uses electrodes to measure the electrical activity of the cardiac cycle [109]. However, most consumer wearables with the capability to measure cardiac activity use photoplethysmography due to its simplicity, comfort, and cost [110,111]. In contrast to electrocardiography, photoplethysmography uses specific wavelengths of light, such as infrared, to measure blood volumetric changes [110], which can be used to estimate blood circulation and associated HRV [112]. Arguably, the peak-to-peak interval observed in photoplethysmography can be interpreted as the equivalent of the R-R intervals of electrocardiography [112]. However, in terms of real-world application, this is only true under non-movement conditions as photoplethysmography recordings are extremely sensitive to motion artefacts, such as wrist movements [112–115]. For example, a recent study using a clinical-grade electrocardiogram as a benchmark examined several photoplethysmography-based consumer and research-grade wearables under different conditions [113]. In particular, this study reported that wearables had greater measurement error during physical activity than at rest [113]. Within a laboratory setting, Al-Libawy and colleagues [106] have provided indicative evidence of the validity of photoplethysmography-derived features in predicting electrocardiography-derived features. Furthermore, previous biomarker studies have also provided indicative evidence that photoplethysmography-derived frequency-domain features predicted cognitive fatigue [46,47]. Moving forward, future research should examine the reliability and validity of using photoplethysmography in predicting cognitive fatigue beyond laboratory settings. Exploring plausible algorithmic approaches to account for motion artefacts in photoplethysmography would also be imperative within this research area. Prospective researchers should also explore other machine learning methods to predict cognitive fatigue levels to improve accuracy.

## 5. Towards a Biomarker-Informed Machine Learning Model of Cognitive Fatigue

It appears that traditional biomarker research, as well as digital biomarker research, has contributed substantially to our understanding of the physiological features of cognitive fatigue and the degree to which these features can accurately predict varying states of cognitive fatigue. However, it is important to highlight that the traditional biomarker approach predominantly uses statistical modelling, which is viewed as a form of primary data analysis, whereas the digital biomarker approach typically uses data mining, which is considered as a form of secondary data analysis [116]. Due to the ad-hoc nature of data mining, most researchers are very cautious when it comes to the use of data-driven approaches, such as machine learning, as spurious relationships observed within a dataset can be easily misinterpreted [116–119]. Indeed, such data-driven models are usually atheoretical and thus have limited interpretability [120]. On the flip side, discovering novel relationships in unstructured datasets through data-driven approaches can also help further develop and refine current theoretical accounts [121]. Nonetheless, data-driven models have the potential of achieving high predictive power as their primary goal is to maximise “fit” within a given dataset [117].

In the context of cognitive fatigue, the physiological underpinnings have largely been ignored in data-driven models. As evident in the aforementioned digital biomarker studies in the previous section, machine learning approaches have the potential of producing highly accurate predictive models of cognitive fatigue [106–108]. However, the generalisability of these machine learning models should be further evaluated using test datasets. It is also evident in these previous studies that the HRV features selected are fairly inconsistent, which points to the data-driven nature of these models [106–108]. Given that the imbalance in sympathetic and parasympathetic nervous activity has been proposed as a physiological correlate of cognitive fatigue, the knowledge derived from traditional biomarker research is particularly informative and should be incorporated into machine learning models to aid development and validation. Specifically, such domain-specific knowledge can help with the selection of parameters, features, or models, which could result in models that are

more theoretically coherent, physiologically sound, generalisable, and interpretable [122]. In addition, using multiple biomarkers to build a multivariate model could potentially improve overall predictive power [123]. Thus, adopting a hybridised approach, combining domain-specific knowledge gained from traditional biomarker research with modern machine learning approaches, could potentially help researchers to build a more robust and generalisable model of cognitive fatigue. To this end, future researchers could incorporate our current knowledge on the putative biomarkers of cognitive fatigue (e.g., low-frequency power, high-frequency power, and the ratio of low- to high-frequency power) to aid in feature selection or act as a parameter constraint when developing machine learning models of cognitive fatigue. Open-source tools, such as PySiology, are readily available for researchers to extract these physiological features for machine learning [124].

## 6. Issues and Implications

When building a model of cognitive fatigue, potential confounds or closely-related concepts, such as stress and mental workload, should be considered. Stress can be broadly defined as “an emergent process that involves interactions between individual and environmental factors, historical and current events, allostatic states, and psychological and physiological reactivity” [125] (p. 1). Mental workload, in particular, can be viewed as a form of task-related or occupation-related stressor [126]. Mental workload can be objectively defined as the cost of internal resources (i.e., mental effort) to perform at a certain level or complete a task [127–129]. Cognitive overloading due to high levels of mental workload may lead to cognitive fatigue [6]. Not surprisingly, previous studies have demonstrated that various frequency-domain HRV indices, such as low-frequency power, high-frequency power, and the ratio of low- to high-frequency power, were associated with increased stress (for review, see [130]). In addition, previous research has shown that the mental with physical workload condition had significantly lower AVNN and lower vagal modulation, indexed by lower pNN50, as compared to the physical only workload condition [131]. Interestingly, Fairclough and colleagues [44] found a significant interaction effect, whereby low-frequency power is higher in the low mental workload condition as compared to the high mental workload condition during the initial period of a 64-min task. While low-frequency power increased in both conditions over time, the increase in low-frequency power was attenuated in the low mental workload condition at the end of the task, resulting in higher levels of low-frequency power in the high mental workload condition than the low mental workload condition [44]. Overall, there appears to be significant overlap and complex interactions amongst cognitive fatigue, mental workload, and stress.

One argument is that putative biomarkers of a particular psychological state should have a certain level of specificity [82,132]. That is, HRV should be more strongly correlated with cognitive fatigue than with stress or mental workload. However, this approach could be erroneous; first, for many individuals, cognitive fatigue, stress, and mental workload could be inherently associated, and thus disentangling these phenomena might not only result in misclassification but could actually be impossible; second, this approach might result in excessive reductionism and thus “paradigm-bound theories” [133]. Hence, prospective researchers should further examine the interaction amongst cognitive fatigue, stress, and mental workload in relation to HRV changes, aiming to represent both “robust reverse inference” (i.e., predicting behaviour from biological responses and predicting biological responses from behaviour). In the context of building models, stress should be accounted for when predicting cognitive fatigue due to the significant overlap in physiological findings between stress and cognitive fatigue. Given the interaction between mental workload and cognitive fatigue on low-frequency power [44], it appears that mental workload is a potential moderator that needs to be considered. Moreover, previous research comparing cognitive fatigue and mental workload levels found that HRV is a better index of cognitive fatigue, whereas heart rate is a better index of mental workload [134]. Hence, future studies could explore using heart rate, as well as heart rate variability, as a measure of mental

workload in the development of a more generalisable predictive model of cognitive fatigue, serving as a moderator to account for varying levels of mental workload across different tasks and situations.

## 7. Conclusions and Applications

Cognitive fatigue is a mental state characterised by the subjective feelings of tiredness, insufficient energy, difficulty with concentration, and impaired ability to think [1,2]. Traditionally, cognitive fatigue has been assessed through self-report and cognitive task performance. Later, biomarker approaches have been adopted to help understand the physiological underpinnings of cognitive fatigue. In particular, the imbalance of sympathetic and parasympathetic nervous activity has been proposed as a physiological correlate of cognitive fatigue. Indeed, as highlighted in the second section of this paper, various HRV indices have been demonstrated to vary as a function of cognitive fatigue levels, indicating that these HRV measures are putative biomarkers of cognitive fatigue. More recently, researchers have also demonstrated that machine learning approaches are capable of predicting cognitive fatigue using physiological data to a high level of accuracy [106–108]. Given the ubiquity of wearables that can measure cardiovascular activity, it appears that data collected from these devices have the potential of accurately predicting cognitive fatigue through machine learning approaches. However, the use of domain-specific knowledge from traditional biomarker research with novel machine learning approaches is imperative in building a robust and generalisable predictive model of cognitive fatigue.

A robust model of cognitive fatigue would allow for the development of a continuous fatigue monitoring system on wearables, which could be used to alert or remind an individual of the need to rest. For instance, given that cognitive fatigue increases the risk of accidents and errors [14,15], such a device may help mitigate not only minor errors but also potentially major accidents. By incorporating fatigue alleviating interventions with this monitoring system, timely and adequate rest could be objectively quantified and maximised. For instance, a previous study has shown that providing HRV biofeedback to participants attenuated cognitive fatigue during a mentally fatiguing task [135]. These effects have also been observed in the chronic fatigue syndrome population, albeit preliminary, where a pilot study has demonstrated that providing HRV biofeedback has improvements on specific cognitive components of fatigue [136]. Besides biofeedback, exposure to natural sounds has also been shown to have a positive effect on cognitive fatigue recovery [137]. Overall, this area of research is important in advancing our knowledge on not only cognitive fatigue monitoring but also cognitive fatigue recovery, which have significant implications in mitigating and minimising the risk of human errors in cognitively fatiguing situations.

**Author Contributions:** Conceptualization, K.F.A.L., W.-S.G. and G.C.; methodology, K.F.A.L., W.-S.G. and G.C.; investigation, K.F.A.L., W.-S.G. and G.C.; resources, W.-S.G. and G.C.; writing—original draft preparation, K.F.A.L. and G.C.; writing—review and editing, K.F.A.L., W.-S.G. and G.C.; visualization, K.F.A.L.; supervision, G.C. and W.-S.G.; project administration, W.-S.G. and G.C.; funding acquisition, W.-S.G. and G.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research/project is supported by the National Research Foundation, Singapore, and Ministry of National Development, Singapore under its Cities of Tomorrow R&D Programme (CoT Award: COT-V4-2020-1). Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the view of National Research Foundation, Singapore and Ministry of National Development, Singapore.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Boksem, M.A.S.; Tops, M. Mental fatigue: Costs and benefits. *Brain Res. Rev.* **2008**, *59*, 125–139. [[CrossRef](#)] [[PubMed](#)]
2. Kluger, B.M.; Krupp, L.B.; Enoka, R.M. Fatigue and fatigability in neurologic illnesses: Proposal for a unified taxonomy. *Neurology* **2013**, *80*, 409–416. [[CrossRef](#)] [[PubMed](#)]
3. Desmond, P.A.; Hancock, P.A. Active and passive fatigue states. In *Stress, Workload, and Fatigue*, 1st ed.; Hancock, P.A., Desmond, P.A., Eds.; CRC Press: Boca Raton, FL, USA, 2008; pp. 455–465.
4. Grandjean, E. Fatigue in industry. *Occup. Environ. Med.* **1979**, *36*, 175–186. [[CrossRef](#)] [[PubMed](#)]
5. Job, R.F.S.; Dalziel, J. Defining fatigue as a condition of the organism and distinguishing it from habituation, adaptation, and boredom. In *Stress, Workload, and Fatigue*, 1st ed.; Hancock, P.A., Desmond, P.A., Eds.; CRC Press: Boca Raton, FL, USA, 2008; pp. 466–476.
6. Subramanyam, M.; Muralidhara, P.; Muralidhara, P. Mental workload and cognitive fatigue: A study. *IUP J. Manag. Res.* **2013**, *12*, 29–39.
7. van der Linden, D.; Frese, M.; Meijman, T.F. Mental fatigue and the control of cognitive processes: Effects on perseveration and planning. *Acta Psychol.* **2003**, *113*, 45–65. [[CrossRef](#)]
8. Ilies, R.; Huth, M.; Ryan, A.M.; Dimotakis, N. Explaining the links between workload, distress, and work–family conflict among school employees: Physical, cognitive, and emotional fatigue. *J. Educ. Psychol.* **2015**, *107*, 1136–1149. [[CrossRef](#)]
9. Lim, J.; Dinges, D.F. Sleep deprivation and vigilant attention. *Ann. N. Y. Acad. Sci.* **2008**, *1129*, 305–322. [[CrossRef](#)]
10. Gergelyfi, M.; Jacob, B.; Olivier, E.; Zénon, A. Dissociation between mental fatigue and motivational state during prolonged mental activity. *Front. Behav. Neurosci.* **2015**, *9*, 176. [[CrossRef](#)]
11. Holtzer, R.; Shuman, M.; Mahoney, J.R.; Lipton, R.; Verghese, J. Cognitive fatigue defined in the context of attention networks. *Neuropsychol. Dev. Cogn. B Aging Neuropsychol. Cogn.* **2011**, *18*, 108–128. [[CrossRef](#)]
12. Simon, J.; Takács, E.; Orosz, G.; Berki, B.; Winkler, I. Short-term cognitive fatigue effect on auditory temporal order judgments. *Exp. Brain Res.* **2020**, *238*, 305–319. [[CrossRef](#)]
13. Tanaka, M.; Ishii, A.; Watanabe, Y. Effects of mental fatigue on brain activity and cognitive performance: A magnetoencephalography study. *Anat. Physiol.* **2015**, *5*, 1–5. [[CrossRef](#)]
14. Al-Mekhlafi, A.A.; Isha, A.S.N.; Naji, G.M.A. The relationship between fatigue and driving performance: A review and directions for future research. *J. Crit. Rev.* **2020**, *7*, 134–141. [[CrossRef](#)]
15. Goode, J.H. Are pilots at risk of accidents due to fatigue? *J. Saf. Res.* **2003**, *34*, 309–313. [[CrossRef](#)]
16. Johansson, B.; Starmark, A.; Berglund, P.; Rödhölm, M.; Rönnbäck, L. A self-assessment questionnaire for mental fatigue and related symptoms after neurological disorders and injuries. *Brain Inj.* **2010**, *24*, 2–12. [[CrossRef](#)] [[PubMed](#)]
17. Chalder, T.; Berelowitz, G.; Pawlikowska, T.; Watts, L.; Wessely, S.; Wright, D.; Wallace, E.P. Development of a fatigue scale. *J. Psychosom. Res.* **1993**, *37*, 147–153. [[CrossRef](#)]
18. Greenberg, S.; Aislinn, P.; Kirsten, D. Development and validation of the fatigue state questionnaire: Preliminary findings. *Open Psychol. J.* **2016**, *9*, 50–65. [[CrossRef](#)]
19. Haefel, G.J.; Howard, G.S. Self-report: Psychology’s four-letter word. *Am. J. Psychol.* **2010**, *123*, 181–188. [[CrossRef](#)]
20. Penner, I.K.; Raselli, C.; Stöcklin, M.; Opwis, K.; Kappos, L.; Calabrese, P. The fatigue scale for motor and cognitive functions (FSMC): Validation of a new instrument to assess multiple sclerosis-related fatigue. *Mult. Scler.* **2009**, *15*, 1509–1517. [[CrossRef](#)]
21. Schmidt, E.A.; Schrauf, M.; Simon, M.; Fritzsche, M.; Buchner, A.; Kincses, W.E. Drivers’ misjudgement of vigilance state during prolonged monotonous daytime driving. *Accid. Anal. Prev.* **2009**, *41*, 1087–1093. [[CrossRef](#)]
22. Dorrian, J.; Roach, G.D.; Fletcher, A.; Dawson, D. Simulated train driving: Fatigue, self-awareness and cognitive disengagement. *Appl. Ergon.* **2007**, *38*, 155–166. [[CrossRef](#)]
23. Kikuchi, Y.; Ishii, N. The relationship between self-awareness of fatigue symptoms and working conditions in female nurses. *Sangyo Eiseigaku Zasshi* **2015**, *57*, 230–240. [[CrossRef](#)]
24. Brown, I.D. Driver fatigue. *Hum. Factors* **1994**, *36*, 298–314. [[CrossRef](#)]
25. Brown, I.D. Prospects for technological countermeasures against driver fatigue. *Accid. Anal. Prev.* **1997**, *29*, 525–531. [[CrossRef](#)]
26. Anwer, S.; Li, H.; Antwi-Afari, M.F.; Umer, W.; Wong, A.Y.L. Evaluation of physiological metrics as real-time measurement of physical fatigue in construction workers: State-of-the-art review. *J. Constr. Eng. Manag.* **2021**, *147*, 03121001. [[CrossRef](#)]
27. Younis, E.M.G.; Kanjo, E.; Chamberlain, A. Designing and evaluating mobile self-reporting techniques: Crowdsourcing for citizen science. *Pers. Ubiquitous Comput.* **2019**, *23*, 329–338. [[CrossRef](#)]
28. Lorist, M.M.; Bezdan, E.; ten Caat, M.; Span, M.M.; Roerdink, J.B.T.M.; Maurits, N.M. The influence of mental fatigue and motivation on neural network dynamics; an EEG coherence study. *Brain Res.* **2009**, *1270*, 95–106. [[CrossRef](#)]
29. Schwid, S.R.; Tyler, C.M.; Scheid, E.A.; Weinstein, A.; Goodman, A.D.; McDermott, M.P. Cognitive fatigue during a test requiring sustained attention: A pilot study. *Mult. Scler.* **2003**, *9*, 503–508. [[CrossRef](#)] [[PubMed](#)]
30. Tanaka, M.; Ishii, A.; Watanabe, Y. Neural effects of mental fatigue caused by continuous attention load: A magnetoencephalography study. *Brain Res.* **2014**, *1561*, 60–66. [[CrossRef](#)] [[PubMed](#)]
31. Stroop, J.R. Studies of interference in serial verbal reactions. *J. Exp. Psychol.* **1935**, *18*, 643–662. [[CrossRef](#)]
32. Simon, J.R.; Wolf, J.D. Choice reaction time as a function of angular stimulus-response correspondence and age. *Ergonomics* **1963**, *6*, 99–105. [[CrossRef](#)]

33. Wang, C.; Ding, M.; Kluger, B.M. Change in intraindividual variability over time as a key metric for defining performance-based cognitive fatigability. *Brain Cogn.* **2014**, *85*, 251–258. [[CrossRef](#)]
34. Wylie, G.R.; Genova, H.M.; DeLuca, J.; Dobryakova, E. The relationship between outcome prediction and cognitive fatigue: A convergence of paradigms. *Cogn. Affect. Behav. Neurosci.* **2017**, *17*, 838–849. [[CrossRef](#)]
35. Liu, J.-P.; Zhang, C.; Zheng, C.-X. Estimation of the cortical functional connectivity by directed transfer function during mental fatigue. *Appl. Ergon.* **2010**, *42*, 114–121. [[CrossRef](#)] [[PubMed](#)]
36. O’Keefe, K.; Hodder, S.; Lloyd, A. A comparison of methods used for inducing mental fatigue in performance research: Individualised, dual-task and short duration cognitive tests are most effective. *Ergonomics* **2020**, *63*, 1–12. [[CrossRef](#)]
37. Zhou, F.; Alsaïd, A.; Blommer, M.; Curry, R.; Swaminathan, R.; Kochhar, D.; Talamonti, W.; Tijerina, L.; Lei, B. Driver fatigue transition prediction in highly automated driving using physiological features. *Expert Syst. Appl.* **2020**, *147*, 113204. [[CrossRef](#)]
38. Möckel, T.; Beste, C.; Wascher, E. The effects of time on task in response selection—An ERP study of mental fatigue. *Sci. Rep.* **2015**, *5*, 10113. [[CrossRef](#)] [[PubMed](#)]
39. Samuel, I.B.H.; Wang, C.; Burke, S.E.; Kluger, B.; Ding, M. Compensatory neural responses to cognitive fatigue in young and older adults. *Front. Neural Circuits* **2019**, *13*, 12. [[CrossRef](#)] [[PubMed](#)]
40. Wang, C.; Trongnetrpunya, A.; Samuel, I.B.H.; Ding, M.; Kluger, B.M. Compensatory neural activity in response to cognitive fatigue. *J. Neurosci.* **2016**, *36*, 3919–3924. [[CrossRef](#)]
41. Tran, Y.; Craig, A.; Craig, R.; Chai, R.; Nguyen, H. The influence of mental fatigue on brain activity: Evidence from a systematic review with meta-analyses. *Psychophysiology* **2020**, *57*, e13554. [[CrossRef](#)]
42. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. *Clin. Pharmacol. Ther.* **2001**, *69*, 89–95. [[CrossRef](#)] [[PubMed](#)]
43. Egelund, N. Spectral analysis of heart rate variability as an indicator of driver fatigue. *Ergonomics* **1982**, *25*, 663–672. [[CrossRef](#)] [[PubMed](#)]
44. Fairclough, S.H.; Venables, L.; Tattersall, A. The influence of task demand and learning on the psychophysiological response. *Int. J. Psychophysiol.* **2005**, *56*, 171–184. [[CrossRef](#)]
45. Li, Z.; Jiao, K.; Chen, M.; Yang, Y.; Wang, C.; Qi, S. Spectral analysis of heart rate variability as a quantitative indicator of driver mental fatigue. *SAE Tech. Pap.* **2002**, 2002-01-0090:1–2002-01-0090:5. [[CrossRef](#)]
46. Mascord, D.J.; Heath, R.A. Behavioral and physiological indices of fatigue in a visual tracking task. *J. Saf. Res.* **1992**, *23*, 19–25. [[CrossRef](#)]
47. Mizuno, K.; Tanaka, M.; Yamaguti, K.; Kajimoto, O.; Kuratsune, H.; Watanabe, Y. Mental fatigue caused by prolonged cognitive load associated with sympathetic hyperactivity. *Behav. Brain Funct.* **2011**, *7*, 17. [[CrossRef](#)] [[PubMed](#)]
48. Tanaka, M.; Mizuno, K.; Tajima, S.; Sasabe, T.; Watanabe, Y. Central nervous system fatigue alters autonomic nerve activity. *Life Sci.* **2009**, *84*, 235–239. [[CrossRef](#)]
49. Tanaka, M.; Mizuno, K.; Yamaguti, K.; Kuratsune, H.; Fujii, A.; Baba, H.; Matsuda, K.; Nishimae, A.; Takesaka, T.; Watanabe, Y. Autonomic nervous alterations associated with daily level of fatigue. *Behav. Brain Funct.* **2011**, *7*, 46. [[CrossRef](#)]
50. Zhang, C.; Zheng, C.-X.; Yu, X.-L. Automatic recognition of cognitive fatigue from physiological indices by using wavelet packet transform and kernel learning algorithms. *Expert Syst. Appl.* **2009**, *36*, 4664–4671. [[CrossRef](#)]
51. Zhang, C.; Yu, X. Estimating mental fatigue based on electroencephalogram and heart rate variability. *Pol. J. Med. Phys. Eng.* **2010**, *16*, 67–84. [[CrossRef](#)]
52. McCorry, L.K. Physiology of the autonomic nervous system. *Am. J. Pharm. Educ.* **2007**, *71*, 78. [[CrossRef](#)] [[PubMed](#)]
53. Porges, S.W. The polyvagal perspective. *Biol. Psychol.* **2007**, *74*, 116–143. [[CrossRef](#)]
54. Thayer, J.F.; Lane, R.D. A model of neurovisceral integration in emotion regulation and dysregulation. *J. Affect. Disord.* **2000**, *61*, 201–216. [[CrossRef](#)]
55. Saul, J.P. Beat-to-beat variations of heart rate reflect modulation of cardiac autonomic outflow. *Physiology* **1990**, *5*, 32–37. [[CrossRef](#)]
56. Peltola, M.A. Role of editing of R–R intervals in the analysis of heart rate variability. *Front. Physiol.* **2012**, *3*, 148. [[CrossRef](#)]
57. Benarroch, E.E. The central autonomic network: Functional organization, dysfunction, and perspective. *Mayo Clin. Proc.* **1993**, *68*, 988–1001. [[CrossRef](#)]
58. Benarroch, E.E. Central autonomic control. In *Primer on the Autonomic Nervous System*, 3rd ed.; Robertson, D., Biaggioni, I., Burnstock, G., Low, P.A., Paton, J.F.R., Eds.; Elsevier: Amsterdam, The Netherlands, 2012; pp. 9–12.
59. Hansen, A.L.; Johnsen, B.H.; Thayer, J.F. Vagal influence on working memory and attention. *Int. J. Psychophysiol.* **2003**, *48*, 263–274. [[CrossRef](#)]
60. Massaro, S.; Pecchia, L. Heart rate variability (HRV) analysis: A methodology for organizational neuroscience. *Organ. Res. Methods* **2019**, *22*, 354–393. [[CrossRef](#)]
61. Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology. Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *Circulation* **1996**, *93*, 1043–1065. [[CrossRef](#)]
62. Shaffer, F.; Ginsberg, J.P. An overview of heart rate variability metrics and norms. *Front. Public Health* **2017**, *5*, 258. [[CrossRef](#)] [[PubMed](#)]
63. Melillo, P.; Bracale, M.; Pecchia, L. Nonlinear heart rate variability features for real-life stress detection. Case study: Students under stress due to university examination. *Biomed. Eng. Online* **2011**, *10*, 96. [[CrossRef](#)]

64. Fiskum, C.; Andersen, T.G.; Bornas, X.; Aslaksen, P.M.; Flaten, M.A.; Jacobsen, K. Non-linear heart rate variability as a discriminator of internalizing psychopathology and negative affect in children with internalizing problems and healthy controls. *Front. Physiol.* **2018**, *9*, 561. [[CrossRef](#)]
65. Kleiger, R.E.; Stein, P.K.; Bigger, J.T. Heart rate variability: Measurement and clinical utility. *Ann. Noninvasive Electrocardiol.* **2005**, *10*, 88–101. [[CrossRef](#)]
66. Bonaduce, D.; Marciano, F.; Petretta, M.; Migaux, M.L.; Morgano, G.; Bianchi, V.; Salemme, L.; Valva, G.; Condorelli, M. Effects of converting enzyme inhibition on heart period variability in patients with acute myocardial infarction. *Circulation* **1994**, *90*, 108–113. [[CrossRef](#)] [[PubMed](#)]
67. Fleisher, L.A.; Frank, S.M.; Sessler, D.I.; Cheng, C.; Matsukawa, T.; Vannier, C.A. Thermoregulation and heart rate variability. *Clin. Sci.* **1996**, *90*, 97–103. [[CrossRef](#)]
68. Taylor, J.A.; Carr, D.L.; Myers, C.W.; Eckberg, D.L. Mechanisms underlying very-low-frequency RR-interval oscillations in humans. *Circulation* **1998**, *98*, 547–555. [[CrossRef](#)] [[PubMed](#)]
69. Bernardi, L.; Leuzzi, S.; Radaelli, A.; Passino, C.; Johnston, J.A.; Sleight, P. Low-frequency spontaneous fluctuations of R-R interval and blood pressure in conscious humans: A baroreceptor or central phenomenon? *Clin. Sci.* **1994**, *87*, 649–654. [[CrossRef](#)]
70. Pagani, M.; Lombardi, F.; Guzzetti, S.; Rimoldi, O.; Furlan, R.; Pizzinelli, P.; Sandrone, G.; Malfatto, G.; Dell’Orto, S.; Piccaluga, E. Power spectral analysis of heart rate and arterial pressure variabilities as a marker of sympatho-vagal interaction in man and conscious dog. *Circ. Res.* **1986**, *59*, 178–193. [[CrossRef](#)]
71. Pomeranz, B.; Macaulay, R.J.B.; Caudill, M.A.; Kutz, I.; Adam, D.; Gordon, D.; Kilborn, K.M.; Barger, A.C.; Shannon, D.C.; Cohen, R.J.; et al. Assessment of autonomic function in humans by heart rate spectral analysis. *Am. J. Physiol. Heart Circ. Physiol.* **1985**, *248*, H151–H153. [[CrossRef](#)] [[PubMed](#)]
72. Rahman, F.; Pechnik, S.; Gross, D.; Sewell, L.; Goldstein, D.S. Low frequency power of heart rate variability reflects baroreflex function, not cardiac sympathetic innervation. *Clin. Auton. Res.* **2011**, *21*, 133–141. [[CrossRef](#)] [[PubMed](#)]
73. Malliani, A.; Lombardi, F.; Pagani, M. Power spectrum analysis of heart rate variability: A tool to explore neural regulatory mechanisms. *Br. Heart J.* **1994**, *71*, 1–2. [[CrossRef](#)]
74. Mourot, L.; Bouhaddi, M.; Perrey, S.; Cappelle, S.; Henriët, M.-T.; Wolf, J.-P.; Rouillon, J.-D.; Regnard, J. Decrease in heart rate variability with overtraining: Assessment by the Poincaré plot analysis. *Clin. Physiol. Funct. Imaging* **2004**, *24*, 10–18. [[CrossRef](#)] [[PubMed](#)]
75. Mourot, L.; Bouhaddi, M.; Perrey, S.; Rouillon, J.-D.; Regnard, J. Quantitative poincaré plot analysis of heart rate variability: Effect of endurance training. *Eur. J. Appl. Physiol.* **2004**, *91*, 79–87. [[CrossRef](#)]
76. Tulppo, M.P.; Mäkikallio, T.H.; Takala, T.E.S.; Seppänen, T.; Huikuri, H.V. Quantitative beat-to-beat analysis of heart rate dynamics during exercise. *Am. J. Physiol. Heart Circ. Physiol.* **1996**, *271*, H244–H252. [[CrossRef](#)]
77. De Vito, G.; Galloway, S.D.R.; Nimmo, M.A.; Maas, P.; McMurray, J.J.V. Effects of central sympathetic inhibition on heart rate variability during steady-state exercise in healthy humans. *Clin. Physiol. Funct. Imaging* **2002**, *22*, 32–38. [[CrossRef](#)]
78. Bolea, J.; Laguna, P.; Remartínez, J.M.; Rovira, E.; Navarro, A.; Bailón, R. Methodological framework for estimating the correlation dimension in HRV signals. *Comput. Math. Methods Med.* **2014**, *2014*, 129248. [[CrossRef](#)]
79. Toichi, M.; Sugiura, T.; Murai, T.; Sengoku, A. A new method of assessing cardiac autonomic function and its comparison with spectral analysis and coefficient of variation of R–R interval. *J. Auton. Nerv. Syst.* **1997**, *62*, 79–84. [[CrossRef](#)]
80. Goldberger, J.J. Sympathovagal balance: How should we measure it? *Am. J. Physiol. Heart Circ. Physiol.* **1999**, *276*, H1273–H1280. [[CrossRef](#)] [[PubMed](#)]
81. Ritsner, M.S.; Gottesman, I.I. Where do we stand in the quest for neuropsychiatric biomarkers and endophenotypes and what next? In *The Handbook of Neuropsychiatric Biomarkers, Endophenotypes and Genes Volume I: Neuropsychological Endophenotypes and Biomarkers*; Ritsner, M.S., Ed.; Springer: Dordrecht, The Netherlands, 2009; pp. 3–21.
82. Woo, C.-W.; Chang, L.J.; Lindquist, M.A.; Wager, T.D. Building better biomarkers: Brain models in translational neuroimaging. *Nat. Neurosci.* **2017**, *20*, 365–377. [[CrossRef](#)]
83. Babrak, L.M.; Menetski, J.; Rebhan, M.; Nisato, G.; Zinggeler, M.; Brasier, N.; Baerenfaller, K.; Brenzikofer, T.; Baltzer, L.; Vogler, C.; et al. Traditional and digital biomarkers: Two worlds apart? *Digit. Biomark.* **2019**, *3*, 92–102. [[CrossRef](#)]
84. Califf, R.M. Biomarker definitions and their applications. *Exp. Biol. Med.* **2018**, *243*, 213–221. [[CrossRef](#)]
85. Seshadri, D.R.; Li, R.T.; Voos, J.E.; Rowbottom, J.R.; Alfes, C.M.; Zorman, C.A.; Drummond, C.K. Wearable sensors for monitoring the physiological and biochemical profile of the athlete. *NPJ Digit. Med.* **2019**, *2*, 72. [[CrossRef](#)]
86. Rykov, Y.; Thach, T.; Dunleavy, G.; Roberts, A.C.; Christopoulos, G.; Soh, C.; Car, J. Activity tracker-based metrics as digital markers of cardiometabolic health in working adults: Cross-sectional study. *JMIR mHealth uHealth* **2020**, *8*, e16409:1–e16409:17. [[CrossRef](#)]
87. Yarkoni, T.; Westfall, J. Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspect. Psychol. Sci.* **2017**, *12*, 1100–1122. [[CrossRef](#)] [[PubMed](#)]
88. Dwyer, D.B.; Falkai, P.; Koutsouleris, N. Machine learning approaches for clinical psychology and psychiatry. *Annu. Rev. Clin. Psychol.* **2018**, *14*, 91–118. [[CrossRef](#)] [[PubMed](#)]
89. Murphy, K.P. Machine learning: What and why? In *Machine Learning: A Probabilistic Perspective*; The MIT Press: Cambridge, MA, USA, 2012; pp. 1–2.

90. Bzdok, D.; Meyer-Lindenberg, A. Machine learning for precision psychiatry: Opportunities and challenges. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **2018**, *3*, 223–230. [[CrossRef](#)] [[PubMed](#)]
91. Choy, G.; Khalilzadeh, O.; Michalski, M.; Do, S.; Samir, A.E.; Pianykh, O.S.; Geis, J.R.; Pandharipande, P.V.; Brink, J.A.; Dreyer, K.J. Current applications and future impact of machine learning in radiology. *Radiology* **2018**, *288*, 318–328. [[CrossRef](#)]
92. Jordan, M.I.; Mitchell, T.M. Machine learning: Trends, perspectives, and prospects. *Science* **2015**, *349*, 255–260. [[CrossRef](#)]
93. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning*, 2nd ed.; Springer: New York, NY, USA, 2009.
94. Ripley, B.D. *Pattern Recognition and Neural Networks*; Cambridge University Press: Cambridge, UK, 2005.
95. Ballabio, D.; Grisoni, F.; Todeschini, R. Multivariate comparison of classification performance measures. *Chemometr. Intell. Lab. Syst.* **2018**, *174*, 33–44. [[CrossRef](#)]
96. Botchkarev, A. Evaluating performance of regression machine learning models using multiple error metrics in azure machine learning studio. *SSRN Electron. J.* **2018**, 1–16. [[CrossRef](#)]
97. Botchkarev, A. A new typology design of performance metrics to measure errors in machine learning regression algorithms. *Interdiscip. J. Inf. Knowl. Manag.* **2019**, *14*, 45–76. [[CrossRef](#)]
98. Chai, T.; Draxler, R.R. Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature. *Geosci. Model Dev.* **2014**, *7*, 1247–1250. [[CrossRef](#)]
99. Ferri, C.; Hernández-Orallo, J.; Modroiu, R. An experimental comparison of performance measures for classification. *Pattern Recognit. Lett.* **2009**, *30*, 27–38. [[CrossRef](#)]
100. Sokolova, M.; Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* **2009**, *45*, 427–437. [[CrossRef](#)]
101. Bent, B.; Wang, K.; Grzesiak, E.; Jiang, C.; Qi, Y.; Jiang, Y.; Cho, P.; Zingler, K.; Ogbeide, F.I.; Zhao, A.; et al. The digital biomarker discovery pipeline: An open-source software platform for the development of digital biomarkers using mHealth and wearables data. *J. Clin. Transl. Sci.* **2021**, *5*, e19:1–e19:8. [[CrossRef](#)]
102. Kamišalić, A.; Fister, I.; Turkanović, M.; Karakatič, S. Sensors and functionalities of non-invasive wrist-wearable devices: A review. *Sensors* **2018**, *18*, 1714. [[CrossRef](#)] [[PubMed](#)]
103. Dunn, J.; Runge, R.; Snyder, M. Wearables and the medical revolution. *Per. Med.* **2018**, *15*, 429–448. [[CrossRef](#)] [[PubMed](#)]
104. Reinertsen, E.; Clifford, G.D. A review of physiological and behavioral monitoring with digital sensors for neuropsychiatric illnesses. *Physiol. Meas.* **2018**, *39*, 05TR01. [[CrossRef](#)] [[PubMed](#)]
105. Tortelli, R.; Rodrigues, F.B.; Wild, E.J. The use of wearable/portable digital sensors in Huntington’s disease: A systematic review. *Parkinsonism Relat. Disord.* **2021**, *83*, 93–104. [[CrossRef](#)] [[PubMed](#)]
106. Al-Libawy, H.; Al-Ataby, A.; Al-Nuaimy, W.; Al-Tae, M.A. HRV-based operator fatigue analysis and classification using wearable sensors. In Proceedings of the 2016 13th International Multi-Conference on Systems, Signals & Devices (SSD), Leipzig, Germany, 21–24 March 2016; pp. 268–273.
107. Tsunoda, K.; Chiba, A.; Yoshida, K.; Watanabe, T.; Mizuno, O. Predicting changes in cognitive performance using heart rate variability. *IEICE Trans. Inf. Syst.* **2017**, *100*, 2411–2419. [[CrossRef](#)]
108. Huang, S.; Li, J.; Zhang, P.; Zhang, W. Detection of mental fatigue state with wearable ECG devices. *Int. J. Med. Inform.* **2018**, *119*, 39–46. [[CrossRef](#)]
109. AlGhatrif, M.; Lindsay, J. A brief review: History to understand fundamentals of electrocardiography. *J. Community Hosp. Intern. Med. Perspect.* **2012**, *2*, 14383. [[CrossRef](#)] [[PubMed](#)]
110. Castaneda, D.; Esparza, A.; Ghamari, M.; Soltanpur, C.; Nazeran, H. A review on wearable photoplethysmography sensors and their potential future applications in health care. *Int. J. Biosens. Bioelectron.* **2018**, *4*, 195–202. [[CrossRef](#)] [[PubMed](#)]
111. Sviridova, N.; Sakai, K. Human photoplethysmogram: New insight into chaotic characteristics. *Chaos Solitons Fractals* **2015**, *77*, 53–63. [[CrossRef](#)]
112. Lu, S.; Zhao, H.; Ju, K.; Shin, K.; Lee, M.; Shelley, K.; Chon, K.H. Can photoplethysmography variability serve as an alternative approach to obtain heart rate variability information? *J. Clin. Monit. Comput.* **2008**, *22*, 23–29. [[CrossRef](#)]
113. Bent, B.; Goldstein, B.A.; Kibbe, W.A.; Dunn, J.P. Investigating sources of inaccuracy in wearable optical heart rate sensors. *NPJ Digit. Med.* **2020**, *3*, 18. [[CrossRef](#)] [[PubMed](#)]
114. Schuurmans, A.A.T.; de Looft, P.; Nijhof, K.S.; Rosada, C.; Scholte, R.H.J.; Popma, A.; Otten, R. Validity of the Empatica E4 wristband to measure heart rate variability (HRV) parameters: A comparison to electrocardiography (ECG). *J. Med. Syst.* **2020**, *44*, 190. [[CrossRef](#)]
115. Yu, C.; Liu, Z.; McKenna, T.; Reisner, A.T.; Reifman, J. A method for automatic identification of reliable heart rates calculated from ECG and PPG waveforms. *J. Am. Med. Inform. Assoc.* **2006**, *13*, 309–320. [[CrossRef](#)]
116. Hand, D.J. Data mining: Statistics and more? *Am. Stat.* **1998**, *52*, 112–118. [[CrossRef](#)]
117. Hand, D.J. Statistics and data mining. *ACM SIGKDD Explor. Newsl.* **1999**, *1*, 16–19. [[CrossRef](#)]
118. L’Heureux, A.; Grolinger, K.; Elyamany, H.F.; Capretz, M.A.M. Machine learning with big data: Challenges and approaches. *IEEE Access* **2017**, *5*, 7776–7797. [[CrossRef](#)]
119. Smith, G. The paradox of big data. *SN Appl. Sci.* **2020**, *2*, 1041. [[CrossRef](#)]
120. Chin-Yee, B.; Upshur, R. Three problems with big data and artificial intelligence in medicine. *Perspect. Biol. Med.* **2019**, *62*, 237–256. [[CrossRef](#)]

121. Adjerid, I.; Kelley, K. Big data in psychology: A framework for research advancement. *Am. Psychol.* **2018**, *73*, 899–917. [[CrossRef](#)] [[PubMed](#)]
122. He, Q.P.; Wang, J. Application of systems engineering principles and techniques in biological big data analytics: A review. *Processes* **2020**, *8*, 951. [[CrossRef](#)]
123. Robotti, E.; Manfredi, M.; Marengo, E. Biomarkers discovery through multivariate statistical methods: A review of recently developed methods and applications in proteomics. *J. Proteom. Bioinform.* **2013**, *S3*, 1–20. [[CrossRef](#)]
124. Gabrieli, G.; Azhari, A.; Esposito, G. PySiology: A python package for physiological feature extraction. In *Neural Approaches to Dynamics of Signal Exchanges; Smart Innovation, Systems and Technologies*; Esposito, A., Faundez-Zanuy, M., Morabito, F.C., Pasero, E., Eds.; Springer: Singapore, 2020; Volume 151, pp. 395–402.
125. Epel, E.S.; Crosswell, A.D.; Mayer, S.E.; Prather, A.A.; Slavich, G.M.; Puterman, E.; Mendes, W.B. More than a feeling: A unified view of stress measurement for population science. *Front. Neuroendocrinol.* **2018**, *49*, 146–169. [[CrossRef](#)]
126. López-Núñez, M.I.; Rubio-Valdehita, S.; Diaz-Ramiro, E.M.; Aparicio-García, M.E. Psychological capital, workload, and burnout: What's new? The impact of personal accomplishment to promote sustainable working conditions. *Sustainability* **2020**, *12*, 8124. [[CrossRef](#)]
127. Hart, S.G. Nasa-task load index (NASA-TLX); 20 years later. In Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting, San Francisco, CA, USA, 16–20 October 2006; pp. 904–908.
128. Hart, S.G.; Staveland, L.E. Development of NASA-TLX (task load index): Results of empirical and theoretical research. In *Human Mental Workload; Advances in Psychology*; Hancock, P.A., Meshkati, N., Eds.; North Holland Publishing Company: Amsterdam, The Netherlands, 1988; Volume 52, pp. 139–183.
129. Xie, B.; Salvendy, G. Prediction of mental workload in single and multiple tasks environments. *Int. J. Cogn. Ergon.* **2000**, *4*, 213–242. [[CrossRef](#)]
130. Kim, H.-G.; Cheon, E.-J.; Bai, D.-S.; Lee, Y.H.; Koo, B.-H. Stress and heart rate variability: A meta-analysis and review of the literature. *Psychiatry Investig.* **2018**, *15*, 235–245. [[CrossRef](#)]
131. Taelman, J.; Vandeput, S.; Vlemincx, E.; Spaepen, A.; Van Huffel, S. Instantaneous changes in heart rate regulation due to mental load in simulated office work. *Eur. J. Appl. Physiol.* **2011**, *111*, 1497–1505. [[CrossRef](#)]
132. Lee, K.F.A.; Fox, A.M.; Notebaert, L. The effects of anxiety, depressive, and obsessive-compulsive subclinical symptoms on performance monitoring. *Int. J. Psychophysiol.* **2020**, *158*, 362–369. [[CrossRef](#)] [[PubMed](#)]
133. Varoquaux, G.; Poldrack, R.A. Predictive models avoid excessive reductionism in cognitive neuroimaging. *Curr. Opin. Neurobiol.* **2019**, *55*, 1–6. [[CrossRef](#)]
134. Trutschel, U.; Heinze, C.; Sirois, B.; Golz, M.; Sommer, D.; Edwards, D. Heart rate measures reflect the interaction of low mental workload and fatigue during driving simulation. In Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications AutomotiveUI '12, Portsmouth, NH, USA, 17–19 October 2012; pp. 261–264.
135. Brown, D.M.Y.; Bray, S.R. Heart rate biofeedback attenuates effects of mental fatigue on exercise performance. *Psychol. Sport Exerc.* **2019**, *41*, 70–79. [[CrossRef](#)]
136. Windthorst, P.; Mazurak, N.; Kuske, M.; Hipp, A.; Giel, K.E.; Enck, P.; Nieß, A.; Zipfel, S.; Teufel, M. Heart rate variability biofeedback therapy and graded exercise training in management of chronic fatigue syndrome: An exploratory pilot study. *J. Psychosom. Res.* **2017**, *93*, 6–13. [[CrossRef](#)] [[PubMed](#)]
137. Abbott, L.C.; Taff, D.; Newman, P.; Benfield, J.A.; Mowen, A.J. The influence of natural sounds on attention restoration. *J. Park Recreat. Admi.* **2016**, *34*, 5–15. [[CrossRef](#)]