

Open camera or QR reader and
scan code to access this article
and other resources online.



Determination of the Effects of Transcutaneous Auricular Vagus Nerve Stimulation on the Heart Rate Variability Using a Machine Learning Pipeline

Anna Tarasenko, MSci,¹ Stefano Guazzotti, PhD,^{1,2} Thomas Minot, MEng,¹
Mikheil Oganessian, MEng,¹ and Nikolai Vysokov, PhD¹

Abstract

Background: We are all aware of day-to-day healthy stress, but, when sustained for long periods, stress is believed to lead to serious physical and mental health issues.

Materials and Methods: In this study, we investigated the potential effects of transcutaneous auricular vagus nerve stimulation (taVNS) on stress processing as reflected in the electrocardiogram (ECG)-derived biomarkers of stress adaptability. Stress reflecting biomarkers included a range of heart rate variability metrics: standard deviation of N-N intervals (SDNN), root mean squared of successive differences in heartbeat intervals (RMSSD), low-frequency component, high-frequency component and their ratio (LF, HF, and LF/HF).

In addition, we created a machine learning model capable of distinguishing between the stimulated and nonstimulated conditions from the ECG-derive data from various subjects and states. The model consisted of a deep convolutional neural network, which was trained on R-R interval (RRI) data extracted from ECG and time traces of LF, HF, LF/HF, SDNN, and RMSSD.

Results: Only LF/HF ratio demonstrated a statistically significant change in response to stimulation. Although the LF/HF ratio is expected to increase during exposure to stress, we have observed that stimulation during exposure to stress counteracts this increase or even reduces the LF/HF ratio. This could be an indication that the vagus nerve stimulation decreases the sympathetic activation during stress inducement.

Our Machine Learning model achieved an accuracy of 70% with no significant variations across the three states (baseline, stress, and recovery). However, training an analogous neural network to identify the states (baseline, stress, and recovery) proved to be unsuccessful.

Conclusion: Overall, in this study, we showed further evidence of the beneficial effect of taVNS on stress processing. Importantly we have also demonstrated the promising potential of ECG metrics as a biomarker for the development of closed-loop stimulation systems.

Keywords: tVNS, taVNS, HRV, AI, ML, stress

¹BrainPatch Ltd., London, United Kingdom.

²School of Physics and CRANN Institute, Trinity College Dublin, Dublin, Ireland.

© Anna Tarasenko et al. 2022; Published by Mary Ann Liebert, Inc. This Open Access article is distributed under the terms of the Creative Commons License [CC-BY] (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been updated on July 29, 2022 after first online publication of July 26, 2022 to reflect Open Access, with copyright transferring to the author(s), and a Creative Commons License (CY-BY) added (<http://creativecommons.org/licenses/by/4.0>).

Introduction

WE ARE ALL AWARE of day-to-day healthy stress, but, when sustained for long periods, stress is believed to lead to serious physiological and mental health issues.^{1–3} At the time of writing this article, the global pandemic and the changes imposed by related restrictions are having a profound effect on the mental health and well-being of people around the world.⁴ This created an unprecedented demand to develop effective tools for measuring and controlling stress to prevent stress-related disorders and disabilities. This is a complex task requiring a multifaceted approach and in this article, we investigated the potential for noninvasive brain stimulation and heart rate-derived parameters for this purpose. However, to recapitulate stress in a well-controlled environment, we had to limit the definition of stress to one that is specific to our research as a “lasting unpleasant emotional sensation triggered by an external visual cue.”

When looking at heart rate-derived parameters, one of the key features of a healthy heart is the ability to quickly react to various stressors. This property is controlled by a network of complex mechanisms and can be jointly referred to as heart rate variability (HRV). Various measures were proposed to both quantify and analyze HRV.⁵ These include time-domain analysis techniques (e.g., standard deviation of N-N intervals [SDNN], root mean squared of successive differences in heartbeat intervals [RMSSD]), and frequency domain techniques of analyzing HRV band powers and their ratios (e.g., high frequency [HF], low frequency [LF], and LF/HF ratio).

HRV is obtained by calculating the time intervals between the heartbeats—RR intervals. It is believed that inter-beat intervals reflect the balance between the sympathetic and parasympathetic nervous systems. Therefore, the aforementioned measures are associated with the activation or deactivation in either sympathetic or parasympathetic nervous systems.^{5–7} Studies investigating the effects of stress on HRV have shown that it results in a decrease in the RR intervals, increase in LF/HF ratio, mixed results in changes to SDNN, consistent decrease in RMSSD.^{8–10} Therefore, potentially these can be used as biomarkers for stress.^{10–12}

However, there is a high degree of contradiction in research with regard to the direction of change of these parameters and their relative importance for stress discrimination.¹³ Therefore, for the purposes of this study, we did not limit ourselves to one particular metric and included all of them in the analysis pipeline. In the literature, there is a small number of studies that used these features to design machine learning (ML) models, which would identify stressful and resting states.^{14,15} Nevertheless, there is a lack of studies that look at the change in HRV factors before, during, and after stress inducement within the same sample. As a result, we have a poor understanding of the dynamics of HRV throughout stressful situations.

One of the key structures responsible for the activity of the parasympathetic nervous system is the vagus nerve. Because of its extensive network of innervated internal organs, it is able to affect vascular, digestive, and immune activities.^{16–18} The functional magnetic resonance imaging studies of vagus nerve stimulation (VNS) demonstrate a decrease in activation of the limbic system after the stimulation.^{19,20} This is important because the limbic system is often hyperactive in people with psychiatric disorders such as major depressive disorder and general anxiety disorder.²¹ It has been suggested that the vagus nerve is able to affect hypothalamic-pituitary-adrenal (HPA) axis, which is responsible for stress response and is affected in the cases of emotional dysregulation.²²

At present, invasive VNS has been Food and Drug Administration (FDA) approved as a treatment for refractory major depression and intractable epilepsy.²³ Moreover, the noninvasive analogue was widely investigated as to whether it can bring similar benefits without surgical risks. There are data to suggest positive effects in depression, anxiety, irritable bowel syndrome, rheumatoid arthritis, tinnitus, and Crohn’s disease.^{16,18,22–28} Nevertheless, an important caveat is that although it might seem that there is a positive relationship between the degree of vagal activation and well-being, this notion has been challenged. There is evidence to suggest a nonlinear relationship between the vagal tone and well-being.²⁹ This further demonstrates the need for a personalized system of stimulation rather than generic protocols. Our study is one step toward the development of such a system. Therefore, the aims of our study were as follows:

- To investigate potential effects of transcutaneous auricular vagus nerve stimulation (taVNS) on stress processing as reflected in the biomarkers of stress adaptability calculated from electrocardiogram (ECG).
- To create an ML model capable of distinguishing between the stimulated and nonstimulated conditions from the ECG data, and generalized to various subjects and states.

Materials and Methods

Details of materials and methods, including information on ML algorithms are in the Supplementary Materials. In brief, within-subject design was used; ECG was recorded from healthy volunteers during a stressful video paradigm with and without taVNS (Fig. 1). Experimental design broadly adhered to the recommendations put forward in the literature.³⁰ All the participants were debriefed about the experimental aims and procedures; informed consent was obtained before the beginning of the experiment. The design of the experiment and all associated procedures adhered to the WMA Declaration of Helsinki. The ECG data were extracted and analyzed using conventional methods as well as using deep learning. The results are presented in the following section.

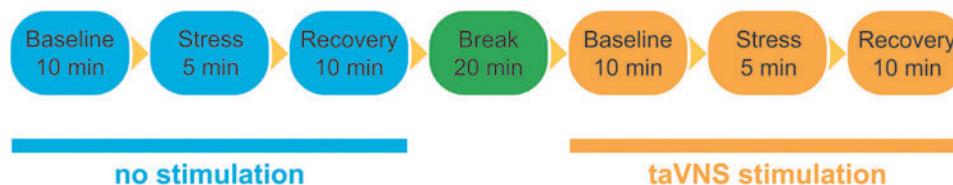


FIG. 1. Experimental design. There were two sessions: no stimulation and stimulation. Each lasted 25 min and contained 10 min of baseline, 5 min of stress inducement, and 10 min of recovery. There was also a 20 min break in between the sessions.

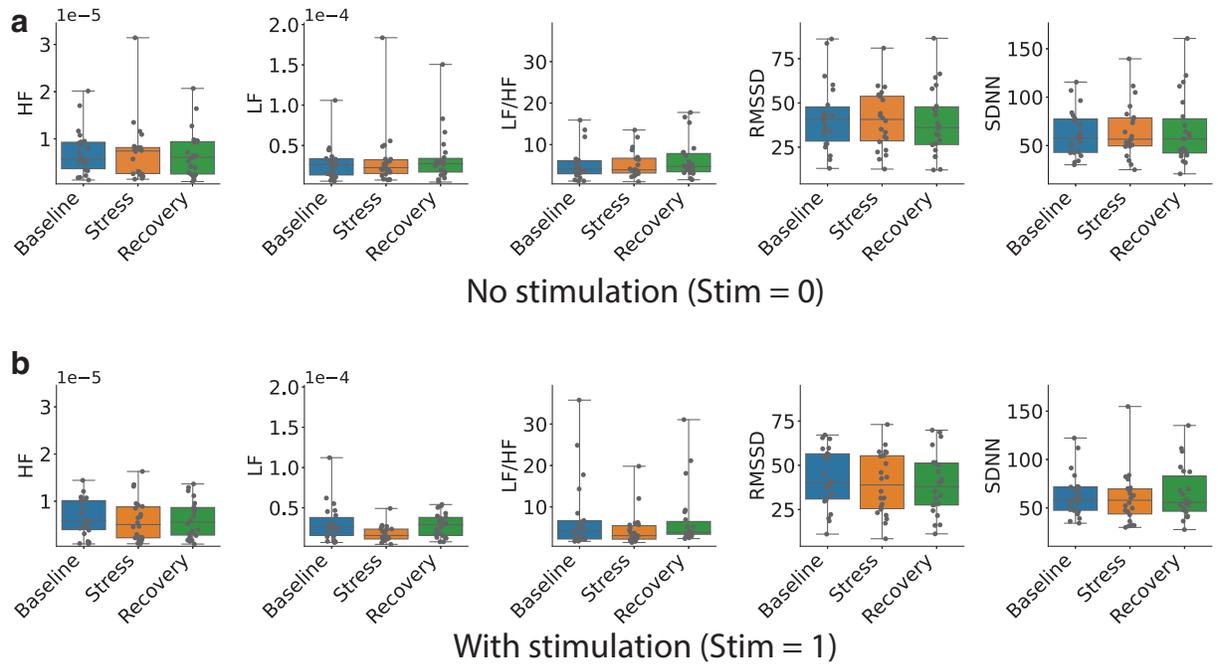


FIG. 2. Distribution of metric values derived from ECG as described in section S1.6 in Supplementary Material across subjects for each *state* for both conditions: with (b) and without (a) stimulation. The values do not have units for the frequency domain metrics (LF, HF, and LF/HF), and they represent spectral power for the respective bands. For the time domain metrics (RMSSD and SDNN), the values are in milliseconds. The box and whiskers refer to distribution parameters as described in section S1.7 in Supplementary Material. ECG, electrocardiogram; HF, high frequency; LF, low frequency; RMSSD, root mean squared of successive differences in heartbeat intervals; SDNN, standard deviation of N-N intervals.

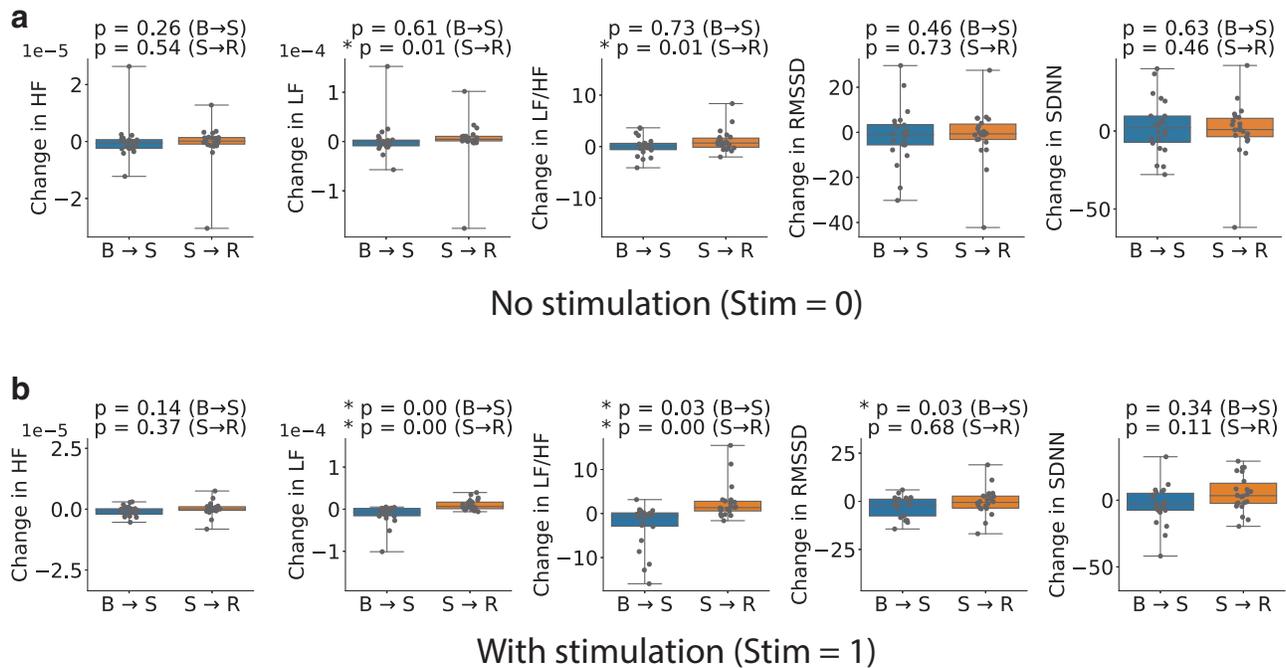


FIG. 3. Distribution of the *changes* in values of ECG-derived metrics between *neighboring states*, that is, baseline to stress ($B \rightarrow S$) and stress to recovery ($S \rightarrow R$) for both conditions: with (b) and without (a) stimulation. The values do not have units for the frequency domain metrics (LF, HF, and LF/HF), and they represent spectral power for the respective bands. For the time domain metrics (RMSSD and SDNN), the values are in milliseconds. We quantify the statistical significance of the change by testing against the null hypothesis that the population mean of the changes in metric is equal to 0 through the Wilcoxon signed-rank test. The resulting p -values are listed in the title of each panel.

Results

Statistical analysis of the effect of stress on HRV

Before establishing the effects of the stimulation on stress management, it was important to test whether the experimental design could induce stress in the participants and which of the measured ECG-derived parameters were affected by the stress inducement. To accomplish this, we selected the data from the nonstimulated section of the experiments, $Stim=0$, and tested whether stress inducement affected each parameter individually. The distribution of values across the different metrics and stress states are shown in Figure 2a.

The presence of a difference between the three subsequent states was quantified by using a Wilcoxon signed-rank test. It looked at the difference in value between the selected metric for each of the two combinations of states for each subject and tested whether the two samples come from populations with the same distribution. The p -values obtained from the tests are listed in the title of each panel of Figure 3a for comparing the distributions of baseline to stress ($B \rightarrow S$) and stress to recovery ($S \rightarrow R$) metrics. The panels themselves show the distributions of these differences and highlight the position of their quartiles. The same process was then repeated for the data collected during stimulation, leading to Figures 2b and 3b.

There are trends that are commonly reported in the literature after stress inducement, which we were expecting to see in our data. These include increase in LF/HF ratio after stress inducement and decrease in RMSSD.⁸⁻¹²

As can be seen from Figure 3a and b, the observed trends were somewhat different. Figure 2a and b shows that there was a lot of interindividual variability in all the metrics. In the nonstimulated condition (Fig. 3a), there was no statistically significant difference between the baseline and stress in any metric. There was a statistically significant change in LF and LF/HF in the transition between stress and recovery. Both LF and LF/HF increased going from stress to recovery. This is the opposite of what we expected. It might be caused by a delay in stress processing, meaning that people did not get stressed immediately after the beginning of the video but a bit later. In the stimulation condition, the situation was similar. There was a statistically significant increase in the LF/HF ratio between the stress and recovery states. But also, there was a statistically significant decrease in the LF/HF ratio and a decrease in RMSSD between the baseline and stress states.

Statistical analysis of the effect of $tAVNS$ on HRV

Overall, 15 people out of 22 reported feeling more relaxed during stress inducement if the stimulation was applied compared with the absence of stimulation. The rest reported not feeling a difference. To quantify the effect of applying $tAVNS$ stimulation on a subject, while at the same time taking into account interindividual variability, we focused on the differences between the two subsequent states (baseline-stress and stress-recovery) with and without stimulation as derived in Figure 3a and b. For ease of comparison, we have plotted the

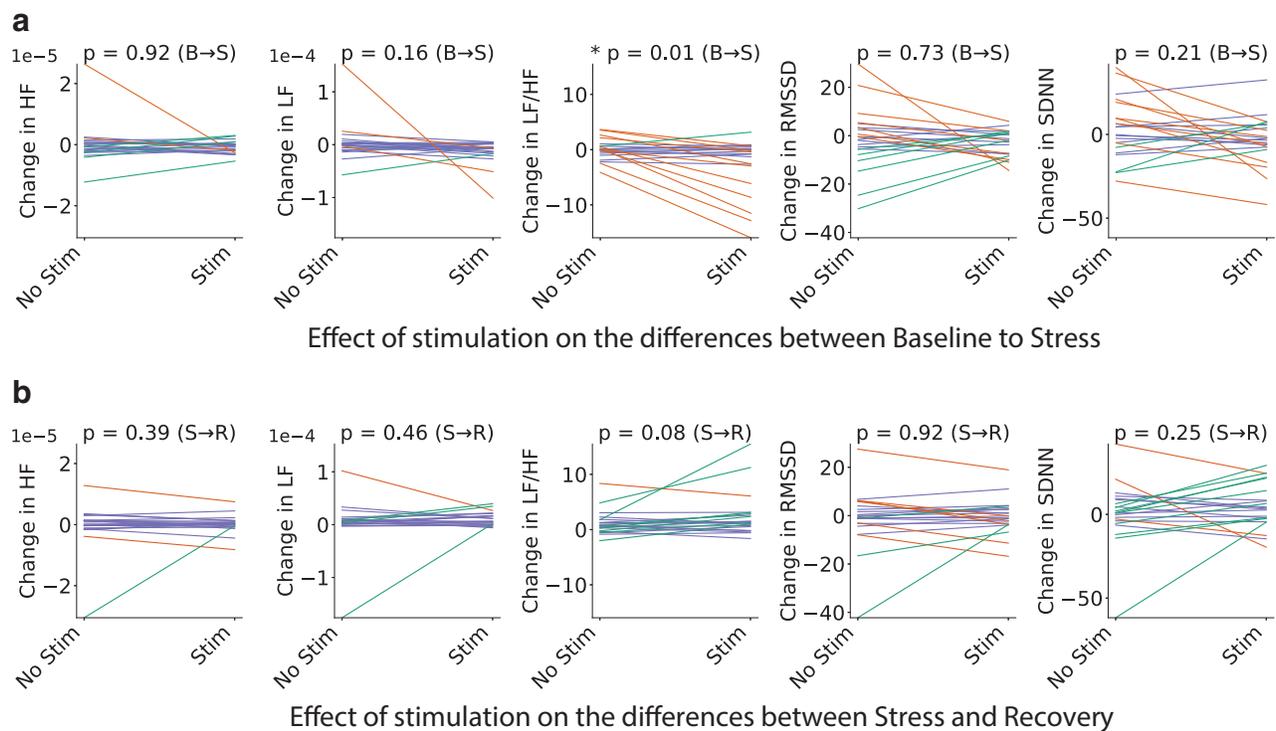


FIG. 4. Effect of stimulation on change in HRV metrics between neighboring states, i.e. baseline to stress (a) and stress to recovery (b). Values on the left show if a metric increased (> 0) or decreased (< 0) in going from one state to the next when no stimulation is applied. Values on the right correspond to the change in the value of the same metric for the stimulated condition. Lines joining the points are colored depending on whether they represent an increase (green), a decrease (orange), or a stationary value (purple). The threshold between increase, stationary, and decrease is set at $0.5 \times \sigma'$, where σ' is the standard deviation obtained from all the changes between the two relevant states with and without stimulation. p -Values in the title come from the Wilcoxon signed-rank test between the values on the left ($Stim=0$) and on the right ($Stim=1$) of each plot. HRV, heart rate variability.

two conditions on the same graph. Figure 4a shows the stimulation's effect on changes between baseline and stress states ($B \rightarrow S$). Similarly, Figure 4b shows the effect of stimulation on changes that occur between stress and recovery ($S \rightarrow R$).

Each point on the left corresponds to a change in the relevant metric for a given subject when no stimulation was applied. Corresponding change with stimulation applied is marked on the right and joined by a line to reflect on the individual effects of the stimulation. For visualization purposes, the color of each line is determined by the difference in metric changes between left ($\text{Stim}=0$) and right ($\text{Stim}=1$) of the plot where orange marks a clear decrease, green a clear increase, and purple a negligible effect, that is, a change smaller than half the standard deviation of all the metric changes between the two relevant stress phases with and without

stimulation. As before, the statistical significance of the change was evaluated through the Wilcoxon signed-rank test of the points on the left and right of each panel, leading to the p -values indicated in the title.

In the literature, there is a proposed concept of responders and nonresponders to stimulation.²⁶ Moreover, it seems that the presence or absence of the response might be related to the baseline LF/HF value. In the LF/HF section of Figure 4a there are two similar-sized groups of subjects whose LF/HF decreases or stays stationary. The former might be considered responders to stimulation given the presence of an effect on the HRV metric. Similarly, the latter could be identified as nonresponders, given the absence of the response. The Pearson correlation coefficient between stimulation effect and baseline (nonstimulated) value was evaluated to test

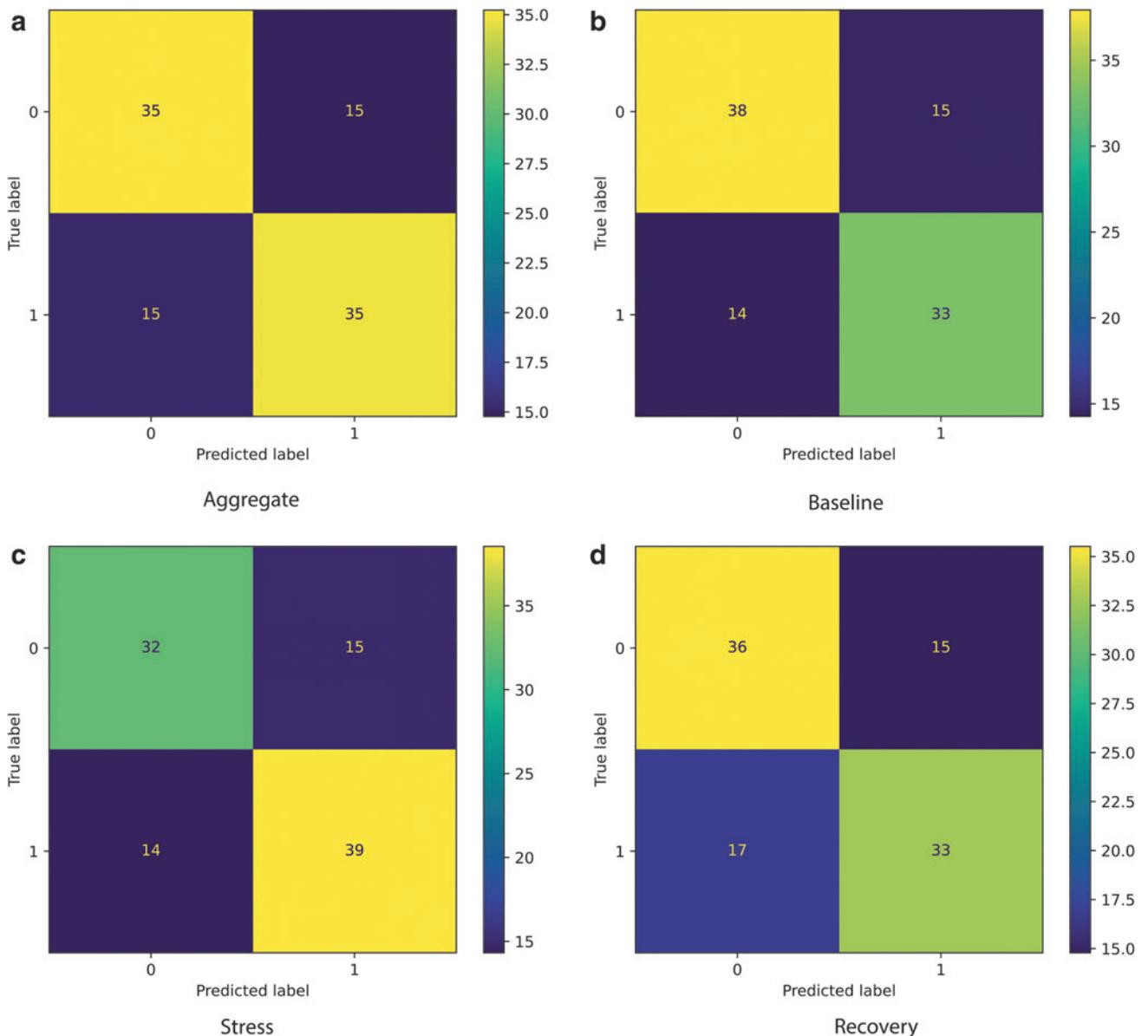


FIG. 5. Confusion matrices for discerning stimulated ($\text{Stim}=1$) and nonstimulated ($\text{Stim}=0$) conditions. The numbers inside the panels represent percentages of the total number of samples (1760 from 22 individuals) in the test set. Panel (a) is an aggregate across all three states, (b) baseline, (c) stress, and (d) recovery. Here the matrices (b, c, and d) are a decomposition of matrix (a).

whether the baseline value can predict the association with a particular category. The results show that there is a weak but statistically significant correlation ($p < 0.05$) in the LF/HF metric between the stimulation effect and the baseline value for both the baseline to stress ($r = -0.64, p = 0.001$) and stress to recovery transitions ($r = 0.51, p = 0.015$).

The negative correlation between the baseline value and the stimulation effect in the stress condition means the higher values of LF/HF in the baseline predict the bigger decrease in the LF/HF ratio in the stress state. Thus, implying that people with the higher baseline LF/HF ratio are more susceptible to the effects of stimulation and can be classified as responders. This is supported by the literature.^{26,31} In turn, the positive correlation of $r = 0.51$ between the baseline LF/HF ratio and the effect of stimulation between the stress and recovery means that the higher the baseline, the higher is the stimulation-induced increase in LF/HF from stress to recovery.

ML: discerning stimulated and nonstimulated conditions

By employing a convolutional neural network trained on the HRV data to discern stimulated versus nonstimulated conditions, we obtained a performance of 70% for both conditions and an F1 score of 69%. As was described in the section S1.8 in Supplementary Material, the two categories of data were fed into the model: temporal RRI series and their transformations in the frequency space (LF, HF, their ratio, RMSSD, and SDNN). Feeding both to the model yields better performance than training separately (data not shown). Figure 5 shows the confusion matrices first totaling all samples and then breaking them into confusion matrices for different states (baseline, stress, and recovery). The proportion of true labels coinciding with the predicted labels is 0.38 and 0.33 for the baseline state, 0.32 and 0.39 for the stress state, and 0.36 and 0.33 for the recovery state.

Interestingly, although misclassification is similar in baseline and stress states, it is somewhat different in the recovery state. The proportion of cases classified as stimulated while coming from the nonstimulated condition is 0.15. In contrast, the proportion of cases classified as nonstimulated while coming from the stimulated condition is 0.17. This slight bias toward the nonstimulated condition shown by the ML classifier might indicate that features correlated with stimulation presence are different in the recovery state compared to baseline and stress.

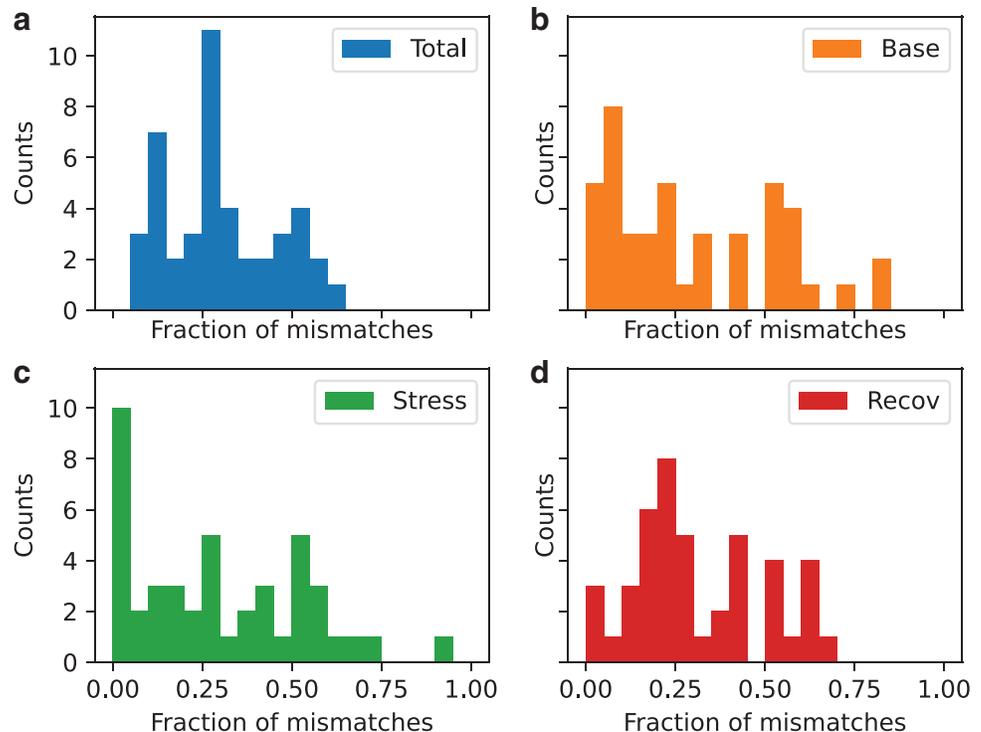
To identify whether the model performed equally well across all individuals in determining whether they were stimulated or not, we identified the number of mismatches for each individual for each condition. To visualize the results and look into the reasons for the misclassifications, we expressed them as a fraction of mismatches and plotted the distribution in Figure 6.

We can see that there is no individual with complete accuracy and only one individual with $>60\%$ mismatches. There are only around seven individuals with $>50\%$ mismatches in total. Overall, in all three states, the distribution of mismatches seems to be broadly distributed with the different degrees of the shift to the left from 50% (Fig. 6a). In particular, in the baseline state, the stimulation is more accurately determined as the distribution is shifted mainly to the left of 50%, with the tallest peak at $\sim 10\%$ (Fig. 6b). Similarly, the stress condition is still left-shifted, with a tall peak at $\sim 5\%$ mismatches and an even broader overall distribution (Fig. 6c). Recovery, however, is a bit worse, with the center of the distribution lying closer to 50% (Fig. 6d).

ML: discerning baseline, stress, and recovery states

On the task of classifying the session state, that is, baseline, stress, and recovery, two experiments were performed: baseline/stress binary classification and three-way baseline/stress/recovery classification. We built the train and test set by including the samples corresponding to the states of baseline and

FIG. 6. (a) Distribution of mismatches between the true and predicted labels in stimulation condition classification across individual subjects, including the source of these mismatches across the three different states. The percentage of mismatches was calculated for each session (2 conditions and 22 subjects) across the 40 cross-validation iterations and placed into histogram bins 5% wide. (b–d) A similar histogram where the data were restricted to those laying in the baseline, stress, and recovery state, respectively.



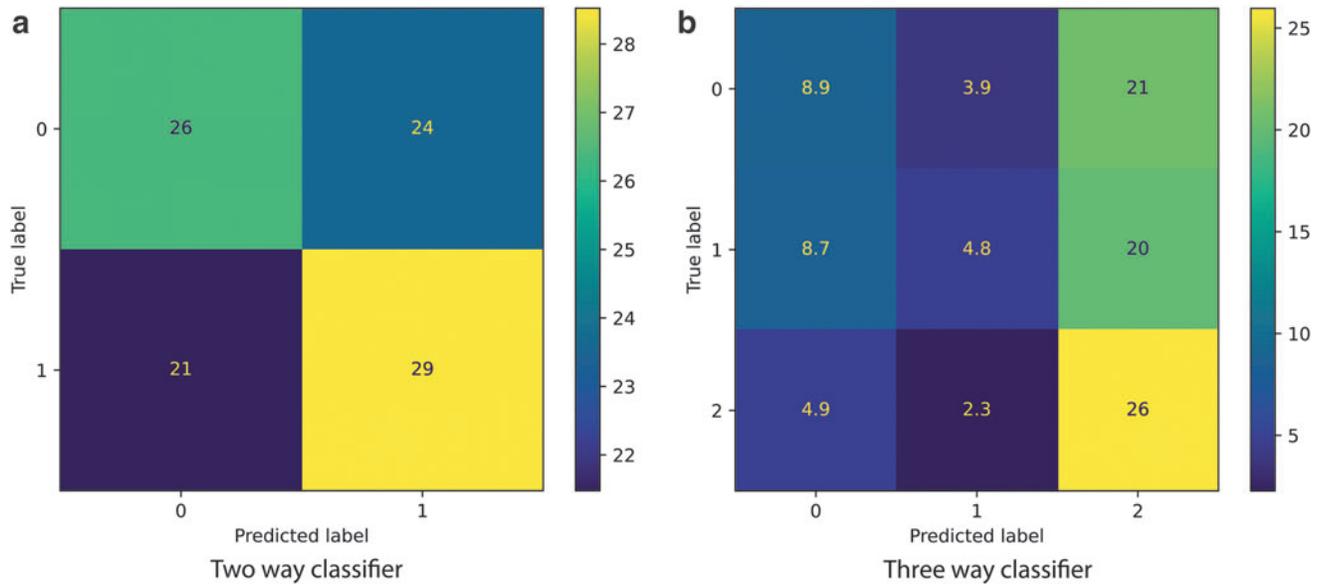


FIG. 7. (a) Confusion matrix for discerning baseline and stress states in a binary classification mode across the stimulated and nonstimulated conditions. In this figure the labels are Baseline = 0, Stress = 1, Recovery = 2. The numbers inside the panels represent percentages of the total number of samples (1760 across 22 individuals) in the test set. (b) Confusion matrix for discerning baseline, recovery, and stress states in a three-way classification mode across the stimulated and nonstimulated conditions.

stress for the binary classification task and the recovery state for the three-way task. This way, each state is represented equally in the test set, with each test set containing 1760 samples.

Using this model, we achieved 55% accuracy and f1 scores for the two-way classification (Fig. 7a) and 40% accuracy, 35% f1 scores for the three-way classification (Fig. 7b).

Under the assumption that data generated by the two-way classifier can be described by the Poisson distribution, the number of correctly identified examples would be $N_2 \approx 968 \pm 31$, where 31 is the statistical uncertainty. Meanwhile, the random chance performance would add up to 880 (half of the total number of examples). This means that the obtained performance is slightly better than chance. Similarly, for the three-way classifier, the number of correctly identified examples is $N_3 \approx 704 \pm 27$ as compared with 587 obtained by chance. Hence its performance is also slightly better than chance. Interestingly, the classifier seemed to overfit the recovery state and incorrectly predict it for both stress and baseline state.

Discussion

First, in this study, we have provided further evidence for the beneficial effects of taVNS reflected in HRV metrics. In Figure 4a we show that the stimulation is effective at reducing the expected LF/HF increase in response to stress and even reducing it below the baseline in several cases. This is important because we know from the literature that there is a disproportional increase in LF/HF in response to stress in people with affective disorders, such as anxiety.^{32,33}

Our results are in line with the findings on the effects of the transcutaneous VNS (tVNS) on the cervical branch of the vagus nerve. It was shown that tVNS decreases sympathetic activation in response to stress in healthy individuals^{31,34} as well as the people suffering from post-traumatic stress disorder (PTSD).³⁵ All our participants tolerated the stimulation well, with only one reporting a light muscle tension in the neck region during the

stimulation, which disappeared as soon as the stimulation stopped. This goes in line with mounting evidence of the safety of tVNS in the course of prolonged stimulation periods.³⁶

Second, we have created and trained an algorithm that can discern between stimulation and no stimulation conditions in various states (baseline, stress, and recovery) with a high degree of reliability (70%). Because the research in this area still remains limited, in particular in different affective states, it is difficult to conclude the meaning of this accuracy value as there is no benchmark. Based on the spread of the distribution of mismatches shown in Figure 6, it might be that the features associated with stimulation are the most prominent in the baseline condition (Fig. 6b), with stress condition following closely. In the recovery state, features seem to be the least pronounced. The fact that the distributions of mismatches vary across the different states might be indicative of the fact that the stimulation effects also vary in different states. This is under the assumption that the features learned by the model are indicative of the beneficial effects of stimulation.

The state dependence of the response to the taVNS we noted in Figure 4a and in the correlation analysis has been noted before in the rat studies using invasive VNS. VNS can reduce the overarousal present in rats that are stressed at the baseline, however, has no effect on the arousal in the non-stressed rats.³⁷ Vagus nerve receives numerous afferent projections that indicate the state of the viscera. If the subject is already in a calm state, with a high level of parasympathetic activity, the stimulation may not be as effective. This fits well with the polyvagal theory that focuses on the adaptive properties of the physiological state.³⁸ According to this theory, the physiological state defines the range of reactions that are possible to a given stimulus. There is also an interesting hypothesis that the release of adrenalin in response to stressful situations affects the brain through the vagus nerve.³⁹

The released adrenalin might bind to the beta-adrenergic receptors on the vagus nerve, which then increases the release

of norepinephrine in the brain as a homeostatic mechanism to counteract stress. This was proposed as a mechanism of effectiveness of tVNS in the fear extinction in PTSD. This could potentially explain why the differences between baseline and stress states were enhanced after the stimulation. Nevertheless, this hypothesis has only been tested in animal models with invasive VNS. In addition, it is important to mention a lack of studies on continuous changes in HRV after stress inducement. These are essential to understand the effects of stimulation on the stress response, especially given the contradictory findings about the importance of HRV markers in stress identification.

The ability to discern the presence of the stimulation from the HRV data using a small section (1 min) might imply that stimulation effects can be used as optimization features for the closed-loop system. Moreover, this might also mean that this model can be further adapted for online data processing to adjust the stimulation parameters immediately. The degree of individual calibration of the model for every new user requires further investigation. We envision the need for the collection of preliminary ECG data in a range of different stress scenarios to generate a versatile model. However, going beyond 1 min for the online processing will impose a further constraint on window length for the fast Fourier transform, as that determines how far the metrics lag behind real time.

In this study, we have used simple previously reported²⁶ stimulation parameters. However, optimal and desirable stimulation may require more complex biphasic patterns that would be the scope of our further investigations.

Another important part of the system that needs to be developed in parallel is the state recognition system. However, we found that both state classifiers (two-way and a three-way) are significantly worse than the stimulation condition classification model. From Figure 7a it seems the data set contains some features that allow distinguishing between the baseline and stress (and recovery) states; however, these are not prominent enough for the ML algorithm to reach a good accuracy. This leads us to believe that higher performance could be achieved by improving the quality and diversity of the collected data.

Alternatively, it might be due to the experimental design flaws, which meant that the induced stress was not sufficient to be discerned from the HRV metrics. Another possibility is that the participants were already stressed at baseline as they were expecting to see the stressful video. Using a nonstressful video as the baseline instead of the black screen might allow a clearer distinction. This, however, requires a more systematic study. Alternative stress-inducing protocols may give the model more information about the person's reaction to stress, thus enabling better recognition.

In principle, a state recognition system should conduct continuous labeling of states, rather than simple baseline-stress-recovery. To achieve this, one would need to reframe the model into a regression rather than classification. Therefore, the outcome would be a *probability* of belonging to different states. This would put the effects of the stimulation into the context of the state of a person and, therefore, enable one to bring a user closer to the desired state through the use of stimulation. However, the usefulness of HRV metrics for such purposes remains an open question given the poor performance of state classifier.

Thus, interpretation of our results has to be taken with a range of considerations that directly impact the generalizability and validity of conclusions that could be drawn.

Experimental design considerations

There are a wide variety of stress inducement protocols used in research. However, their common drawback is the lack of realistic scenarios. Common stress paradigms include cold water immersion, needle pricking, and electric shock.^{40,41} The advantage of these protocols is that they are more generic and less dependent on individual factors such as personality and background. Nevertheless, apart from the obvious ethical implications, these paradigms are difficult to administer remotely. Therefore, in the context of this study, alternatives were used.

Although the stressful video paradigm used was easy to administer remotely, it is similarly not representative of the real-life stressors most people face. Further development would be to explore other modalities and sensory inputs for inducing specific mental states, especially combining them to produce a more immersive environment (videos, auditory and haptic inputs, and virtual reality). Our experimental paradigm was aiming to recreate the scenario of acute stress and demonstrate the change in the response to it. The effects of tVNS in the conditions of chronic stress require further investigation and more complex experimental designs.

Another important consideration is the duration of stress inducement: it was twice shorter than the baseline and recovery, thereby providing much less information for the training of the ML model. This potentially might contribute to the poor performance on the state recognition task. This is further supported by the fact that the classifier seemed to overfit the recovery state slightly (Fig. 7b).

Finally, in our design we were comparing response to a stressful video without and with stimulation within the same individuals to reduce interindividual variability assuming that any placebo effects of the stimulation were relatively small. This was done adhering to the recommendation on conduction of the HRV-based studies.³⁰ A double-blinded placebo controlled study involving a sham stimulation could help further confirm the effects of the stimulation. However, due to controversies on an appropriate tAVNS sham, and complexities in study design, conducting placebo-controlled study was beyond the scope of this research.

Choice of outcome measures

According to von Rosenberg et al.,¹² the LF/HF ratio, which is one of the key metrics in our model, does not perform equally well in different stress scenarios. Their proposed alternative is to look at the pair of values LF, HF, claiming that different combinations of high or low values of the two metrics are instrumental in identifying different kinds of stress. Yet, other research suggests that LF is a poor marker of sympathetic activity as it is influenced by the activity of both sympathetic and parasympathetic systems.⁴² Hence it might be difficult to interpret.

In addition to the HRV metrics that we used, several new “nonlinear” metrics inspired by the behavior of other nonlinear systems have been proposed, such as the sample entropy and correlation dimension.^{5,43} Even though these metrics can be, at least partially, correlated with some of the other time- or frequency-domain metrics, they could still prove useful in our pipeline as they would lead to expanded feature space for the ML analysis. Furthermore, respiration rates could be measured to control for the contribution of Respiratory Sinus Arrhythmia to HF. Finally, other ways of measuring stress, including galvanic skin response, cortisol levels, and questionnaires, could be added into the model.

Data analysis considerations

Spline interpolation of RRI that we used, as described in Section S1.6 in Supplementary Material, does not ensure a monotonic behavior between sampling points and could, therefore, introduce artefacts in which the interpolant overshoots the sampled values, thus introducing high-frequency components. This is prevented when using more sophisticated interpolation algorithms that ensure monotonic behavior, such as cubic interpolation.⁴⁴

The architecture used for ML training was chosen beforehand and it is possible it may benefit from the hyperparameter fine-tuning. However, this would require a separate validation set as well as a test set.

Since we performed all our analysis over a time window, some of the points in the time traces of different HRV metrics contained information from neighboring states. These can be referred to as “transition periods” between the states. If the fraction of these “transition periods” is significant, the change in the metric’s value between different states could be reduced, making it harder to detect. Furthermore, having a better synchronization between collected data and the start/end of states would greatly improve our ability to remove these “transition periods.” However, this is difficult to arrange from both experimental and biological perspectives.

Conclusion

Our findings in this project further support the promise of using taVNS in a variety of different conditions and applications suggested by previous research. In this study, we used custom-made stimulation and recording devices together with a novel method of data analysis on an example application in acute stress. By demonstrating the feasibility of using small snippets of data for determining the presence of stimulation we are getting closer to the possibility of a closed-loop taVNS system. Creating such a system requires a significant amount of further research, technological development, and trials. However, once developed, we believe that its applications will extend beyond stress and could include medical conditions that are currently trialed with taVNS such as depression, epilepsy, irritable bowel syndrome, and rheumatoid arthritis.

Author Disclosure Statement

At the time of writing, the authors were all employees or subcontractors of BrainPatch Ltd. N.V. is the founder of BrainPatch Ltd. and at the time of writing owned more than 25% of the stock. M.O. and A.T. hold minor share options in BrainPatch Ltd. The authors declare no conflict of interest with the study described in this article.

Funding Information

The study described in this project was funded in great part by Innovate UK grant with the title: “Non-invasive safe electrical stimulation using AI to save lives,” grant no. 106744. In other part, the study was funded by BrainPatch Ltd.

Supplementary Material

Supplementary Materials
Supplementary Figure S1
Supplementary Figure S2

References

1. LeBlanc VR. The effects of acute stress on performance: Implications for health professions education. *Acad Med* 2009;84(10 Suppl.):S25–S33. DOI: 10.1097/acm.0b013e3181b37b8f.
2. Garfin DR, Thompson RR, Holman EA. Acute stress and subsequent health outcomes: A systematic review. *J Psychosom Res* 2018;112:107–113. DOI: 10.1016/j.jpsychores.2018.05.017.
3. von Dawans B, Strojny J, Domes G. The effects of acute stress and stress hormones on social cognition and behavior: Current state of research and future directions. *Neurosci Biobehav Rev* 2021;121:75–88. DOI: 10.1016/j.neubiorev.2020.11.026.
4. Kola L, Kohrt BA, Hanlon C, et al. COVID-19 mental health impact and responses in low-income and middle-income countries: Reimagining global mental health. *Lancet Psychiatry* 2021;8:535–550. DOI: 10.1016/s2215-0366(21)00025-0.
5. Shaffer F, Ginsberg JP. An overview of heart rate variability metrics and norms. *Front Public Health* 2017;5. DOI: 10.3389/fpubh.2017.00258.
6. Draghici AE, Taylor JA. The physiological basis and measurement of heart rate variability in humans. *J Physiol Anthropol* 2016;35. DOI: 10.1186/s40101-016-0113-7.
7. Ernst G. Heart-rate variability—More than heart beats? *Front Public Health* 2017;5. DOI: 10.3389/fpubh.2017.00240.
8. Kim H-G, Cheon E-J, Bai D-S, et al. Stress and heart rate variability: A meta-analysis and review of the literature. *Psychiat Invest* 2018;15:235–245. DOI: 10.30773/pi.2017.08.17.
9. Taelman J, Vandeput S, Vlemincx E, et al. Instantaneous changes in heart rate regulation due to mental load in simulated office work. *Eur J Appl Physiol* 2010;111:1497–1505. DOI: 10.1007/s00421-010-1776-0.
10. Castaldo R, Melillo P, Bracale U, et al. Acute mental stress assessment via short term HRV analysis in healthy adults: A systematic review with meta-analysis. *Biomed Signal Process* 2015;18:370–377. DOI: 10.1016/j.bspc.2015.02.012.
11. Hjortskov N, Rissén D, Blangsted AK, et al. The effect of mental stress on heart rate variability and blood pressure during computer work. *Eur J Appl Physiol* 2004;92:84–89. DOI: 10.1007/s00421-004-1055-z.
12. von Rosenberg W, Chanwimalueang T, Adjei T, et al. Resolving ambiguities in the LF/HF ratio: LF-HF scatter plots for the categorization of mental and physical stress from HRV. *Front Physiol* 2017;8. DOI: 10.3389/fphys.2017.00360.
13. Borges U, Laborde S, Raab M. Influence of transcutaneous vagus nerve stimulation on cardiac vagal activity: Not different from sham stimulation and no effect of stimulation intensity. *PLoS One* 2019;14:e0223848. DOI: 10.1371/journal.pone.0223848.
14. Giannakakis G, Marias K, Tsiknakis M. A stress recognition system using HRV parameters and machine learning techniques. In: 2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW) IEEE; 2019. DOI: 10.1109/aciiw.2019.8925142.
15. Dalmeida KM, Masala GL. HRV features as viable physiological markers for stress detection using wearable devices. *Sens* 2021;21:2873. DOI: 10.3390/s21082873.
16. Shim HJ, Kwak MY, An Y-H, et al. Feasibility and safety of transcutaneous vagus nerve stimulation paired with notched music therapy for the treatment of chronic tinnitus. *J Audiol Otol* 2015;19:159–167. DOI: 10.7874/jao.2015.19.3.159.
17. Berthoud H-R, Neuhuber WL. Functional and chemical anatomy of the afferent vagal system. *Auton Neurosci* 2000;85:1–17. DOI: 10.1016/s1566-0702(00)00215-0.

18. Ellrich J. Transcutaneous auricular vagus nerve stimulation. *J Clin Neurophysiol* 2019;36:437–442. DOI: 10.1097/wnp.0000000000000576.
19. Frangos E, Ellrich J, Komisaruk BR. Non-invasive access to the vagus nerve central projections via electrical stimulation of the external ear: fMRI evidence in humans. *Brain Stimul* 2015;8:624–636. DOI: 10.1016/j.brs.2014.11.018.
20. Kraus T, Kiess O, Hösl K, et al. Cns bold Fmri effects of sham-controlled transcutaneous electrical nerve stimulation in the left outer auditory canal—A pilot study. *Brain Stimul* 2013;6:798–804. DOI: 10.1016/j.brs.2013.01.011.
21. Redlich R, Opel N, Bürger C, et al. The limbic system in youth depression: Brain structural and functional alterations in adolescent in-patients with severe depression. *Neuropsychopharmacol* 2017;43:546–554. DOI: 10.1038/npp.2017.246.
22. Breit S, Kupferberg A, Rogler G, et al. Vagus nerve as modulator of the brain–gut axis in psychiatric and inflammatory disorders. *Front Psychiatry* 2018;9. DOI: 10.3389/fpsy.2018.00044.
23. Stefan H, Kreiselmeier G, Kerling F, et al. Transcutaneous vagus nerve stimulation (t-VNS) in pharmacoresistant epilepsies: A proof of concept trial. *Epilepsia* 2012;53:e115–e118. DOI: 10.1111/j.1528-1167.2012.03492.x.
24. Bonaz B, Sinniger V, Hoffmann D, et al. Chronic vagus nerve stimulation in crohn’s disease: A 6-month follow-up pilot study. *Neurogastroenterol Motil* 2016;28:948–953. DOI: 10.1111/nmo.12792.
25. Koopman FA, Chavan SS, Miljko S, et al. Vagus nerve stimulation inhibits cytokine production and attenuates disease severity in rheumatoid arthritis. *Proc Natl Acad Sci* 2016;113:8284–8289. DOI: 10.1073/pnas.1605635113.
26. Bretherton B, Atkinson L, Murray A, et al. Effects of transcutaneous vagus nerve stimulation in individuals aged 55 years or above: Potential benefits of daily stimulation. *Aging* 2019;11:4836–4857. DOI: 10.18632/aging.102074.
27. Tsaava T, Datta-Chaudhuri T, Addorisio ME, et al. Specific vagus nerve stimulation parameters alter serum cytokine levels in the absence of inflammation. *Bioelectronic Medicine* 2020;6. DOI: 10.1186/s42234-020-00042-8.
28. Bremner JD, Gurel NZ, Wittbrodt MT, et al. Application of noninvasive vagal nerve stimulation to stress-related psychiatric disorders. *J Personal Med* 2020;10:119. DOI: 10.3390/jpm10030119.
29. Kogan A, Gruber J, Shallcross AJ, et al. Too much of a good thing? Cardiac vagal tone’s nonlinear relationship with well-being. *Emotion* 2013;13:599–604. DOI: 10.1037/a0032725.
30. Laborde S, Mosley E, Thayer JF. Heart rate variability and cardiac vagal tone in psychophysiological research—Recommendations for experiment planning, data analysis, and data reporting. *Front Psychol* 2017;08:213. DOI: 10.3389/fpsyg.2017.00213.
31. Clancy JA, Mary DA, Witte KK, et al. Non-invasive vagus nerve stimulation in healthy humans reduces sympathetic nerve activity. *Brain Stimul* 2014;7:871–877. DOI: 10.1016/j.brs.2014.07.031.
32. Pittig A, Arch JJ, Lam CWR, et al. Heart rate and heart rate variability in panic, social anxiety, obsessive–compulsive, and generalized anxiety disorders at baseline and in response to relaxation and hyperventilation. *Int J Psychophysiol* 2013;87:19–27. DOI: 10.1016/j.ijpsycho.2012.10.012.
33. Rottenberg J. Cardiac vagal control in depression: A critical analysis. *Biol Psychol* 2007;74:200–211. DOI: 10.1016/j.bio psycho.2005.08.010.
34. Gurel NZ, Huang M, Wittbrodt MT, et al. Quantifying acute physiological biomarkers of transcutaneous cervical vagal nerve stimulation in the context of psychological stress. *Brain Stimul* 2020;13:47–59. DOI: 10.1016/j.brs.2019.08.002.
35. Gurel NZ, Wittbrodt MT, Jung H, et al. Transcutaneous cervical vagal nerve stimulation reduces sympathetic responses to stress in posttraumatic stress disorder: A double-blind, randomized, sham controlled trial. *Neurobiol Stress* 2020;13:100264. DOI: 10.1016/j.ynstr.2020.100264.
36. Redgrave J, Day D, Leung H, et al. Safety and tolerability of transcutaneous vagus nerve stimulation in humans: A systematic review. *Brain Stimul* 2018;11:1225–1238. DOI: 10.1016/j.brs.2018.08.010.
37. Noble LJ, Chuah A, Callahan KK, et al. Peripheral effects of vagus nerve stimulation on anxiety and extinction of conditioned fear in rats. *Learn Memory* 2019;26:245–251. DOI: 10.1101/lm.048447.118.
38. Porges SW. The polyvagal theory: New insights into adaptive reactions of the autonomic nervous system. *Clev Clin J Med* 2009;76(Suppl 2):S86–S90. DOI: 10.3949/ccjm.76.s2.17.
39. Chen CC, Williams CL. Interactions between epinephrine, ascending vagal fibers, and central noradrenergic systems in modulating memory for emotionally arousing events. *Front Behav Neurosci* 2012;6. DOI: 10.3389/fnbeh.2012.00035.
40. Siedlecka E, Denson TF. Experimental methods for inducing basic emotions: A qualitative review. *Emotion Review* 2018;11:87–97. doi. 10.1177/1754073917749016.
41. Ferreira SO. Emotional activation in human beings: procedures for experimental stress induction [in Spanish]. *Psicol USP* 2019;30. DOI: 10.1590/0103-6564e20180176.
42. Houle MS, Billman GE. Low-frequency component of the heart rate variability spectrum: A poor marker of sympathetic activity. *Am J Physiol-Heart Circ Physiol* 1999;276:H215–H223. DOI: 10.1152/ajpheart.1999.276.1.h215.
43. Stein PK, Reddy A. Non-linear heart rate variability and risk stratification in cardiovascular disease. *Indian Pac Electrophysiol J* 2005;5:210.
44. Fritsch FN, Butland J. A method for constructing local monotone piecewise cubic interpolants. *SIAM J Sci Stat Comput* 1984;5:300–304. DOI: 10.1137/0905021.
45. Goldberger AL, Goldberger ZD, Shvilkin A. How to make basic ECG measurements. In: Goldberger AL, Goldberger ZD, Shvilkin A, eds. *Goldberger’s Clinical Electrocardiography*. Elsevier, 2018: 11–20.
46. Makowski D, Pham T, Lau ZJ, et al. NeuroKit2: A python toolbox for neurophysiological signal processing. *Behav Res Methods* 2021;53:1689–1696. DOI: 10.3758/s13428-020-01516-y.
47. Lipponen JA, Tarvainen MP. A robust algorithm for heart rate variability time series artefact correction using novel beat classification. *J Med Eng Technol* 2019;43:173–181. DOI: 10.1080/03091902.2019.1640306.
48. Hannun AY, Rajpurkar P, Haghpanahi M, et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nat Med* 2019;25:65–69. DOI: 10.1038/s41591-018-0268-3.

Address correspondence to:

Nikolai Vysokov, PhD
BrainPatch Ltd.

Unit 324, Edinburgh House, 170 Kennington Lane
London SE11 5DP
United Kingdom

E-mail: nikolai@brainpatch.ai