Technology and Health Care 27 (2019) S407–S424
DOI 10.3233/THC-199037
IOS Press

Entropy analysis of heart rate variability and its application to recognize major depressive disorder: A pilot study

Sangwon Byun[a], Ah Young Kim[b], Eun Hye Jang[b], Seunghwan Kim[b], Kwan Woo Choi[c,d], Han Young Yu[b,*] and Hong Jin Jeon[c,*]
[a]*Department of Electronics Engineering, Incheon National University, Incheon 22012, Korea*
[b]*Bio-Medical IT Convergence Research Division, Electronics and Telecommunications Research Institute, Daejeon 34129, Korea*
[c]*Department of Psychiatry, Depression Center, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul 06351, Korea*
[d]*Department of Psychiatry, Korea University Anam Hospital, Korea University College of Medicine, Seoul 02841, Korea*

Abstract.
BACKGROUND: The current method to evaluate major depressive disorder (MDD) relies on subjective clinical interviews and self-questionnaires.
OBJECTIVE: Autonomic imbalance in MDD patients is characterized using entropy measures of heart rate variability (HRV). A machine learning approach for screening depression based on the entropy is demonstrated.
METHODS: The participants experience five experimental phases: baseline (BASE), stress task (MAT), stress task recovery (REC1), relaxation task (RLX), and relaxation task recovery (REC2). The four entropy indices, approximate entropy, sample entropy, fuzzy entropy, and Shannon entropy, are extracted for each phase, and a total of 20 features are used. A support vector machine classifier and recursive feature elimination are employed for classification.
RESULTS: The entropy features are lower in the MDD group; however, the disease does not have a significant effect. Experimental tasks significantly affect the features. The entropy did not recover during REC1. The differences in the entropy features between the two groups increased after MAT and showed the largest gap in REC2. We achieved 70% accuracy, 64% sensitivity, and 76% specificity with three optimal features during RLX and REC2.
CONCLUSION: Monitoring of HRV complexity changes when a subject experiences autonomic arousal and recovery can potentially facilitate objective depression recognition.

Keywords: Heart rate variability (HRV), major depressive disorder (MDD), entropy, machine learning, depression, feature selection, support vector machine (SVM), recursive feature elimination (RFE), mental task, autonomic nervous system (ANS)

1. Introduction

Nonlinear dynamics and information theory have been applied to physiological phenomena to under-

*Corresponding authors: Han Young Yu, Bio-Medical IT Convergence Research Division, Electronics and Telecommunications Research Institute, Daejeon 34129, Korea. Tel.: +82 42 860 5905; Fax: +82 42 860 6594; E-mail: uhan0@etri.re.kr; Hong Jin Jeon, Department of Psychiatry, Depression Center, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea. Tel.: +82 2 3410 3586; Fax: +82 2 3410 0050; E-mail: jeonhj@ skku.edu.

0928-7329/19/$35.00 © 2019 – IOS Press and the authors. All rights reserved
This article is published online with Open Access and distributed under the terms of the Creative Commons Attribution Non-Commercial License (CC BY-NC 4.0).

stand the complex behavior of underlying biological systems [1]. In particular, cardiovascular regulation is known to originate from nonlinear regulatory subsystems mediated by an autonomic nervous system (ANS) [2,3]: The interaction between sympathetic and vagal nerve stimulation on the heart rate control is not simply static, but rather dynamic and nonlinear [4]. Cardiac regulatory dynamics have been generally investigated using time series data extracted from electrocardiogram (ECG), i.e., the time intervals between consecutive R-peaks in the ECG signal (RRI). Heart rate variability (HRV), which represents variations in the RRIs, has been used to assess and model autonomic cardiac regulation [5]. Although HRV has been analyzed traditionally using linear methods, such as time- and frequency-domain analyses [5], growing evidence has demonstrated that linear HRV measures may not correctly represent the complex dynamics of heartbeat regulation modulated by the ANS [1–3] and that linear HRV features show a relatively higher inter-subject variability than nonlinear HRV indices [6], suggesting the importance of nonlinear HRV analysis.

Based on these results, nonlinear HRV analyses, including entropy, Lyapunov exponent, detrended fluctuation analysis, symbolic dynamics, or Poincaré plots, have been adopted to evaluate ANS activity in various clinical diseases. For example, nonlinear measures were used to model heartbeat dynamics in patients with ventricular arrhythmia [7], diabetes [8], Parkinson's disease [9], and sleep apnea [10], as they differ considerably for healthy subjects. These results suggest that nonlinear and non-stationary characteristics of cardiac regulation can reveal relevant information on the pathological states of the ANS. Reduced complexity in HRV has been generally interpreted as an outcome of the pathological ANS dynamics of diseases, indicating the subject's lack of ability to adapt to a changing environment [11].

Among these applications based on HRV, assessment of psychiatric disorders has great potential because conventional methods to evaluate most psychiatric disorders rely on verbal interviews and scores from subjective self-questionnaires [12]. Therefore, previous studies have attempted to find a more objective method to evaluate psychiatric disorders, using plasma biomarkers [13], functional imaging [14], salivary cortisol [15], vocal acoustic features [16], and physiological signals, such as electroencephalogram [17], and electrodermal activity [18]. In addition, nonlinear HRV measures have been used to study psychiatric disorders, such as major depressive disorder (MDD) [6], bipolar disorder [19,20], and schizophrenia [21]. Given that HRV measurement is non-invasive and easy-to-quantify, this technique can facilitate objective evaluation and accurate identification of patients with psychiatric disorders. Furthermore, with recent advances in machine learning methods, automated diagnosis systems for psychiatric disorders have been demonstrated based on nonlinear HRV analysis. For example, Valenza et al. demonstrated the feasibility of predicting mood changes between depression and euthymia in bipolar patients using nonlinear HRV features [22]. HRV entropy combined with other physical activity measures has been successfully used to classify patients with schizophrenia from healthy controls [23].

In this study, we demonstrate the entropy analysis of HRV measured from patients with MDD and its application for automated MDD detection using machine learning approach. Entropy-based measures have been widely used to quantify complexity in various biomedical time series datasets, including RRI data [2,3,24]. Entropy analysis can be applied to non-stationary data and is more resistant to noise than the linear HRV indices. Furthermore, compared to other nonlinear measures such as the fractal dimension, correlation dimension, and Lyapunov exponents, entropy measures require a shorter length of data [25,26]. This characteristic suggests that entropy measures may provide more reliable results from short-term HRV analysis (e.g., 5 min) than other nonlinear HRV features.

Depression is a common disease with a significant economic burden, which is characterized by a persistent low mood, fatigue, insomnia, anhedonia, restlessness, and substantial weight loss or gain [27,28].

Major depression is also associated with an increased risk of coronary heart disease, and the disturbed ANS system is one of the mechanisms underlying a link between MDD and cardiovascular disease [29]. Untreated severe depression can lead to suicidal behavior, which is a serious social issue [30], affirming the need to develop an objective diagnosis technique for MDD that can compensate the conventional methods. Our study is inspired by several previous studies, which have demonstrated that HRV-based automated diagnosis systems for MDD using machine learning methods [31–35]. The aforementioned studies applied neuro-fuzzy networks, support vector machine (SVM), linear discrimination analysis, logistic regression, and Bayesian networks on HRV features to differentiate MDD patients from healthy controls. These studies achieved accuracies between 71–95%. However, their results were partially or entirely dependent on linear HRV features. Machine learning-based diagnosis of MDD considering nonlinear HRV features has not been studied in detail. To the best of our knowledge, classification of major depression based solely on entropy features has not been reported to date.

We consider four entropy measures: approximate entropy (ApEn), sample entropy (SampEn), fuzzy entropy (FuzzEn), and Shannon entropy (ShanEn). Pincus first proposed the concept of ApEn to assess time series irregularity and to represent signal randomness [3]. SampEn and FuzzEn were developed later as improved versions of ApEn [24,36]. In addition, ShanEn, which is based on information theory and used to evaluate the distribution complexity of heartbeat signal samples, has also been introduced [37]. Entropy features measured from patients with MDD have been previously reported. For example, the HRV ApEn and SampEn values obtained from MDD patients are significantly lower than those of healthy subjects [38–41]. In addition, the HRV ShanEn is significantly reduced in patients with MDD [42] or bipolar disorders [43], suggesting that entropy measures are substantially affected by depressive symptoms.

Therefore, we first aim to analyze the entropy features measured from patients with MDD and healthy controls. In addition to the measurement of the basal ANS activity, we consider mental stress to examine the ANS reactivity to stimuli and the recovery from the stimuli. Previous studies have demonstrated dysfunctional autonomic responses to stimuli such as emotional cues and stress, resulting in attenuated or exaggerated HRV reactivity [44–48]. Furthermore, patients with MDD showed weaker stress recovery than the healthy controls [49]. The second aim is to demonstrate automated differentiation of patient with MDD using the entropy features as input data. To this end, feature selection is performed by support vector machine-recursive feature elimination (SVM-RFE), and four machine learning algorithms are employed. We expect that multiple autonomic alterations during the measurement may improve the discriminative power of HRV features [50].

2. Materials and methods

2.1. Participants

All participants were recruited at the Samsung Medical Center, Seoul, Korea. A total of 33 MDD patients and 33 healthy controls participated in the experiment, and they were matched for age and gender. A senior psychiatrist evaluated the patients with MDD based on the DSM-IV criteria. The depressive episode severity was evaluated using the Hamilton depression rating score (HAMD) [51]. Healthy subjects who had no history of psychiatric disorders were recruited to be part of the control group. The experimental procedure was explained to all participants, and they submitted signed informed consent. As compensation, $50 was paid to each participant. This study was approved by the Institutional Review Board of the Samsung Medical Center, Seoul, Korea (No. 2015-07-151), and performed according to the relevant guidelines.

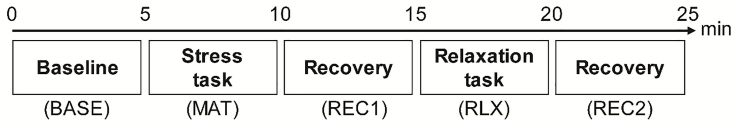


Fig. 1. Experimental protocol for ECG measurement. ECG signals were recorded during five consecutive phases, each of which had 5 min duration.

2.2. Experimental procedure

All subjects entered a room and sat in an armchair with a headrest. The experimenter explained the measurement procedure in detail while attaching the electrodes and sensors. The protocol used in this study was designed to assess autonomic reactivity to mental stress and relaxation tasks (Fig. 1) [50]. The entire protocol consisted of five consecutive phases: the baseline (BASE), mental stress task (MAT), recovery from the stress task (REC1), relaxation task (RLX), and recovery from the relaxation task (REC2) phases. Each phase was 5-min long. During BASE, the subjects were instructed to rest and minimize their movement to measure basal ANS activity. During MAT, a mental arithmetic test was conducted by the subjects, who were asked to serially subtract 7 starting from 500 and to report answers verbally [52]. The mental arithmetic test has been widely used as a psychological stressor to perturb the cardiovascular system and measure autonomic reactivity in depressed patients [45]. During REC1, the subjects were instructed to relax to allow autonomic recovery from the mental task. During RLX, the subjects were asked to rest while watching 10 30-s long images of natural scenery presented on a PC monitor. In the final phase (REC2), the image presentation was stopped, and the subjects were instructed to rest to allow autonomic recovery from the relaxation task.

2.3. Physiological measurement

Electrocardiogram (ECG) signals were measured using the ProComp Infiniti system (SA7500, Thought Technology, Canada). Three Ag/AgCl electrodes were attached on the wrists and an ankle in a lead II configuration. The measured signal was amplified, band-pass filtered, and sampled at 256 Hz. We used Kubios HRV Premium software (Kubios, www.kubios.com) to analyze the ECG signals. This system implements an in-house QRS detection algorithm based on the Pan-Tompkins method [53,54]. All detected R-peaks were inspected by the same operator to maintain consistency. A piecewise cubic spline interpolation method was used to correct artifacts. Finally, the entropy features were calculated separately from the individual phases using the RRI data.

2.4. Approximate entropy

Pincus first suggested ApEn to measure time-series data irregularity. In this work, we followed the originally proposed ApEn definition [3].

First, a sequence of vectors $\boldsymbol{x}(1), \boldsymbol{x}(2), \ldots, \boldsymbol{x}(N-m+1)$ in $\boldsymbol{R}^m$ is formed, defined by

$$\boldsymbol{x}(i) = [u(t_i), u(t_{i+1}), \ldots, u(t_{i+m-1})]. \tag{1}$$

Then, for each $i, 1 \leqslant i \leqslant N-m+1$,

$$C_i^m(r) = \frac{\text{number of } j \text{ such that } d[\boldsymbol{x}(i), \boldsymbol{x}(j)] \leqslant r}{N-m+1}, \tag{2}$$

where $d[\boldsymbol{x}(i), \boldsymbol{x}(j)]$ is defined by

$$d[\boldsymbol{x}(i), \boldsymbol{x}(j)] = \max(|u(t_{i+k-1}) - u(t_{j+k-1})|, 1 \leqslant k \leqslant m). \tag{3}$$

Next, we define

$$\Phi^m(r) = (N - m + 1)^{-1} \sum_{i=1}^{N-m+1} \log C_i^m(r). \tag{4}$$

Finally, ApEn is defined as

$$\text{ApEn}(m, r, N) = \Phi^m(r) - \Phi^{m+1}(r). \tag{5}$$

The value of ApEn depends on three parameters, i.e., the embedding dimension, m; the tolerance value, r; and the data length, N. In the present study, m was set to 2 and r was selected as 0.2 times the standard deviation of the data, as they are commonly used for clinical HRV data [3,55].

2.5. Sample entropy

SampEn was proposed to decrease bias in ApEn, which is caused by self-comparison of vector $\boldsymbol{x}(i)$ in the calculation of $C_i^m(r)$ [24]. To eliminate this self-comparison, $C_i^m(r)$ is defined as

$$C_i^m(r) = \frac{\text{number of } j \text{ such that } d[x(i), x(j)] \leqslant r}{N - m + 1}, (i \neq j). \tag{6}$$

Next, we define

$$C^m(r) = (N - m + 1)^{-1} \sum_{i=1}^{N-m+1} C_i^m(r). \tag{7}$$

Then, SampEn is defined as

$$\text{SampEn}(m, r, N) = \log \frac{C^m(r)}{C^{m+1}(r)}. \tag{8}$$

In this study, the values of m and r were the same as those for the ApEn calculation [55].

2.6. Fuzzy entropy

FuzzEn was introduced by Chen et al. to reduce the parameter dependence of SampEn [36]. The vector similarity in ApEn and SampEn is defined based on the Heaviside function. However, FuzzEn is derived from the concept of fuzzy sets, with the Heaviside function being replaced by an exponential function. First, all vectors are normalized by removing a baseline

$$u_0(t_i) = \frac{1}{m} \sum_{j=0}^{m-1} u(t_{i+j}). \tag{9}$$

Therefore, $d[\boldsymbol{x}(i), \boldsymbol{x}(j)]$ is defined by

$$d[\boldsymbol{x}(i), \boldsymbol{x}(j)] = \max(|(u(t_{i+k-1}) - u_0(t_i)) - (u(t_{j+k-1}) - u_0(t_j))|, 1 \leqslant k \leqslant m). \tag{10}$$

The similarity degree, $D_{ij}^m(r)$, is calculated as

$$D_{ij}^m(r) = \exp(-(d[\boldsymbol{x}(i), \boldsymbol{x}(j)])^n / r), (i \neq j). \tag{11}$$

The following relation is then defined:

$$\Phi^m(r) = (N-m)^{-1} \sum_{i=1}^{N-m} \left((N-m-1)^{-1} \sum_{j=1, j\neq i}^{N-m} D_{ij}^m(r) \right). \tag{12}$$

Finally, FuzzEn is defined as

$$\text{FuzzEn}(m, n, r, N) = \log \frac{\Phi^m(r)}{\Phi^{m+1}(r)}. \tag{13}$$

In this study, the values of m and r were the same as those for the SampEn calculation, and n was set to 2 [56].

2.7. Shannon entropy

ShanEn is used to assess the complexity of the RRI series distribution based on information theory [57, 58]. First, $\boldsymbol{RR} = \{RR(i), i = 1, 2, \ldots, N\}$ is defined, where N is the series length. The RR series is then coarse-grained using a uniform quantization procedure. It is spread on ξ quantization bins of amplitude equal to $\varepsilon = (RR_{\max} - RR_{\min})/\xi$. The values in each bin are replaced by integers ranging from 0 to $\xi - 1$. $\boldsymbol{RR}$ becomes a quantized integer series: $\boldsymbol{RR}^\xi = \{RR^\xi(i), i = 1, 2, \ldots, N\}$. Then, a pattern of L delayed samples is defined: $RR_L^\xi(i) = (RR^\xi(i), RR^\xi(i-1), \ldots, RR^\xi(i-L+1))$, with $\boldsymbol{RR}_L^\xi = \{RR_L^\xi(i), i = 1, 2, \ldots, N-L+1\}$. Finally, ShanEn is defined as

$$\text{ShanEn}(L, \xi) = -\sum_{i=1}^{N-L+1} p(RR_L^\xi(i)) \log(p(RR_L^\xi(i))), \tag{14}$$

where $p(RR_L^\xi(i))$ is the probability of the $RR_L^\xi(i)$ pattern formed by the current value, $RR_L^\xi(i)$, and by the $L-1$ past samples, $RR_{L-1}^\xi(i-1)$. In the present study, L and ξ were set to 3 and 6, respectively, as they are commonly used for clinical HRV data [37,59].

2.8. Leave-one-out procedure and feature selection

Figure 2 shows an overview of the data processing procedure. A leave-one-out (LOO) method was used to split the data and evaluate performance of the classifier. We used an SVM-RFE algorithm as a feature selection method [60]. The RFE algorithm is a backward sequential selection method, which eliminates the least relevant features individually during each iteration. For this study, the feature importance was decided by the weights calculated from an SVM algorithm with a linear kernel function. All procedures were performed using MATLAB R2018a (MathWorks, Natick, MA). The feature selection was performed using Library for SVM (LIBSVM) [61] and the RFE algorithm from the Mathworks File Exchange [62].

For the LOO procedure, the entropy data were normalized by subtracting the median and dividing by the mean absolute deviation (MAD). Four supervised machine learning algorithms, i.e., SVM, linear discriminant analysis (LDA), k-nearest neighbors (k-NN), and Naïve Bayes (NB), were implemented to classify the control and MDD groups using the entropy features as input data. (These classifiers are discussed in more detail in the following subsection.) To ensure all classifiers were trained using the same training set and the same feature subset, SVM-RFE was applied to a given training set and the feature ranking was determined. Then, the ranking information and the same training set were used

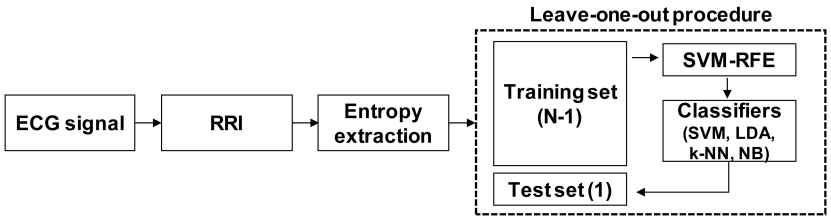


Fig. 2. Overview of data processing procedure.

to train the SVM, LDA, k-NN, and NB classifiers. The prediction performance was evaluated using the unseen test set. This process was repeated $N = 66$ times. Then, the overall performance measures were assessed by averaging the results from all 66 folds. Cross-validation was performed external to the feature selection process, so as to represent the prediction performance accurately. We used the accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) as performance measures. We also computed the average rank for each feature over all folds to assess the relative importance of the feature.

2.9. Classifiers

We present a brief description of each classifier here. Further details on these machine learning algorithms can be found in the literature [63–65]. An SVM algorithm is based on the margin maximization principle [65]. The input data is mapped into a high-dimensional feature space, and in this space, an optimal separating hyperplane is constructed, which maximizes the interclass distances [66]. We used a linear kernel function for both the SVM-RFE and the classifier model. We employed a fixed value of the regularization parameter $C = 0.9$ based on a grid search performed for an optimal C value in the range of 0.01–100.

An LDA assumes a normal distribution of the data with an equal covariance matrix [67]. It seeks a projection that maximizes the ratio of the between-class variance to the within-class variance [63]. As the LDA requires very low computational power and is simple to implement, it is suitable for classifying various physiological signals [67,68].

Construction of a k-NN model involves storage of the training dataset only [64]. To predict a class label for a new data point, this algorithm finds the k closest data points in the training set according to a distance metric [68]. When more than one neighbor is considered, the majority class of the k-nearest neighbors is assigned to the new data point [68]. In this study, we used the Euclidean distance as a distance metric and tested four different odd numbers for k (1, 3, 5, and 7). The best performance was achieved with $k = 7$, and we report the results from this model herein.

An NB classifier is a probabilistic model derived from Bayes' theorem, which relies on two assumptions: (1) the predictive attributes are assumed to be conditionally independent given the class; (2) it is assumed that no hidden attributes affect the prediction process [69]. Under those assumptions, the Bayesian classifier uses a training dataset to estimate the parameters of the feature probability distribution. Then, it calculates the posterior probability of the newly introduced test data and chooses the class with the maximum value. In the present study, we assumed that the entropy measures followed a Gaussian distribution.

2.10. Statistical analyses

Statistical analyses were performed using MATLAB R2018a, SPSS 25 (SPSS Inc., Chicago, IL, USA),

Table 1
Demographic and clinical characteristics of control and MDD groups

Factors	Control ($N = 33$)	MDD ($N = 33$)	p value
Sex (%)			1.000 (χ^2)
M	9 (27%)	9 (27%)	
F	24 (73%)	24 (73%)	
Age (SD)	40.21 (15.16)	40.18 (16.10)	0.753
Education, years (SD)	14.42 (2.40)	13.15 (3.23)	0.163
Marital status (%)			0.124 (χ^2)
Single	14 (42%)	14 (42%)	
Married	19 (58%)	14 (42%)	
Divorced	0 (0%)	2 (6%)	
Bereavement	0 (0%)	3 (9%)	
BMI (SD)	23.19 (3.13)	22.45 (5.41)	0.832
HAMD (SD)	1.91 (1.83)	17.42 (7.00)	$< 0.001^{***}$

The factors of the two groups were compared using the Mann-Whitney U test ($^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$), except for the sex, marital status, and smoking factors, which were compared using the chi-square test (χ^2). Abbreviations: BMI, body mass index; HAMD, Hamilton depression rating score.

and R software 3.5.1 (The R Foundation for Statistical Computing, Vienna, Austria). The age, years of education, body mass index (BMI), and HAMD were compared between the MDD and control groups using Mann-Whitney U tests as these factors were not normally distributed. The sex and marital status were compared between the two groups using chi-square tests. ApEn, SampEn, FuzzEn, ShanEn, and RRI were tested for the assumptions of normality and homogeneity of variance required for an analysis of variance (ANOVA) using the Shapiro-Wilk test and Levene's test, respectively. SampEn, FuzzEn, and RRI satisfied both assumptions. Therefore, we used a repeated-measures ANOVA to test the effects of group and task on those features. In contrast, ApEn and ShanEn violated the normality assumption. For those features, we used the non-parametric equivalent of a repeated-measures ANOVA to test the effects of group and task using the R statistics package "nparLD" [70]. This package included a rank-based non-parametric method, which utilized estimators of relative treatment effects to test longitudinal data in factorial experiments [71]. Hence, ANOVA-type statistics for examination of the hypotheses were provided. Statistical tests of the main effects were followed by post-hoc pairwise comparisons between different tasks, which were corrected using the Bonferroni method. In all statistical tests, a P-value of less than or equal to 0.05 was considered significant.

3. Results

3.1. Demographic and clinical characteristics of participants

The participants for the present study were 33 healthy subjects (24 females) and 33 MDD patients (24 females). Descriptive statistics of the demographic and clinical characteristics of the participants are summarized in Table 1. No significant differences were observed for sex, age, years of education, marital status, or BMI. The MDD group showed significantly higher HAMD than the control group ($P < 0.001$).

3.2. RRI measurement

We examined the influence of group and task on RRI to understand the effects of the MDD and our experimental protocol on the autonomic activity of the participants. The RRI changes during measurement are shown in Fig. 3. We used a repeated-measures ANOVA to statistically test the effects of group

Table 2
A. Effects of group and task on RRI tested with repeated-measures ANOVA; B. The simple main effect of the task was significant for both MDD and control groups ($P < 0.001$)

A.						
Feature	Group effect (MDD vs. control)		Task effect (5 tasks)		Interaction (group × task)	
	F	P-value	F	P-value	F	P-value
RRI	0.10	0.754	77.08	$< 0.001^{***}$	3.59	0.036^{*}

B.			
Pairwise comparison	Bonferroni adjusted P		
	MDD	Control	
BASE vs. MAT	$< 0.001^{***}$	$< 0.001^{***}$	
BASE vs. REC1	1.000	1.000	
BASE vs. RLX	0.659	1.000	
BASE vs. REC2	0.518	1.000	
MAT vs. REC1	$< 0.001^{***}$	$< 0.001^{***}$	
MAT vs. RLX	$< 0.001^{***}$	$< 0.001^{***}$	
MAT vs. REC2	$< 0.001^{***}$	$< 0.001^{***}$	
REC1 vs. RLX	$< 0.005^{**}$	$< 0.044^{*}$	
REC1 vs. REC2	$< 0.004^{**}$	$< 0.044^{*}$	
RLX vs. REC2	1.000	1.000	

Post-hoc pairwise comparisons between tasks were performed for MDD and control groups separately, and P-values were corrected using the Bonferroni method ($^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$).

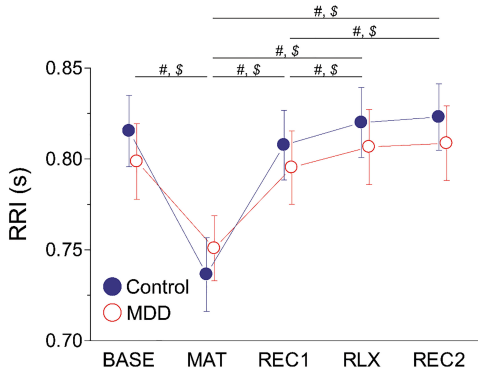


Fig. 3. Mean ± SD values of RRI. The effects of group and task on RRI were tested using a repeated-measures ANOVA. The main effect of group was non-significant, but there were significant main effects of task ($P < 0.001$) and interaction ($P = 0.036$). The simple main effect of the task was significant for both MDD and control groups ($P < 0.001$). Post-hoc pairwise comparisons between tasks were performed for MDD and control groups separately, and P-values were corrected using the Bonferroni method ($^{\#}P < 0.05$ for MDD, $^{\$}P < 0.05$ for control).

and task (Table 2). The main effect of group was non-significant, but significant main effects of task ($P < 0.001$) and interaction ($P = 0.036$) on RRI were found. Therefore, we tested the simple main effect of the task. For both MDD and control groups, the effect of the task was significant ($P < 0.001$). Post-hoc comparisons between tasks were performed separately for two groups. In both groups, pairwise comparisons revealed that the RRI was significantly reduced during MAT compared to the other phases. In addition, the RRI from REC1 was significantly shorter than those from RLX and REC2.

3.3. Entropy analyses

Similar to the RRI, we tested the effects of group and task on each entropy measure. The ApEn, Sam-

Table 3
Mean (SD) values of ApEn, SampEn, FuzzEn, and ShanEn measured during each phase in control ($N = 33$) and MDD groups ($N = 33$)

Phase	BASE	MAT	REC1	RLX	REC2
Entropy	ApEn				
Control	1.105 (0.104)	1.122 (0.140)	1.077 (0.115)	1.119 (0.082)	1.118 (0.078)
MDD	1.097 (0.102)	1.135 (0.116)	1.063 (0.103)	1.076 (0.110)	1.059 (0.113)
Entropy	SampEn				
Control	1.472 (0.266)	1.443 (0.383)	1.369 (0.311)	1.459 (0.237)	1.450 (0.215)
MDD	1.427 (0.277)	1.412 (0.281)	1.307 (0.265)	1.351 (0.261)	1.300 (0.253)
Entropy	FuzzEn				
Control	0.819 (0.254)	0.708 (0.271)	0.713 (0.226)	0.791 (0.201)	0.744 (0.181)
MDD	0.769 (0.221)	0.691 (0.213)	0.672 (0.223)	0.753 (0.223)	0.680 (0.209)
Entropy	ShanEn				
Control	3.431 (0.417)	3.418 (0.534)	3.264 (0.371)	3.441 (0.354)	3.283 (0.344)
MDD	3.340 (0.452)	3.389 (0.412)	3.188 (0.385)	3.208 (0.433)	3.139 (0.450)

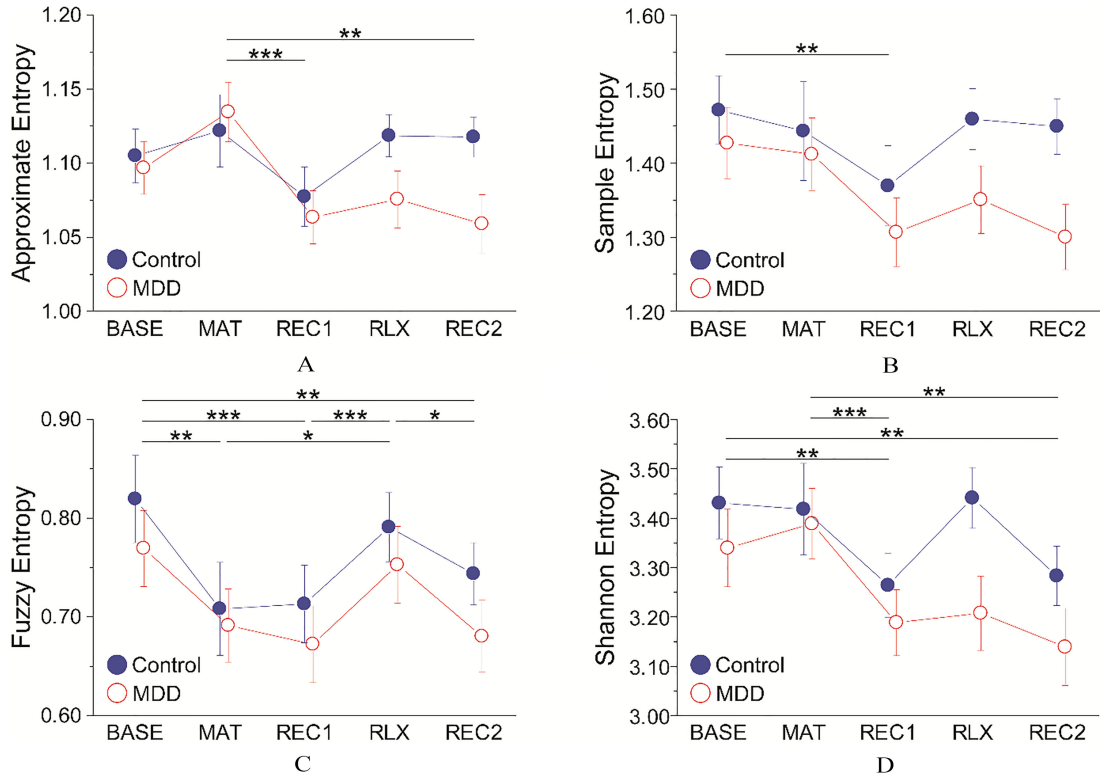


Fig. 4. Mean ± SD values of A. ApEn, B. SampEn, C. FuzzEn, and D. ShanEn. All four entropy features were significantly affected by the main effect of task. However, there were no significant effects of group or interaction on any entropy measures. Post-hoc pairwise comparisons between tasks were corrected using the Bonferroni method ($^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$).

pEn, FuzzEn, and ShanEn changes during measurement are presented in Fig. 4 and Table 3. SampEn and FuzzEn were tested with repeated-measures ANOVA. However, we used the non-parametric equivalent of a repeated-measures ANOVA for ApEn and ShanEn, as they violated the normality assumption (for further details, see the Methods). All four entropy features were significantly affected by the main

Table 4
A. Effects of group and task on entropy. Statistical analyses for SampEn and FuzzEn were performed with a repeated-measures ANOVA. Statistical analyses for ApEn and ShanEn were performed with the non-parametric equivalent of a repeated-measures ANOVA using the R statistics package "nparLD". B. Post-hoc pairwise comparisons between tasks were corrected using the Bonferroni method ($^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$)

A.

Feature	Group effect (MDD vs. control)		Task effect (5 tasks)		Interaction (group × task)	
	F	P-value	F	P-value	F	P-value
ApEn	1.55	0.213	6.56	$< 0.001^{**}$	1.63	0.166
SampEn	2.08	0.154	3.67	0.010^{*}	1.14	0.337
FuzzEn	0.78	0.382	8.70	$< 0.001^{***}$	0.30	0.854
ShanEn	1.27	0.259	7.74	$< 0.001^{***}$	1.28	0.279

B.

Pairwise comparison	Bonferroni adjusted P			
	ApEn	SampEn	FuzzEn	ShanEn
BASE vs. MAT	0.088	1.000	0.005^{**}	1
BASE vs. REC1	0.079	0.008^{**}	$< 0.001^{***}$	0.001^{**}
BASE vs. RLX	1.000	1.000	1.000	1
BASE vs. REC2	1.000	0.068	0.001^{**}	0.003^{**}
MAT vs. REC1	$< 0.001^{***}$	0.234	1.000	$< 0.001^{***}$
MAT vs. RLX	0.051	1.000	0.025^{*}	1
MAT vs. REC2	0.003^{**}	1.000	1.000	0.003^{**}
REC1 vs. RLX	0.273	0.205	$< 0.001^{***}$	0.091
REC1 vs. REC2	1.000	1.000	1.000	1
RLX vs. REC2	1.000	1.000	0.011^{*}	0.185

effect of task (Table 4). However, there were no significant effects of group or interaction on any entropy measures. Pairwise comparisons between tasks revealed the following (Table 4): ApEn was significantly higher during MAT than in REC1 and REC2. SampEn was reduced significantly during REC1 compared to BASE. The BASE FuzzEn was significantly higher than those for MAT, REC1, and REC2. The RLX FuzzEn was significantly higher than those for MAT, REC1, and REC2. The BASE ShanEn was significantly higher than those for REC1 and REC2. The MAT ShanEn was significantly higher than those for REC1 and REC2.

3.4. Classification of control and MDD participants using entropy features

Four supervised machine learning algorithms (SVM, LDA, k-NN, and NB) were implemented to classify healthy and MDD participants using a total of 20 entropy features (i.e., four entropy features from five phases) as input data. SVM-RFE performed feature selection The performance of classification were evaluated using the LOO procedure, as shown in Fig. 2.

Figure 5 shows the prediction accuracy as a function of the number of features selected by the SVM-RFE method. The best performance results for each classifier are summarized in Table 5. SVM outperformed the LDA, k-NN, and NB classifiers and achieved 70% accuracy, 64% sensitivity, and 76% specificity using the three most relevant features. The best performance measures of LDA were similar to those of SVM, and were also based on the same three optimal features. The k-NN accuracy peaked with the first feature, but use of additional features did not yield better performance. The NB classifier required nine relevant features to achieve 67% accuracy. When all 20 features were utilized for classification, the SVM, LDA, k-NN, and NB models achieved only 50%, 42%, 39%, and 59% accuracy, respectively, indicating that the feature selection substantially improved performance of the classifiers.

Table 5
Best performance of each classifier

Classifier	NF	ACC (%)	SEN (%)	SPE (%)	PPV (%)	NPV (%)
SVM	3	70	64	76	72	68
LDA	3	68	64	73	70	67
k-NN	1	67	67	67	67	67
NB	9	67	58	76	70	64

Abbreviations: NF, number of features; ACC, accuracy; SEN, sensitivity; SPE, specificity.

Table 6
Average ranks of 20 entropy measures determined by SVM-RFE

$Rank_{Avg}$	Phase	Entropy
1.00	REC2	SampEn
2.00	REC2	FuzzEn
3.42	RLX	ShanEn
6.11	RLX	ApEn
7.08	BASE	FuzzEn
7.24	MAT	ApEn
7.97	BASE	ApEn
8.47	MAT	FuzzEn
10.32	MAT	ShanEn
10.33	RLX	SampEn
11.73	MAT	SampEn
12.35	REC1	ShanEn
13.09	BASE	SampEn
13.73	RLX	FuzzEn
14.03	REC1	FuzzEn
14.59	REC2	ApEn
15.09	REC1	SampEn
16.12	REC2	ShanEn
16.29	REC1	ApEn
19.05	BASE	ShanEn

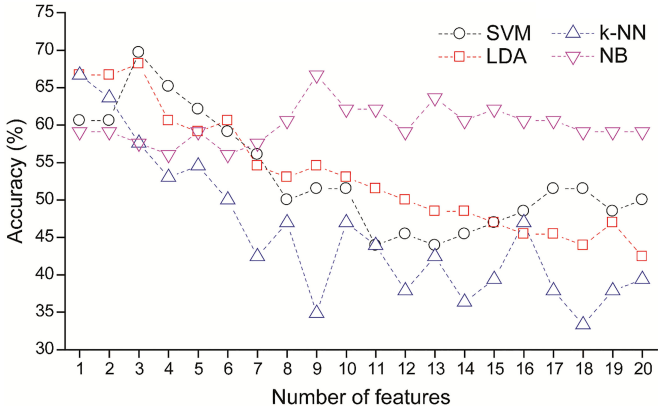


Fig. 5. Average accuracy as function of number of selected features. The features were ranked by SVM-RFE.

The average ranks of the 20 entropy features are listed in Table 6. These ranks were computed by SVM-RFE during the LOO procedure. The most relevant feature was SampEn from REC2, followed by FuzzEn from REC2, ShanEn from RLX, ApEn from RLX, and FuzzEn from BASE.

4. Discussion

We have demonstrated differentiation of patients with MDD from healthy controls with 70% accuracy, 64% sensitivity, and 76% specificity, using the three most relevant entropy features. To the best of our knowledge, the present study is the first to classify MDD patients and healthy controls using only HRV entropy features. The feature ranking determined by SVM-RFE revealed that the features extracted from the REC2 and RLX phases were highly relevant, suggesting that the entropy measured when a subject is experiencing autonomic stress and recovery can improve the discriminative power of HRV entropy features.

First, we examined whether our experimental protocol induced multiple alterations in the ANS activities of participants during measurements (Fig. 3 and Table 2). The significant decrease in RRI during the MAT phase observed in this study suggests that the mental stress task successfully induced increased sympathetic activity. The participants recovered from this stress during the first recovery (REC1), as indicated by a significant difference in RRI between the MAT and REC1 phases. Interestingly, the RRI values increased significantly in the RLX and REC2 phases compared to the REC1 phase. These results suggest that REC1 was not sufficiently long for the participants to fully recover from the mental stress and, also, that the participants further recovered during the relaxation task (RLX) and the second recovery (REC2).

Previous studies have generally shown that entropy features are significantly reduced in MDD patients compared to healthy participants [38–40]. Conflicting results have also been reported, for which no significant differences in ApEn and SampEn were obtained for MDD and control groups [72–74]. In this study, mean values of entropy features were lower in the MDD group than in the control group except for the ApEn from MAT (Table 3). However, there were no significant differences in entropy features between the control and patient groups (Table 4). This inconsistency is likely a result of the heterogeneous presentation and multifactorial etiology of MDD, which substantially increase variability in the assessment of ANS activity [75].

Our experimental tasks affected the entropy features, but the entropy responded differently than the RRI (Fig. 4 and Table 4). For example, the RRI was significantly reduced during the stress task compared with the baseline, before recovering during the subsequent phases. However, the entropy continued to decrease and did not recover in REC1, suggesting that entropy features (but not RRI) may reflect prolonged sympathetic excitation in the recovery phase. In addition, although the mean RRI values were higher in REC2 than in RLX, the entropy features were lower in REC2 than in RLX. FuzzEn and ShanEn were significantly decreased in REC2 compared to BASE, indicating that they did not achieve full recovery to the initial values, unlike the RRI values. Similar to the RRI, SampEn and FuzzEn were lower in MAT than in BASE. Mental stress typically induces a decrease in HRV complexity, including entropy features [76–78]. The entropy was increased in the RLX phase, indicating that the relaxation task induced vagal activity. These results suggest that the entropy features used in the present study could reflect changes in the balance between sympathetic and parasympathetic cardiac modulation during our experimental protocol, as demonstrated by a previous study [76].

Based on these entropy features, we achieved 70% accuracy in differentiating the patients with MDD using the SVM classifier (Table 5). The feature selection substantially improved the performance of all classifiers. For example, the SVM classifier accuracy was 50% when all features were used, but increased to 70% when the three most relevant features were used (Fig. 5). The top three most relevant features were from REC2 and RLX (Table 6), suggesting that observation of the baseline ANS activity may not be ideal for distinguishing MDD patients from the healthy population. Similarly, Schumann et

al. suggested that the resting HRV evaluates cardiac vagal modulation only, and lacks information on sympathetic modulation [79]. The most relevant feature in the present study was SampEn from REC2. FuzzEn from REC2 and ShanEn from RLX ranked second and third, respectively. Interestingly, the difference in SampEn between the MDD and control groups increased after the MAT phase and showed the largest gap in REC2 (Fig. 4 and Table 3). A similar trend was observed for the other entropy features. For FuzzEn and ShanEn, the differences between the two groups were larger in REC1, RLX, and REC2 compared to MAT. A previous study demonstrated that patients with MDD exhibit reduced autonomic recovery from laboratory stress [49]. This may explain the growing entropy difference between the MDD and control groups after the stress task. Therefore, our feature selection results suggest that analyzing HRV entropy in response to autonomic stimuli and recovery from those stimuli may reflect differences in ANS reactivity between MDD and control subjects. In previous studies on HRV-based classification of depressive symptoms, various autonomic stimuli were also adopted during HRV measurements, such as emotion elicitation [31], random number generation [33], and Ewing tests [34]. This approach revealed the disrupted autonomic cardiac control in depressed patients.

A relatively small number of subjects (MDD = 33, control = 33) is a limitation of our study. Given that major depression exhibits multiple heterogeneous sources of variability, recruiting a large number of subjects is crucial for obtaining reliable results. We are recruiting more patients and will conduct a future study based on the expanded cohort. It is important to note that treatment of depression with antidepressant medications has been associated with changes in HRV. Tricyclic medication has been reported to reduce HRV [80,81]. Findings on the effect of selective serotonin reuptake inhibitors (SSRIs) on HRV have been inconsistent: a longitudinal cohort study found that SSRIs decreased HRV [81], but a meta-analysis of antidepressant treatment trials reported no such change [80]. These findings suggest that antidepressants could be responsible for the results in the present study although direct comparisons with previous literature are difficult due to methodological issues [82].

5. Conclusion

We have found that ANS stimulation during measurement was crucial for revealing an altered heartbeat regulation of depressed patients, as these characteristics were not manifested in the baseline activity. In particular, the differences in the entropy features between the MDD and control groups increased after the stress phase and showed the largest gap in the final recovery phase. Similarly, the feature ranking from the SVM-RFE suggests that the HRV features from the relaxation and the last recovery phases are most relevant in classifying the MDD and control groups. Unlike the RRI, reduced HRV entropy due to mental stress did not recover, suggesting that entropy features may reflect prolonged sympathetic excitation in the recovery phase. This finding suggests that monitoring of HRV complexity changes when a subject is experiencing autonomic arousal and recovery can potentially allow higher-accuracy depressive symptom recognition. Future works can also examine patients with other medical conditions which elicit similar symptoms to those of the MDD, such as dementia [83].

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2017R1C1B5017730) and the Institute for Information and Communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (No. 2015-0-00062, The development of skin adhesive patches for the monitoring and prediction of mental disorders). This research was furthermore supported by the Original Technology Research Program for Brain

Science through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (No. NRF-2016M3C7A1947307; PI HJJ), and the Bio and Medical Technology Development Program of the NRF funded by the Korean government, MSIP (No. NRF-2017M3A9F1027323; PI HJJ).

Conflict of interest

The authors declare that they have no competing interests.

References

[1] Goldberger AL, Amaral LAN, Hausdorff JM, Ivanov PC, Peng C-K, Stanley HE. Fractal dynamics in physiology: Alterations with disease and aging. Proc Natl Acad Sci [Internet]. 2002; 99(Supplement 1): 2466–72. Available from: http://www.pnas.org/cgi/doi/10.1073/pnas.012579499.

[2] Costa M, Goldberger AL, Peng C. Multiscale entropy analysis of complex physiologic time series. Phys Rev Lett. 2002; 89(6): 6–9.

[3] Pincus SM. Approximate entropy as a measure of system complexity. Proc Natl Acad Sci USA [Internet]. 1991; 88(6): 2297–301. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11607165.

[4] Sunagawa K, Kawada T, Nakahara T. Dynamic nonlinear vago-sympathetic interaction in regulating heart rate. Heart Vessels. 1998; 13(4): 157–74.

[5] Malik M, Bigger T, Camm AJ, Kleiger RE, Malliani A, Moss AJ, et al. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Eur Heart J. 1996; 17: 354–81.

[6] Schulz S, Koschke M, Bär K-JJ, Voss A. The altered complexity of cardiovascular regulation in depressed patients. Physiol Meas [Internet]. 2010; 31(3): 303–21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20086275.

[7] Voss A, Kurths J, Kleiner HJ, Witt A, Wessel N, Saparin P, et al. The application of methods of non-linear dynamics for the improved and predictive recognition of patients threatened by sudden cardiac death. Vol. 31, Cardiovascular Research. 1996. 419–33.

[8] Rajendra Acharya U, Faust O, Adib Kadri N, Suri JS, Yu W. Automated identification of normal and diabetes heart rate signals using nonlinear measures. Comput Biol Med [Internet]. 2013; 43(10): 1523–9. Available from: http://dx.doi.org/10.1016/j.compbiomed.2013.05.024.

[9] Valenza G, Orsolini S, Diciotti S, Citi L, Scilingo EP, Guerrisi M, et al. Assessment of spontaneous cardiovascular oscillations in Parkinson's disease. Biomed Signal Process Control [Internet]. 2016; 26: 80–9. Available from: http://dx.doi.org/10.1016/j.bspc.2015.12.001.

[10] Al-Angari HMHM, Sahakian AV a. V. Use of sample entropy approach to study heart rate variability in obstructive sleep apnea syndrome. IEEE Trans Biomed Eng. 2007; 54(10): 1900–4.

[11] de la Torre-Luque A, Bornas X, Balle M, Fiol-Veny A. Complexity and nonlinear biomarkers in emotional disorders: A meta-analytic study. Neurosci Biobehav Rev [Internet]. 2016; 68: 410–22. Available from: http://dx.doi.org/10.1016/j.neubiorev.2016.05.023.

[12] Jacob A. Limitations of clinical psychiatric diagnostic measurements ayden. J Neurol Disord [Internet]. 2013; 1(122). Available from: http://www.esciencecentral.org/journals/2329-6895/2329-6895-1-122.digital/2329-6895-1-122.html.

[13] Setoyama D, Kato TA, Hashimoto R, Kunugi H, Hattori K, Hayakawa K, et al. Plasma metabolites predict severity of depression and suicidal ideation in psychiatric patients-a multicenter pilot analysis. PLoS One. 2016; 11(12): 1–16.

[14] Kim K, Kim SW, Myung W, Han CE, Fava M, Mischoulon D, et al. Reduced orbitofrontal-thalamic functional connectivity related to suicidal ideation in patients with major depressive disorder. Sci Rep [Internet]. 2017; 7(1): 1–11. Available from: http://dx.doi.org/10.1038/s41598-017-15926-0.

[15] Taylor CB, Conrad A, Wilhelm FH, Neri E, Delorenzo A, Kramer MA, et al. Psychophysiological and cortisol responses to psychological stress in depressed and nondepressed older men and women with elevated cardiovascular disease risk. Psychosom Med [Internet]. 2006; 68(4): 538–46. Available from: http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00006842-200607000-00005.

[16] Mundt JC, Vogel AP, Feltner DE, Lenderking WR. Vocal acoustic biomarkers of depression severity and treatment response. Biol Psychiatry [Internet]. 2012; 72(7): 580–7. Available from: http://dx.doi.org/10.1016/j.biopsych.2012.03.015.

[17] Khodayari-Rostamabad A, Reilly JP, Hasey GM, de Bruin H, MacCrimmon DJ. A machine learning approach using EEG data to predict response to SSRI treatment for major depressive disorder. Clin Neurophysiol [Internet]. 2013; 124(10): 1975–85. Available from: http://dx.doi.org/10.1016/j.clinph.2013.04.010.

[18] Jandl M, Steyer J, Kaschka WP. Suicide risk markers in major depressive disorder: a study of electrodermal activity and event-related potentials. J Affect Disord [Internet]. 2010; 123(1–3): 138–49. Available from: http://dx.doi.org/10.1016/j.jad.2009.09.011.
[19] Valenza G, Nardelli M, Bertschy G, Lanata A, Scilingo EP. Mood states modulate complexity in heartbeat dynamics: A multiscale entropy analysis. EPL [Internet]. 2014; 107(1): 18003. Available from: http://iopscience.iop.org/0295-5075/107/1/18003/article/.
[20] Valenza G, Citi L, Barbieri R. Instantaneous nonlinear assessment of complex cardiovascular dynamics by laguerre-volterra point process models. In: Proc Annu Int Conf IEEE Eng Med Biol Soc EMBS. 2013; pp. 6131–4.
[21] Mujica-Parodi LR, Yeragani V, Malaspina D. Nonlinear complexity and spectral analyses of heart rate variability in medicated and unmedicated patients with schizophrenia. Neuropsychobiology. 2005; 51(1): 10–5.
[22] Valenza G, Nardelli M, Lanata A, Gentili C, Bertschy G, Kosel M, et al. Predicting mood changes in bipolar disorder through heartbeat nonlinear dynamics. IEEE J Biomed Heal Informatics. 2016; 20(4): 1034–43.
[23] Reinertsen E, Osipov M, Liu CY, Kane JM, Petrides G, Clifford GD. Continuous assessment of schizophrenia using heart rate and accelerometer data. Physiol Meas. 2017; 38(7): 1456–71.
[24] Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. Am J Physiol Heart Circ Physiol. 2000; 278(6): H2039–49.
[25] Eckmann JP, Ruelle D. Fundamental limitations for estimating dimensions and Lyapunov exponents in dynamical systems. Phys D Nonlinear Phenom. 1992; 56(2–3): 185–7.
[26] Bolea J, Laguna P, Remartínez JM, Rovira E, Navarro A, Bailón R. Methodological Framework for Estimating the Correlation Dimension in HRV Signals. Comput Math Methods Med. 2014; 2014.
[27] World Health Organization. The Global Burden of Disease: 2004 update [Internet]. 2004 Update. 2008. Available from: http://www.who.int/healthinfo/global_burden_disease/2004_report_update/en/index.html.
[28] Luppa M, Heinrich S, Angermeyer MC, König H-H, Riedel-Heller SG. Cost-of-illness studies of depression. J Affect Disord [Internet]. 2007; 98(1–2): 29–43. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0165032706003302.
[29] Dhar AK, Barton DA. Depression and the link with cardiovascular disease. Front Psychiatry. 2016; 7(MAR): 1.
[30] Franklin JC, Ribeiro JD, Fox KR, Bentley KH, Kleiman EM, Jaroszewski ACAC, et al. Risk factors for suicidal thoughts and behaviors: a meta-analysis of 50 years of research. Psychol Bull [Internet]. 2016; (in press)(2): 187–232. Available from: http://doi.apa.org/getdoi.cfm?doi=10.1037/bul0000084.
[31] Zhang ZX, Tian XW, Lim JS. New algorithm for the depression diagnosis using HRV: A neuro-fuzzy approach. Proc 2011 Int Symp Bioelectron Bioinformatics, ISBB 2011. 2011; 283–6.
[32] Roh T, Hong S, Yoo HJ. Wearable depression monitoring system with heart-rate variability. Conf Proc. Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Annu Conf. 2014; 2014: 562–5.
[33] Matsui T, Kakisaka K, Shinba T. Impaired parasympathetic augmentation under relaxation in patients with depression as assessed by a novel non-contact microwave radar system. J Med Eng Technol. 2016; 40(1): 15–9.
[34] Kuang D, Yang R, Chen X, Lao G, Wu F, Huang X, et al. Depression recognition according to heart rate variability using Bayesian Networks. J Psychiatr Res [Internet]. 2017; 95: 282–7. Available from: https://doi.org/10.1016/j.jpsychires.2017.09.012.
[35] Sun G, Shinba T, Kirimoto T, Matsui T. An objective screening method for major depressive disorder using logistic regression analysis of heart rate variability data obtained in a mental task paradigm. Front Psychiatry. 2016; 7(NOV): 1–7.
[36] Chen W, Wang Z, Xie H, Yu W. Characterization of surface EMG signal based on fuzzy entropy. IEEE Trans Neural Syst Rehabil Eng. 2007; 15(1): 266–72.
[37] Porta A, Guzzetti S, Montano N, Furlan R, Pagani M, Malliani A, et al. Entropy, entropy rate, and pattern classification as tools to typify complexity in short heart period variability series. IEEE Trans Biomed Eng. 2001; 48(11): 1282–91.
[38] Vigo DE, Siri LN, Ladrón De Guevara MS, Martínez-Martínez JA, Fahrer RD, Cardinali DP, et al. Relation of depression to heart rate nonlinear dynamics in patients = 60 years of age with recent unstable angina pectoris or acute myocardial infarction. Am J Cardiol [Internet]. 2004; 93(6): 756–60. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15019886.
[39] Berger S, Kliem A, Yeragani V, Bär KJ. Cardio-respiratory coupling in untreated patients with major depression. J Affect Disord [Internet]. 2012; 139(2): 166–71. Available from: http://dx.doi.org/10.1016/j.jad.2012.01.035.
[40] Leistedt SJ, Linkowski P, Lanquart JP, Mietus JE, Davis RB, Goldberger AL, et al. Decreased neuroautonomic complexity in men during an acute major depressive episode: analysis of heart rate dynamics. Transl Psychiatry [Internet]. 2011; 1(7): e27. Available from: http://dx.doi.org/10.1038/tp.2011.23.
[41] Valenza G, Garcia RG, Citi L, Scilingo EP, Tomaz CA, Barbieri R. Nonlinear digital signal processing in mental health: Characterization of major depression using instantaneous entropy measures of heartbeat dynamics. Front Physiol. 2015; 6(MAR): 1–8.
[42] Tonhajzerova I, Ondrejka I, Javorka K, Turianikova Z, Farsky I, Javorka M. Cardiac autonomic regulation is impaired in

girls with major depression. Prog Neuro-Psychopharmacology Biol Psychiatry [Internet]. 2010; 34(4): 613–8. Available from: http://dx.doi.org/10.1016/j.pnpbp.2010.02.023.
[43] Voss A, Baier V, Schulz S, Bar KJ. Linear and nonlinear methods for analyses of cardiovascular variability in bipolar disorders. Bipolar Disord. 2006; 8(5 I): 441–52.
[44] Kibler JL, Ma M. Depressive Symptoms and Cardiovascular Reactivity to Laboratory Behavioral Stress. 2004; 11(2): 81–7.
[45] Carroll D, Phillips AC, Hunt K, Der G. Symptoms of depression and cardiovascular reactions to acute psychological stress: Evidence from a population study. Biol Psychol. 2007; 75(1): 68–74.
[46] Appelhans BM, Luecken LJ. Heart rate variability as an index of regulated emotional responding. Rev Gen Psychol. 2006; 10(3): 229–40.
[47] Nugent AC, Bain EE, Thayer JF, Sollers JJ, Drevets WC. Heart rate variability during motor and cognitive tasks in females with major depressive disorder. Psychiatry Res – Neuroimaging [Internet]. 2011; 191(1): 1–8. Available from: http://dx.doi.org/10.1016/j.pscychresns.2010.08.013.
[48] Shinba T. Altered autonomic activity and reactivity in depression revealed by heart-rate variability measurement during rest and task conditions. Psychiatry Clin Neurosci. 2014; 68(3): 225–33.
[49] Salomon K, Clift A, Rottenberg J. Major Depressive Disorder Is Associated With Attenuated Cardiovascular Reactivity and Impaired Recovery Among Those Free of Cardiovascular Disease. 2009; 28(2): 157–65.
[50] Kim AY, Jang EH, Kim S, Choi KW, Jeon HJ, Yu HY, et al. Automatic detection of major depressive disorder using electrodermal activity. Sci Rep [Internet]. 2018; 8(1): 17030. Available from: http://www.nature.com/articles/s41598-018-35147-3.
[51] Hamilton MAX. Development of a rating scale for primary depressive illness. Br J Clin Psychol. 1967; 6(4): 278–96.
[52] Zarjam P, Epps J, Chen F, Lovell NH. Estimating cognitive workload using wavelet entropy-based features during an arithmetic task. Comput Biol Med [Internet]. 2013; 43(12): 2186–1295. Available from: http://dx.doi.org/10.1016/j.compbiomed.2013.08.021.
[53] Tarvainen MP, Niskanen JP, Lipponen JA, Ranta-aho PO, Karjalainen PA. Kubios HRV – heart rate variability analysis software. Comput Methods Programs Biomed [Internet]. 2014; 113(1): 210–20. Available from: http://dx.doi.org/10.1016/j.cmpb.2013.07.024.
[54] Pan J, Tompkins WJ. A real-time QRS detection algorithm. IEEE Trans Biomed Eng [Internet]. 1985; BME-32(3): 230–6. Available from: http://ieeexplore.ieee.org/document/4122029/.
[55] Yentes JM, Hunt N, Schmid KK, Kaipust JP, McGrath D, Stergiou N. The Appropriate Use of Approximate Entropy and Sample Entropy with Short Data Sets. 2013; 41(2): 349–65.
[56] Chen W, Zhuang J, Yu W, Wang Z. Measuring complexity using FuzzyEn, ApEn, and SampEn. Med Eng Phys. 2009; 31(1): 61–8.
[57] Porta A, Baselli G, Liberati D, Montano N, Cogliati C, Gnecchi-Ruscone T, et al. Measuring regularity by means of a corrected conditional entropy in sympathetic outflow. Biol Cybern. 1998; 78(1): 71–8.
[58] Viola AU, Tobaldini E, Chellappa SL, Casali KR, Porta A, Montano N. Short-term complexity of cardiac autonomic control during sleep: REM as a potential risk factor for cardiovascular system in aging. PLoS One. 2011; 6(4).
[59] de Oliveira EA, da Silva AKF, Christofaro DGD, Vanzella LM, Gomes RL, Vanderlei FM, et al. Influence of Type 1 Diabetes on the Symbolic Analysis and Complexity of Heart Rate Variability in Young Adults. Arq Bras Cardiol [Internet]. 2018; 1–11. Available from: http://www.gnresearch.org/doi/10.5935/abc.20180117.
[60] Guyon I. Gene Selection for Cancer Classification. 2002; 389–422.
[61] Fan R, Chen P, Lin C. Working Set Selection Using Second Order Information for Training Support Vector Machines. 2005; 6: 1889–918.
[62] Yan K, Zhang D. Feature selection and analysis on correlated gas sensor data with recursive feature elimination. Sensors Actuators, B Chem [Internet]. 2015; 212: 353–63. Available from: http://dx.doi.org/10.1016/j.snb.2015.02.025.
[63] Fukunaga K. Introduction to Statistical Pattern Recognition. San Diego, CA, USA: Academic Press; 1990.
[64] Müller AC, Guido S. Introduction to Machine Learning with Python: A Guide for Data Scientists [Internet]. O'Reilly Media, Incorporated; 2017. Available from: https://books.google.co.kr/books?id=q5pnAQAACAAJ.
[65] Vapnik VN. Statistical learning theory. Adapt Learn Syst Signal Process Commun Control [Internet]. 1998; 2: 1–740. Available from: http://eu.wiley.com/WileyCDA/WileyTitle/productCd-0471030031.html.
[66] Vapnik VN. The Nature of Statistical Learning Theory. Springer; 2000.
[67] Lotte F, Congedo M, Lécuyer A, Lamarche F, Arnaldi B. A review of classification algorithms for EEG-based brain-computer interfaces. J Neural Eng. 2007; 4(2).
[68] Ravelo-García AG, Kraemer JF, Navarro-Mesa JL, Hernández-Pérez E, Navarro-Esteva J, Juliá-Serdá G, et al. Oxygen saturation and RR intervals feature selection for sleep apnea detection. Entropy. 2015; 17(5): 2932–57.
[69] Moisen GG. Classification and regression trees. Encycl Ecol. 2008; (2000): 582–8.
[70] Noguchi K, Gel YR, Brunner E, Konietschke F. nparLD: An R software package for the nonparametric analysis of longitudinal data in factorial experiments. J Stat Softw. 2012; 50(12): 1–23.

[71] Brunner E, Puri ML. Nonparametric methods in factorial designs. Stat Pap. 2001; 42(1): 1–52.
[72] Moon E, Lee S-H, Kim D-H, Hwang B. Comparative study of heart rate variability in patients with schizophrenia, bipolar disorder, post-traumatic stress disorder, or major depressive disorder. Clin Psychopharmacol Neurosci [Internet]. 2013; 11(3): 137–43. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3897762&tool=pmcentrez&rendertype=abstract.
[73] Chang JS, Yoo CS, Yi SH, Her JY, Choi HM, Ha TH, et al. An integrative assessment of the psychophysiologic alterations in young women with recurrent major depressive disorder. Psychosom Med. 2012; 74(5): 495–500.
[74] Baumert M, Lambert GW, Dawood T, Lambert EA, Esler MD, McGrane M, et al. Short-term heart rate variability and cardiac norepinephrine spillover in patients with depression and panic disorder. Am J Physiol Heart Circ Physiol. 2009; 297(2): H674–9.
[75] Jentsch MC, Van Buel EM, Bosker FJ, Gladkevich AV, Klein HC, Oude Voshaar RC, et al. Biomarker approaches in major depressive disorder evaluated in the context of current hypotheses. Biomark Med. 2015; 9(3): 277–97.
[76] 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 [Internet]. 2011; 10(1): 96. Available from: http://www.biomedical-engineering-online.com/content/10/1/96.
[77] Schubert C, Lambertz M, Nelesen RA, Bardwell W, Choi JB, Dimsdale JE. Effects of stress on heart rate complexity-A comparison between short-term and chronic stress. Biol Psychol. 2009; 80(3): 325–32.
[78] Visnovcova Z, Mestanik M, Javorka M, Mokra D, Gala M, Jurko A, et al. Complexity and time asymmetry of heart rate variability are altered in acute mental stress. Physiol Meas. 2014; 35(7): 1319–34.
[79] Schumann A, Andrack C, Bär KJ. Differences of sympathetic and parasympathetic modulation in major depression. Prog Neuro-Psychopharmacology Biol Psychiatry [Internet]. 2017; 79: 324–31. Available from: http://dx.doi.org/10.1016/j.pnpbp.2017.07.009.
[80] Kemp AH, Quintana DS, Gray MA, Felmingham KL, Brown K, Gatt JM. Impact of depression and antidepressant treatment on heart rate variability: a review and meta-analysis. Biol Psychiatry [Internet]. 2010; 67(11): 1067–74. Available from: http://dx.doi.org/10.1016/j.biopsych.2009.12.012.
[81] Licht CMM, De Geus EJC, Van Dyck R, Penninx BWJH. Longitudinal evidence for unfavorable effects of antidepressants on heart rate variability. Biol Psychiatry [Internet]. 2010; 68(9): 861–8. Available from: http://dx.doi.org/10.1016/j.biopsych.2010.06.032.
[82] Kemp AH, Quintana DS, Malhi GS. Effects of serotonin reuptake inhibitors on heart rate variability: methodological issues, medical comorbidity, and clinical relevance. Biol Psychiatry [Internet]. 2011; 69(8): e25–6. Available from: http://dx.doi.org/10.1016/j.biopsych.2010.10.035.
[83] Kitching D. Depression in dementia. Aust Prescr. 2015; 38(6): 209–11.