Stress detecting wearable system-Towards a methodology for stress recognition through the physiological analysis

Giorgia Acerbi[1†], Erika Rovini[1†], Stefano Betti[1†], Antonio Tirri[2], Judit Flóra Rónai[2], Antonella Sirianni[2], Jacopo Agrimi[3], Lorenzo Eusebi[4], Filippo Cavallo[1]

Corresponding Author: The corresponding author will receive all communication regarding the contribution
Full name: Giorgia Acerbi
Address of institution: The BioRobotics Institute, Scuola Superiore Sant'Anna; Viale Rinaldo Piaggio, 34, 56025 Pontedera (PI), Italy
Email: g.acerbi@sssup.it
Phone Number: +39 0587 672152

Well-being and active ageing

Keywords: stress detection, heart rate variability, galvanic skin response, psychometric instruments

Acknowledgments: This work was supported by research funding provided by Trans.Safe (AmbienT Response to Avoid Negative Stress and enhance SAFEty, www.transsafe.eu) project - 6th call of the Ambient Assisted Living Joint Programme (AAL JP) with the topic "ICT-based Solutions for Supporting Occupation in Life of Older Adults".

---

[1] Giorgia Acerbi; Erika Rovini; Stefano Betti; Filippo Cavallo;

The BioRobotics Institute, Scuola Superiore Sant'Anna; Viale Rinaldo Piaggio, 34, 56025 Pontedera (PI), Italy; e-mail: g.acerbi@sssup.it, e.rovini@sssup.it; s.betti@sssup.it; f.cavallo@sssup.it

[2] Antonio Tirri; Judit Flóra Rónai; Antonella Sirianni;

Telecom Italia | TIM WHITE Joint Open Lab; Via Cardinale Pietro Maffi, 27 Pisa, Italy; e-mail: antonio.tirri@telecomitalia.it; juditflora.ronai@telecomitalia.it

[3] Jacopo Agrimi;

Life Science Institute, Scuola Superiore Sant'Anna; Piazza Martiri della Libertà, 33, 56127 Pisa (PI), Italy; e-mail: j.agrimi@sssup.it

[4] Lorenzo Eusebi;

Telecom Italia | TIM S-Cube Joint Open Lab; Via Camillo Golgi, 42 Milano, Italy; e-mail: lorenzo.eusebi@telecomitalia.it

[†] Equal contribution to the work

1 Introduction

Stress is a physiological response to the mental, emotional, or physical challenge and it can be defined as the reaction of a person to the environmental requests or influences (Sun et al., 2010). Stress conditions can cause physical and emotional exhaustion that leads to symptoms such as headaches, stomach complaints and difficulties in sleeping. In the last years the impact of stress on the society has been increased. A study conducted by the American Institute of Stress (Statistic Brain Research Institute, NY) has shown as in 2015 the 48% of people feels that their stress condition has increased over the past five years. 77% of people regularly experiences physical symptoms caused by stress with a negative impact on their personal and professional life (Statistic Brain, 2015). The influence of stress and its consequences on society concerns also the economic aspect. According to the recent EU-funded project 2013, the cost to Europe of work-related stress and depression was estimated to be €617 billion annually. The total amount includes loss of productivity, health care costs and social welfare costs (EU-OSHA, 2016). The early detection of stress can positively affect personal wellbeing and society affluence.

Traditionally, the level of personal stress has been established using some psychometric instruments and scales (Ulstein et al., 2007), which are subjective. Subsequently the correlation between the variation of the physiological signals and stress was investigated in order to make the measurement more objective.

1.1 Physiological Signals and Stress Concept

Physiological phenomena are extremely correlated with stress and anxiety, such as heart rate variability and galvanic skin response. Human stress response can be described through Psychoneuroimmunology that tries to link together the physiological systems involved in the stress response: the nervous system, the endocrine system and the immune system (Seaward, 1999).

Several studies have shown that stress has an impact on the Autonomic Nervous System (ANS) (Watkins et al., 1999). The ANS provides a rapidly responding mechanism to control a wide range of functions and organs, including heart, skin resistance, digestive tract, lungs, bladder and blood vessels (Tsigos & Chrousos, 1994). The ANS has two components, the sympathetic nervous system (SNS) and the parasympathetic nervous system (PNS). In particular, the response "fight-or-flight" is associated with SNS, through the release of adrenaline and noradrenaline (Seaward, 1999), while PNS is involved in relaxation process. Stress response is structured into 3 main stages: immediate effects of stress involve the SNS, with releasing of adrenaline and noradrenaline in 2-3 seconds; intermediate effects are characterized by 20-30 seconds time activity, in which adrenal medulla releases epinephrine and norepinephrine.

That is why alteration of physiological signals and variables can be related to a change of stress condition such as cardiac activity (de Santos Sierra et al., 2011), electrodermal activity (EDA) (Park, 2009; Haapalainen et al., 2010), electro-myographic activity (Lundberg et al., 1994), breathing (Rottenberg, 2002), skin temperature (Karthikeyan et al., 2012), electrical brain activity (Lim & Chia, 2015), eye blink (Haak et al., 2009). In particular SNS and PNS regulate the EDA, the heart rate variability (HRV) and the brain waves that are commonly used in literature to investigate the levels of stress during different tasks (Sharma & Gedeon, 2012).

1.1.1 Electrodermal activity

Psycho physiological measures have been recently used in HRI studies, in which, in addition to HRV, Galvanic Skin Response (GSR) has been used. The neural mechanism and pathways involved in the central control of electrodermal activity are numerous and complex. EDA is related to the level of arousal elicited by an extended range of psychological and emotional states with either positive or negative valence. Different studies investigating anxiety, anger, fear and also joy experiences report increased EDA (Ritz et al., 2000; Stemmler et al., 2001). It is also an indicator of the cognitive load, stress and arousal (Park, 2009; Haapalainen et al., 2010), because of the variation of the skin electrical resistance in response to various emotional stimuli. When a subject is under mental stress, sweat gland activity is activated and increases skin conductance (SC). Since the sweat glands are also controlled by the SNS, SC acts as an indicator for sympathetic activation due to the stress reaction (Sun et al., 2010).

GSR has already been used in previous works in combination with other physiological parameters. For example, SC has been combined with electro cardiac activity, electromyographic activity and respiration activity in order to monitor drivers' behaviours through open roads (Haley & Picard, 2005). In particular, the parameters provided were the number of stressors in a given temporal window, the sum of the amplitude of all the stressors counted in that temporal window, the sum of the response durations and the sum of the areas under the peaks counted as stressors. SC has been also used together with speech activity (Kurniawan et al., 2013), measuring mean, minimum and standard deviation of the GSR signal, the sum of all the amplitudes of the peaks and the rising time in order to distinguish different stress levels. Finally, the integration of GSR, HRV and accelerometer data has been implemented in the work of Sun et al. (2010), with the aim to differentiate between physical activity and mental stress. In particular, electrodermal activity has been analysed through three main parameters: the number of the stressor, the related amplitude and the sum of the duration of the responses.

1.1.2 Electro cardiac activity

There are two types of neuro-modulatory receptors in cardiac cells: one is for acetylcholine (SNS) and the other is norepinephrine (PNS). These receptors interact with inhibitory or excitatory proteins, which, through chemical exchanges, can modify the Calcium concentration in the heart cells membrane and inhibit or stimulate heart rate (HR) and the strength of contraction (Clifford, 2002). HR describes the cardiac activity when the ANS attempts to tackle with the human body demands depending on the stimuli received. Concretely, ANS reacts against a stressing stimulus provoking an increase in blood volume within the veins, so rest of the body can react properly, increasing the number of heartbeats (de Santos Sierra et al., 2011). In confirmation of this aspects, over recent years clinical researches have shown that one of the most important indicators of stress is HRV. It is the variation in the time interval between one heartbeat and the next one. To study the effect of SNS and PNS activities, starting from ECG signal, it is necessary to analyse the HRV signal both in time and frequency domains. Generally cardiac parameters as mean of Inter-Beat Interval (IBI), HR, signal power in low frequency (LF) and high frequency (HF) bands are used to analyse stress. The HRV analysis has already been used in different studies to detect stress in various condition as mental task (Taelman et al., 2009), high workload (Orsila et al., 2008) , car driving (Healey & Picard, 2005) and other common daily tasks.

1.2 The aim of the study

This paper presents an experimental methodology to collect and analyse physiological data to detect the stress status of the user. The methodology has been applied in a test for the Trans.Safe (The AmbienT Response to Avoid Negative Stress and enhance SAFEty) European research project which has the aim to detect stress levels, through the monitoring and interpretation of physiological signals.

EDA and HRV were the physiological signals measured during the tests since they are two of the most important indicators of stress (see par. 1.1.1 and 1.1.2) and they can be revealed through portable and non-invasive devices. Thus, the stress detection activity carried out in this experimentation has been performed through a combination of two wearable sensors, Shimmer GSR Sensor and Zephyr BioHarness™.

A new experimental protocol for the collection of physiological data in different conditions has been defined. It consisted of alternated stages of rest and stress induction phases combined with the administration of psychometric instruments. Subsequently signal processing techniques and machine learning algorithms for classification have been implemented. Then, the data collected was properly processed and analysed in order to build a classifier of the user's stress status.

Since the main goal of this study is the detection and monitoring of stress with a satisfactory accuracy level using a ultra-low invasiveness system, it is reasonable to think that improving the comfort, the user could wear the system for a long time, both during work or daily activities.

In the future a such system could act as a portable system control, like medical devices as cardiac holter or pressure monitoring devices (24 hours). Furthermore it could also be useful for the user in order to predict the rise of stress and act to reduce it. The feedback for stress presence and any suggestion or intervention to decrease it would allow benefits for the user's health and a reduction in health care costs for stress-related illnesses.

2 Materials and Methods

In this section the sensor devices used for the acquisition of physiological signals, the experimental protocol developed and adopted and the methodology chosen for data analysis are described in detail.

2.1 Instrumentation

The choice of the wearable sensor devices to be included into the test has been performed according to two criteria: accuracy of measurements and unobtrusiveness of the sensors. There are several devices on the market that claim the measurement of cardiac and electro-dermal activity in a unobtrusive way. Unfortunately, not all these devices are accurate enough for a reliable assessment of stress conditions. In order to find a reasonable trade-off, we selected two devices: Zephyr BioHarness™3 and Shimmer GSR Sensor (Fig. 1).

Figure 1. Zephyr BioHarness™ 3on the left and Shimmer GSR Sensor on the right

Zephyr BioHarness™3 (BH3) (Medtronic, 2015) is a Bluetooth chest belt capable of retrieving signals derived from the ECG such Heart Rate and R-R Intervals. The ECG signal is sampled at 250 Hz. Moreover, the BH3 is able to collect other signals such as breathing rate, posture information and skin temperature. For the data analysis and the development of the stress detection algorithm the Inter-Beat-Interval data provided by the device has been used. The GSR Module developed by Shimmer (Shimmer, 2016) is a wearable sensor composed by two special finger electrodes and a main unit that streams data related to the galvanic skin response with a sample frequency of 51.2 Hz using a Bluetooth connection.

2.2 Participants

Twelve voluntary students (3 men, 9 women) with a mean age of 26.0 years old (SD= 4.8 years, range = 21-30 years old) participated on purpose in this study. All the participants did not meet the exclusion criteria that consisted in neurological disorders that made unable the subjects to complete the mental tasks proposed or cardiac diseases that could deface the physiological response in electro cardiac activity.

Participants completed the experimental session in the Scuola Superiore Sant'Anna (Pisa, Italy) and in the Telecom Italia WHITE Joint Open Lab (Pisa, Italy). Written informed consent was obtained from all the participants before starting the tests.

2.3 Experimental Protocol

The experimental protocol was intended to put the subjects in a state of emotional and cognitive stress, in order to measure the variations of their physiological parameters induced by stress.

The experimentation consisted in three phases: a baseline, a stress induction and a recovery stage. During baseline the subjects relaxed in a separate room, for 10 minutes, without using mobile phone, without music or external sounds, without stimuli and without closing their eyes. This phase was indispensable in order to acquire the personal baseline of each subject, since physiological parameters show a wide inter-subjects variability. At the end of baseline recording, the psychologist administered psychometric instruments to the participants to obtain a subjective perception about the level of stress, anxiety and drowsiness. Then the subjects performed the stress phase, during about 15-20 minutes, completing a series of extremely demanding cognitive tests handed out by the psychologist in order to induce the stress. People were not aware that this phase was part of the experiment: the psychologist indeed pretended to be sent by University to detect the intelligence quotient (IQ) for a poll. The investigator assumed a very aggressive behaviour towards the subject, behaving rude and correcting the person even when the he accomplished the task properly. Furthermore, the user performed the required tasks by listening a noisy sound in background that simulates high intensity traffic jam. At the end of this phase, the subjects filled out the psychometric instruments again. Afterwards a recovery period of 10 minutes was performed, in the same conditions as in the baseline phase.

During the whole experimental session (baseline, stress and recovery phases), the tested subjects wore the kit of wearable sensors described in par.2.1, in order to record electro cardiac and electrodermal activities.

2.3.1 Tests for stress induction

The aim of this experimental protocol was to arouse stress in tested subjects that would produce major changes in the level of physiological signals. For this reason, in the experimental protocol the stress induction phase consisted of two paths: (i) the use of validated neuropsychological tests that caused a great cognitive effort; (ii) the creation of a stressful social situation that would put the subject under pressure causing a strong emotional reaction.

The five different following tasks (Fig. 2), were executed by the tested subjects:

- **Digit Span:** it is a common measure of short-term memory to evaluate working memory's number storage capacity. In the test of Reverse Digit Span a list of random numbers was read out loud to the person who had to immediately repeat it in a backward order. The test began with two or three numbers and if the subject correctly repeated them, longer lists are given, increasing until the person commits errors.
- **Stroop Color Test:** it is a common test to measure selective and divided attention, cognitive flexibility and processing speed (Lansbergen et al., 2007). This test is a demonstration of interference in the reaction time of a task in which the subject was asked to read out loud and as fast as possible either the written word or the ink col-

or. Three different experimental protocols have been used to conduct this test: (i) the names of colors appeared in black ink and the user had to read the written word (names of color); (ii) block circles of different colors were presented and the subjects had to say the names of colors; (iii) the written words (names of colors) were printed in a different ink than the color named and the participants had to say the ink color. When the name of a colour is printed in a colour not denoted by the name, naming the colour of that word requires more time and the subject is more prone to errors.

- **Corsi Reverse**: The Corsi block-tapping test is a psychological test that assesses visuo-spatial short term memory. The experiment is done typically by using a wooden base where nine identical spatially separated blocks are present. In the Reverse Corsi Test (Gillet, 2007) the experimenter indicated a sequence of blocks by tapping them and the subject was requested to reproduce the spatial succession of boxes in the reverse way. The sequence started out simple, usually using two or three blocks, increasing until the subject commits errors.
- **Kohs Block Design Test:** this is a performance test designed to be an IQ test and to measure visual-spatial skills (Barbeau, 1980). The subject was asked to replicate the patterns displayed on a series of test cards by using colored cubes (each side has a single color or two colors divided by a diagonal line). The first test cards represented simple patterns (4 blocks) and increased in complexity as the test progress (16 blocks).
- **Tower of Hanoi**: this is a mathematical game, common to test problem solving and executive capacity of the subject (Miyake et al., 2000). It is composed by three rods and a number of disks of different sizes which can slide onto any rod. The subject had to move the entire stack to another rod, following simple rules: only one disk could be moved at a time; only the upper disk from one of the stacks could be moved and placed on top of another stack; no disk could be placed on top of a smaller disk. The minimum number of moves required to solve the game is ($2n–1$), where n is the number of disks.

Figure 2. Stress induction test set administered during the experimental session

2.3.2 Psychometric Instruments

In order to measure the emotional state and the level of stress of the subjects, the following psychometric instruments were administered before and after stress induction phase:

- **State-Trait Anxiety Inventory (STAI)**: this scale is one of the most frequently, reliable and sensitive used measures of anxiety in applied psychology research. In this study the short-form of the STAI scale was used, consisting of only six items (STAI-6) since the objective was to establish the level of stress and anxiety produced during the stress phase (Marteau et al., 1992). Higher STAI scores suggest higher levels of anxiety.
- **Karolinska Sleepiness Scale (KSS):** it is one of the most common sleepiness state tests and it is a 9-point Likert scale based on a self-reported assessment of the person's level of drowsiness at the moment (Åkerstedt & Gillberg, 1990). The subject had to choose his level of sleepiness from 1="very alert" to 9="very sleepy". KSS was originally developed to constitute a one-dimensional scale of sleepiness and was validated against alpha and theta electroencephalographic activity (Kaida et al., 2006).
- **Shortened State Stress Questionnaire (SSSQ):** The 24-item SSSQ (Helton, 2004), based on the 90 Question Dundee Stress State Questionnaire (DSSQ), provides a rapid, reliable, self-report assessment of the three primary stress dimensions: distress, task engagement and worry (Pfaff et al., 2012).

2.3 Data Analysis

The physiological data acquired during the whole experimentation have been offline analysed using Matlab® R2012a.

The acquired data have been examined for baseline phase (10 minutes of recording), stress phase (ranging from 15 to 20 minutes, depending from the attitude and behaviour of the tested subject) and recovery phase (10 minutes). Thus, for each phase, we obtained a dataset composed by a set of GSR features for each participant and another dataset consisting of features extracted from HRV signal. All these data were analysed in order to investigate variations in physiological parameters that could be attributed to stress statutes of the tested subjects.

2.3.1 Galvanic Skin Response (GSR)

The EDA has been recorded using Shimmer GSR sensor which provides as output the galvanic resistance, that has been converted into galvanic skin conductance. In the features extraction algorithm, the signal has been analyzed with temporal windows of 2 minutes, after a filtering process, using a moving average filter. The features extraction algorithm is based on startle detection that can lead to a set of computable features. To detect startles, the derivative has been calculated in order to consider the variation in the amplitude of GSR signal, and a threshold has been applied (TH_{GSR} = 0.005 µS has found to be adequate for the 12 subjects). Once the response was detected, the zero-crossing of the derivative preceding and following the response were identified as the onset and end of the startle (Haley & Picard, 2000). Starting from the startle detection, the following parameters have been calculated (Table 1):

Table 1. Features extracted from GSR signal calculated within a temporal window.

Feature Name	**Description**
Num_Startle	Number of the stressors
Sum_Amplitude	Sum of the amplitude of the stressors
Sum_RiseTime	Sum of the rise duration of the stressors
Sum_RecTime	Sum of the decrease duration of the stressors
Rise_Rate	Mean value of the rise duration of the stressors
Decay_Rate	Mean value of the decrease duration of the stressors
Area_GSR	Mean of the area under each stressor
Mean_GSR	Mean value of GSR signal
Std_GSR	Standard deviation of GSR signal

Some of the computed GSR features are shown, on a typical electrodermal signal, in figure 3.

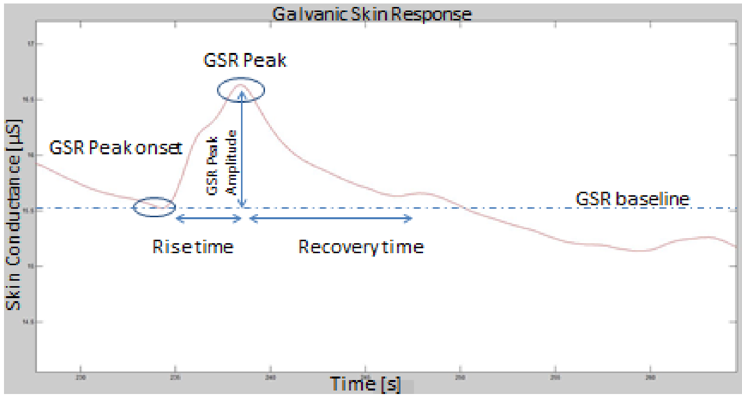

Figure 3: Typical EDA signal, with some of the computed features

2.3.2 Electro cardiac activity

Electro cardiac activity has been recorded using the chest belt Zephyr BioHarness™ BH3. The device provides as output the raw ECG signal and the HRV data that specifies the temporal distance between a beat and the following one. Starting from Inter-Beat-Interval (IBI), the algorithm to extract the main features has been developed. The IBI signal has been modified identifying and correcting ectopic rhythm, which is an irregular heart rhythm due to a premature heartbeat. The analysis of cardiac signal has been structured investigating both the time domain and the frequency domain. Regarding the time domain, the following parameters have been selected and computed (see Table 2)

Table 2. Features extracted from HRV signal in the temporal domain.

Feature Name	Description
IBI_mean	Mean of Inter-Beat-Interval corresponding to R-to-R interval
SDNN	Standard deviation of all Normal RR intervals (NN intervals)
HR_mean	Mean of Heart Rate
SDHR	Standard deviation of the Heart Rate
RMSSD	Square root of the mean of the squared differences between adjacent normal RR intervals
pNN50	Percentage of differences between adjacent normal RR intervals exceeding 50 ms
#ECT	Number of ectopic intervals (abnormal RR intervals)

%ECT	Percentage of ectopic intervals on the total number of RR intervals

By identifying and correcting ectopic rhythm, a Normal-to-Normal (NN) interval sequence appropriate for HRV analysis is obtained. Since the NN interval sequence is an irregularly sampled time sequence, for spectral analysis it had to be therefore converted to an equidistantly sampled sequence (Mali et al., 2014). After a smoothing of the signal, the NN interval sequence has been resampled at 4Hz. For the analysis in frequency domain, the following parameters have been computed (see Table 3):

Table 3. Features extracted from HRV signal in the frequency domain.

Feature Name	Description
Peak VLF	Frequency peak in very low frequency (VLF) range (0.04–0.15 Hz)
Area VLF	Signal power by Power Spectral Density (PSD) in VLF
%VLF	Percentage of signal power in the VLF respect to the total signal power
Peak LF	Frequency peak in low frequency (LF) range (0.04–0.15 Hz)
Area LF	Signal power by PSD in LF
%LF	Percentage of signal power in the LF respect to the total signal power
Peak HF	Frequency peak in high frequency (HF) range (0.15–0.4 Hz)
Area HF	Signal power by PSD in HF
%HF	Percentage of signal power in the HF respect to the total signal power
LF/HF	Ratio between LF and HF powers

2.3.3 Data processing and Statistical Analysis

After extracting features from physiological signals, Kolmogorov-Smirnov test was applied in order to verify the normal distribution of data. A non-parametric statistical analysis was used because the test showed data were not normally distributed. Then, Kruskal-Wallis (KW) test was used for comparing data acquired in baseline phase and those recorded during stress phase in order to verify a significant difference (p-value<0.05) on the basis of the extracted parameters. Furthermore the linear correlation between the significant parameters was calculated using the Pearson's coefficient. If the value of correlation between two features was at least rho=0.8, the less significant one was deleted. Then, the remaining features were used for Principal Component Analysis (PCA) in order to identify how the groups investigated, related to the three different phases of the experimental protocol, could be visualized and separated in the space of the principal components (PCs). Finally, the most important PCs, that included more than 80% of the overall variance of data, were taken into account in order to train and test a Support Vector Machine (SVM) classifier which had to be able to correctly classify a subject as stressed or not-stressed.

Regarding the analysis of the psychometric instruments, a T-test has been conducted in order to assess if significant differences between after and before the stress induction phase could be revealed.

Finally, a linear regression analysis has been implemented with the aim to look for a correlation between the results obtained by the psychometric instruments administered and the physiological parameters measured.

3 Results and Discussion

In this section the results obtained from both the analysis of physiological data and the psychometric instruments are reported and widely discussed, examining the most important features extracted, the evaluation of the psychometric instruments and the algorithm for data classification.

3.1 Physiological Parameters Assessment

Features extracted by physiological parameters are reported in Tables 4-5 both for baseline phase and stress phase as mean values and standard deviations. Furthermore p-values, calculated with KW test for non parametric data, are also disclosed because they represent if there are significant differences between the two investigated groups.

Table 4. Features extracted from GSR signal: mean values ± standard deviations and significance.

Parameters	Baseline			Stress			p-value	
Num_Startle (#)	15.97	±	4.71	17.95	±	3.06	0.119	
Sum_amplitude (µS)	11.18	±	12.40	11.18	±	6.79	0.453	
Sum_RiseTime (s)	32.25	±	6.94	41.25	±	5.61	0.004	*
Sum_RecTime (s)	61.96	±	8.73	64.81	±	2.52	0.488	
Rise_Rate (µS/s)	3.61	±	0.90	3.08	±	0.39	0.141	
Decay_Rate (µS/s)	9.36	±	4.32	5.59	±	1.21	0.003	*
Area_GSR (s·µS)	2.55	±	1.95	2.11	±	1.16	0.773	
Mean_GSR (µS)	11.56	±	5.34	16.83	±	5.99	0.028	*
Std_GSR (µS)	1.62	±	1.33	1.30	±	0.73	0.862	

* Significant difference between groups ($p<0.05$)

Table 5. Features extracted from HRV signal: mean values ± standard deviations and significance.

Parameters	Baseline			Stress			p-value	
IBI_mean (s)	0.788	±	0.126	0.642	±	0.096	0.005	*
SDNN (s)	0.065	±	0.021	0.071	±	0.025	0.544	
HR_mean (bpm)	78.45	±	12.38	95.54	±	13.69	0.005	*
SDHR (bpm)	6.43	±	1.15	10.48	±	3.88	0.001	*
RMSSD (s)	0.04	±	0.02	0.03	±	0.01	0.018	*
pNN50 (%)	22.89	±	19.44	7.35	±	4.98	0.043	*
#ECT (#)	11.42	±	14.57	25.33	±	24.18	0.182	
%ECT (%)	1.30	±	1.77	1.29	±	1.31	0.885	
Peak VLF (Hz)	0.033	±	0.014	0.011	±	0.017	0.005	*
Area VLF (s^2)	195.03	±	253.25	185.88	±	106.16	0.326	
%VLF (%)	16.93	±	9.87	28.85	±	9.41	0.009	*
Peak LF (Hz)	0.077	±	0.019	0.053	±	0.022	0.016	*
Area LF (s^2)	509.75	±	364.71	317.93	±	147.23	0.184	
%LF (%)	55.07	±	16.94	51.15	±	9.45	0.194	
Peak HF (Hz)	0.219	±	0.078	0.203	±	0.096	0.486	
Area HF (s^2)	284.78	±	276.02	120.66	±	59.22	0.106	
%HF (%)	27.99	±	20.94	20.01	±	8.88	0.525	
LF/HF	3.53	±	2.79	3.27	±	2.28	0.908	

* Significant difference between groups ($p<0.05$)

Significant differences are observed in some parameters, both for features extracted by electro dermal and electro cardiac activities, representing a concrete variation in physiological response to a psychological stress induction.

In particular for the first signal, Sum_RiseTime, Decay_Rate and Mean_GSR are the significant parameters. For the second signal IBI_mean, HR_mean, SDHR, RMSSD and pNN50 are the significant features in the temporal domain, whereas Peak VLF, %VLF and Peak LF are the ones in the frequency domain.

Discussing significant parameters derived by electrodermal activity, Sum_RiseTime is a parameter that gives an indication of how the global GSR level is varying as time progresses. If the sympathetic branch of the ANS is highly aroused, then sweat gland activity also increases. This fact leads to an increase of skin conductance, that can be then a measure of emotional and sympathetic responses. A significant variation of this parameter from baseline to stress phase can be explained as an increase of arousal level of the subject, probably due to an increment of stress level during the execution of the stressor tasks. A significant variation has been observed from baseline phase to stress phase for other two parameters: Decay_Rate and Mean_GSR. Regarding the mean value of GSR, it reflects the variation of the signals in terms of arousal, cognitive load and stress in general. So, an increase of cognitive load corresponds to an increase of the mean value of the signal, related to a bigger sweat gland activity that modifies SC. Finally a considerable variation in decay rate, which represents an indirect measure of the relaxation pattern experienced by the subject (Singh et al., 2012) could mean that when the arousal level is high, the GSR needs more time to assume values similar to baseline ones. So it is reasonable to have a variation of the time needed to obtain a relaxation, during a stress phase, respect to the baseline.

Regarding electro cardiac activity variations in the mean values of IBI and HR from baseline to stress phase are absolutely congruent with an increase of stress level: the number of beats in a minute increases, with a related reduction of the time between a heartbeat and the following one. According to Orsila et al. (2008) in which RMSSD parameter changed its values among different phases of the experimental session described, this parameter presents a variation from baseline to stress phase. The lower value in the stress stage may suggests the subjects' perceived stress was effectively higher during this phase of the protocol. The difference between baseline and stress conditions in pNN50 was expected, as in (Taelman et al, 2009). It is probably due to the short term variability, which is lower with a cognitive task than during rest. Also SDHR changes between the phases, being a measure for long term variability. Analysing frequency domain parameters, it is known that sympathetic and parasympathetic activities are reflected into LF and HF power, so a variation in one of the parameters linked to these frequency contributions is justified. The activation of SNS is indeed reflected in the variation of peak LF, peak VLF and %VLF.

3.1.1 Recovery Phase Analysis

All the features extracted by electro cardiac and electrodermal activity and compared in the previous paragraph taking into account only baseline and stress phases, have been calculated also for recovery stage. Each parameters has been investigated applying the KW test in order to assess if there were significant differences between baseline and recovery phases. None of the twenty-seven parameters resulted significant (Fig. 4), confirming that, after the stress phase, the wearable devices were able to measure that the tested subjects relaxed again and their physiological parameters were analogous to those calculated during baseline phase.

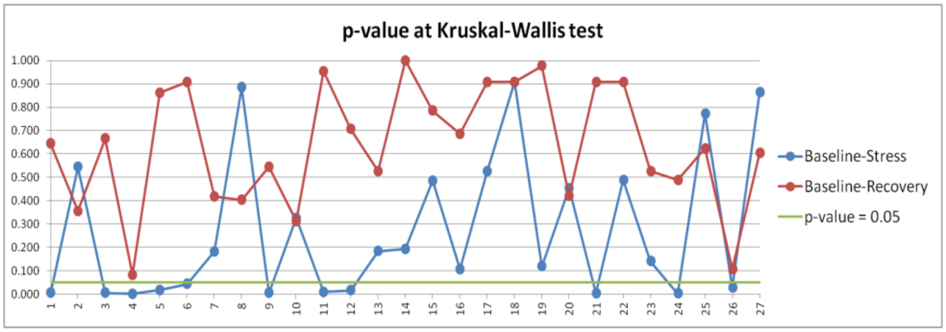

Figure 4. p-value at KW test for each extracted features. Significant parameters are below the green line (p-value=0.05).

This behaviour of the data is showed in Fig. 5, where the PCA results are reported for all the phases of the experimentation.

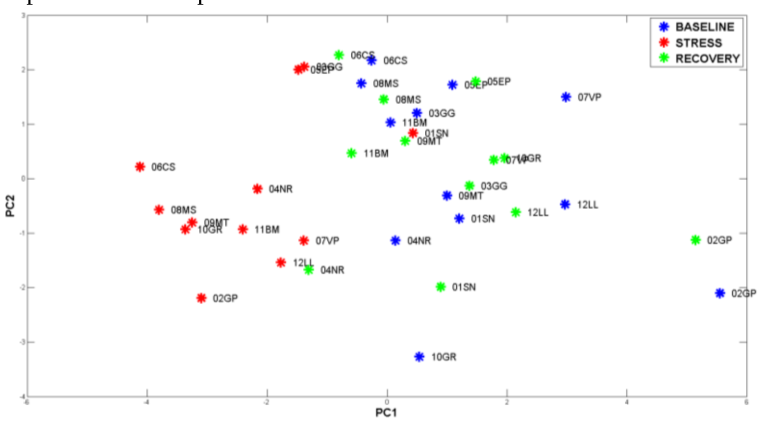

Figure 5. PCA results for Baseline (blue markers), Stress (red) and Recovery (green) phases. The data regarding recovery phase overlap those of the baseline. The first four PCs contain 85.4% of the overall variance.

3.2 Psychometric Instruments Evaluation

A comparison between the scores of the state tests administered before and after stress induction has been performed (Table 6) using the T-test. The KSS scores did not show significant differences between before and after stress induction phase, whereas the STAI-6 scores showed statistically significant differences between the two phases ($p<0.05$) indicating a recognisable level of anxiety in the tested subjects. The SSSQ scores also showed significance differences between pre and post stress induction tests ($p<0.01$) In particular, a highly statistically significant result ($p<0.01$) emerged from a subscale of SSSQ called "distress" that is the most important factor of SSSQ measuring the negative effect of the situation (Helton, 2004). A statistically significant result related to the variation of this subscale could mean that the stress induction phase effectively provided a negative effect on participants.

Table 6. Questionnaires results: mean values ± standard deviations and t-test significance.

Scale	Baseline			Stress			p-value	
KSS	3.8	±	1.3	3.3	±	0.6	0.089	
STAI-6	10.9	±	2.2	13.7	±	4.1	0.031	*
SSSQ	93.4	±	19.5	117.6	±	34.2	0.001	*
Distress	19.9	±	10.9	39.5	±	20.1	0.001	*
Task Management	36.5	±	4.1	35.8	±	6.1	0.639	
Worry	37.0	±	12.9	42.3	±	19.5	0.174	

* Significant difference at T-test between groups ($p<0.05$)

3.3 Correlation between physiological parameters and psychometric instruments

From the analysis of both physiological data and psychometric instruments it has been possible to notice a significant difference among the baseline phase and the stress one, indicating that these are valuable instruments to appreciate the arousal of anxiety and stress. The further step has been to assess the correlation between physiological features obtained from electro cardiac and electro dermal activities and questionnaires, in order to establish if it was possible to classify the stress level using psychometric instruments as reference. Unfortunately, the correlation between these two instruments was not high. The p-values calculated and disclosed in Table 7, did not show a significant correlation between physiological data and psychometric instruments.

Table 7. Correlation between psychometric instruments and physiological data

Scale	R^2	p-value	
KSS	0.77	0.019	*
STAI-6	0.53	0.373	
SSSQ	0.55	0.312	
Distress	0.53	0.357	
Task Management	0.40	0.695	

Worry	0.57	0.278

* Significant statistical values (p<0.05)

Among the psychometric scales used, KSS scale is the most correlated, showing a significant p-value. It is indicated to assess the level of sleepiness of the subject. The correlation with the variation of physiological data could explain that the stress induction phase provided a reduction of sleepiness, increasing the level of alarm and attention. The lack of significant correlation with the other scales has been probably due to the fact that, generally, self-reports provide valuable information but there could be problems with validity. Users of experimental studies often may not answer exactly how they are feeling. Rather, they answer questions as they feel others would answer them, or in a way they think the researcher wants them to answer. Furthermore, the psychometric responses could be dependent on participants' mood and state of mind on the day of the study (Elmes et al., 2011, Burke & Christensen, 2004).

In this study it has been chosen to use physiological measures because, as primary advantage, the participants can not consciously manipulate the activities of their ANS (Kidd & Breazeal, 2005, Picard et al., 2001, McCreadie & Tinker, 2005). Additionally, physiological measures offer a non-invasive method that can be used to determine the stress levels and reactions of participants interacting with technology (Picard et al.2001, Liu et al., 2006). Even if psychometric instruments did not provide a remarkable correlation with physiological response, it is possible to assert that physiological measures provide an indication about the variation in stress level of the tested subjects.

3.4 Data Classification

According to the aim of the paper, a classifier was implemented in order to identify the status of the subjects on the basis of the measured physiological signals. Basically, the classifier should be able to distinguish if a person is stressed or not.

For this purpose, the datasets acquired both in baseline and stress phases were used and, in particular, the parameters resulted significant at the KW test in distinguishing between the two phases have been taken into account (see par. 3.1).

The linear correlation between the significant parameters was calculated using the Pearson's coefficient and results were reported in Table 8.

Table 8. Pearson's coefficient of correlation between significant features

Feature Name	IBI_mean	HR_mean	SDHR	RMSSD	pNN50	peakVLF	%VLF	peakLF	Sum_RiseT ime	Decay_ Rate	Mean_GSR
IBI_mean	1.00	-0.98	-0.55	0.77	0.73	0.34	-0.29	0.28	-0.29	0.56	-0.38
HR_mean	-0.98	1.00	0.62	-0.73	-0.68	-0.35	0.31	-0.30	0.29	-0.50	0.34
SDHR	-0.55	0.62	1.00	-0.22	-0.22	-0.53	0.56	-0.50	0.44	-0.32	0.17
RMSSD	0.77	-0.73	-0.22	1.00	0.96	0.16	-0.23	0.10	-0.20	0.53	-0.27
pNN50	0.73	-0.68	-0.22	0.96	1.00	0.15	-0.28	0.07	-0.22	0.55	-0.35
peakVLF	0.34	-0.35	-0.53	0.16	0.15	1.00	-0.75	0.95	-0.36	0.37	-0.37
%VLF	-0.29	0.31	0.56	-0.23	-0.28	-0.75	1.00	-0.77	0.33	-0.38	0.39
peakLF	0.28	-0.30	-0.50	0.10	0.07	0.95	-0.77	1.00	-0.35	0.39	-0.33
Sum_RiseTime	-0.29	0.29	0.44	-0.20	-0.22	-0.36	0.33	-0.35	1.00	-0.76	0.30
Decay_Rate	0.56	-0.50	-0.32	0.53	0.55	0.37	-0.38	0.39	-0.76	1.00	-0.49
Mean_GSR	-0.38	0.34	0.17	-0.27	-0.35	-0.37	0.39	-0.33	0.30	-0.49	1.00

If the value of correlation between two features was at least 0.80, the less significant one was deleted.

Thus, a reduced number of eight parameters has been selected and used for Principal Component Analysis (PCA) that allowed to visualize the separation between subjects in baseline and stress phases in the space of the PCs as shown in fig. 6.

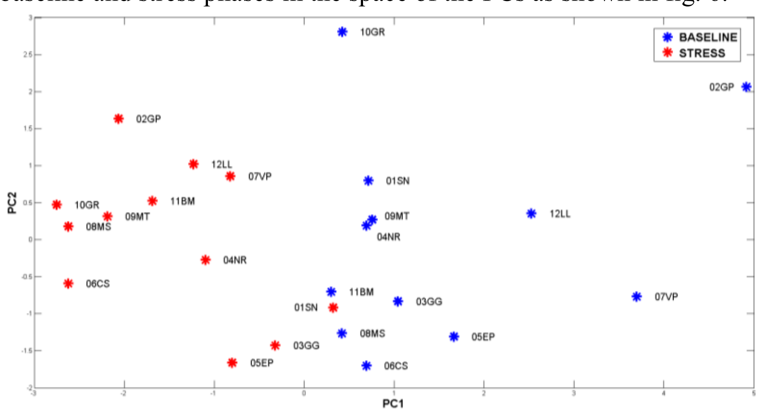

Figure 6. PCA synthesizes the differences in physiological parameters between baseline (blue markers) and stress phase (red markers). The first four PCs contains the 87.8% of the overall variance.

The first four PCs resulted by PCA have been adopted in order to implement a State Vector Machine (SVM) analysis. The aim was to classify the subjects as stressed or not stressed and to assess the quality of the system in terms of sensitivity, specificity and overall accuracy.

A k-fold cross-validation technique (k=5) was implemented to randomly select and assess training and testing datasets. A third order polynomial kernel for SVM algorithm was chosen as the more suitable to recognise the performances of the subjects belonging to baseline or stress phases. To reduce variability, multiple rounds of cross-validation were performed using different random partitions, and the validation results were averaged over the rounds. The algorithm performances resulted in Sensitivity = 93.3%, Specificity = 93.3% and Overall Accuracy 92.0%, thus the classifier showed a very good ability in correctly classifying the subjects in the two phases taken into account.

4 Conclusion

The presented work described features extraction and processing techniques used in HRV and GSR signals. In particular, the aim was to demonstrate the detection and monitoring of stress disease, with a satisfactory accuracy level using a non-obtrusive system. The SVM classifier provided indeed Sensitivity = 93.3%, Specificity = 93.3% and Overall Accuracy 92.0% in distinguishing stressed and not stressed status in the subjects involved in the study. Among the physiological features extracted, significant differences have been observed in some parameters, both for electro dermal activity and electro cardiac activity. This fact can be conferred to a concrete variation in physiological response due to a psychological stress induction. Then, it is possible to conclude that through physiological features it could be feasible to establish if a subject is stressed or not. The significant difference between the scores obtained by the subjects before and after the stress induction in both the STAI-6 questionnaire and SSSQ questionnaire confirms that the stress protocol designed reaches the goal of inducing a cognitive and emotional arousal. The evidence of the efficacy of the protocol is even more evidenced by the results of the distress subscale which seems very effective in evaluating the situational stress experienced by the subject.

Since physiological signals are influenced by a high level of variability among subjects, it is important to collect even small variations of signals in order to calculate the related features. For this purpose it is needed to take into account both the quality and accuracy of the devices used and the precision of the algorithms implemented.

Regarding psychometric instruments, there was not a remarkable correlation between physiological variations and scores obtained from the questionnaires. The only significant p-value obtained was related to the KSS scale, focused on the level of sleepiness, that probably changed among the different phases of the test, with a reduction in the stress induction phase.

In future, the extraction features and classification algorithms will be improved in order to obtain a real time system, able to detect stress levels of the user. The system will suggest also interventions such as physical exercises in order to reduce the stress level. This will support active ageing, allowing also to elderly to work until the retirement, under controlled conditions that could reduce the burden of stress related to the workload on the basis of personalised interventions.

References

1. Åkerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. International Journal of Neuroscience, 52(1-2), 29-37.
2. Barbeau A. (1980) Lecithin in Parkinson's disease. J Neural Transm Suppl. 1980;(16):187-93.
3. Burke, J., & Christensen, L. (2004). Educational research: Quantitative, qualitative, and mixed approaches. Boston: Pearson Education, Inc. Campbell KT, Forge E, Taylor L (2006). The effects of principal centers on professional isolation of school principals. School Leadership Review Summer/Fall, 2(1), 1-15.
4. Clifford, G. D. (2002). Signal processing methods for heart rate variability(Doctoral dissertation, Department of Engineering Science, University of Oxford).
5. de Santos Sierra, A., Ávila, C. S., Pozo, G. B. D., & Casanova, J. G. (2011, October). Stress detection by means of stress physiological template. In Nature and Biologically Inspired Computing (NaBIC), 2011 Third World Congress on (pp. 131-136). IEEE.
6. Elmes, D., Kantowitz, B., & Roediger III, H. (2011). Research methods in psychology. Nelson Education.
7. EU-OSHA (2016). European Agency for Safety and Health at Work. Retrieved from https://osha.europa.eu/en/tools-and-publications/publications/literature_reviews/calculating-the-cost-of-work-related-stress-and-psychosocial-risks
8. Gillett R. (2007) Assessment of working memory performance in self-ordered selection, Cortex. 2007 Nov;43(8):1047-56.
9. Haak, M., Bos, S., Panic, S., & Rothkrantz, L. J. M. (2009). Detecting stress using eye blinks and brain activity from EEG signals. Proceeding of the 1st driver car interaction and interface (DCII 2008), 35-60.
10. Haapalainen, E., Kim, S., Forlizzi, J. F., & Dey, A. K. (2010, September). Psycho-physiological measures for assessing cognitive load. In Proceedings of the 12th ACM international conference on Ubiquitous computing (pp. 301-310). ACM.
11. Healey, J. A., & Picard, R. W. (2005). Detecting stress during real-world driving tasks using physiological sensors. Intelligent Transportation Systems, IEEE Transactions on, 6(2), 156-166.
12. Healey, J., & Picard, R. (2000). SmartCar: detecting driver stress. In Pattern Recognition, 2000. Proceedings. 15th International Conference on (Vol. 4, pp. 218-221). IEEE.
13. Helton, W. S. (2004, September). Validation of a short stress state questionnaire. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 48, No. 11, pp. 1238-1242). SAGE Publications.
14. Kaida, K., Takahashi, M., Åkerstedt, T., Nakata, A., Otsuka, Y., Haratani, T., & Fukasawa, K. (2006). Validation of the Karolinska sleepiness scale against performance and EEG variables. Clinical Neurophysiology, 117(7), 1574-1581.
15. Karthikeyan, P., Murugappan, M., & Yaacob, S. (2012). Descriptive analysis of skin temperature variability of sympathetic nervous system activity in stress.Journal of Physical Therapy Science, 24(12), 1341-1344.

16. Kidd, C. D., & Breazeal, C. (2005, April). Human-robot interaction experiments: Lessons learned. In Proceeding of AISB (Vol. 5, pp. 141-142).
17. Kurniawan, H., Maslov, A. V., & Pechenizkiy, M. (2013, June). Stress detection from speech and galvanic skin response signals. In Computer-Based Medical Systems (CBMS), 2013 IEEE 26th International Symposium on (pp. 209-214). IEEE.
18. Lansbergen MM, Kenemans JL, van Engeland H. (2007) Stroop interference and attention-deficit/hyperactivity disorder: a review and meta-analysis. Neuropsychology. 2007 Mar;21(2):251-62.
19. Lim, C. K. A., & Chia, W. C. (2015). Analysis of Single-Electrode EEG Rhythms Using MATLAB to Elicit Correlation with Cognitive Stress.International Journal of Computer Theory and Engineering, 7(2), 149.
20. Liu, C., Rani, P., & Sarkar, N. (2006, October). Affective state recognition and adaptation in human-robot interaction: A design approach. In Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on (pp. 3099-3106). IEEE.
21. Lundberg, U., Kadefors, R., Melin, B., Palmerud, G., Hassmén, P., Engström, M., & Dohns, I. E. (1994). Psychophysiological stress and EMG activity of the trapezius muscle. International journal of behavioral medicine, 1(4), 354-370.
22. Mali, B., Zulj, S., Magjarevic, R., Miklavcic, D., & Jarm, T. (2014). Matlab-based tool for ECG and HRV analysis. Biomedical Signal Processing and Control, 10, 108-116.
23. Marteau, T. M., & Bekker, H. (1992). The development of a six-item short-form of the state scale of the Spielberger State—Trait Anxiety Inventory (STAI).British Journal of Clinical Psychology, 31(3), 301-306.
24. McCreadie, C., & Tinker, A. (2005). The acceptability of assistive technology to older people. Ageing and society, 25(01), 91-110.
25. Medtronic (2015). Zephyr™ Performance System. Retrieved from http://www.zephyranywhere.com/products/bioharness-3
26. Miyake A, Emerson MJ, Friedman NP (2000) Assessment of executive functions in clinical settings: problems and recommendations. Semin Speech Lang. 2000;21(2):169-83.
27. Orsila, R., Virtanen, M., Luukkaala, T., Tarvainen, M., Karjalainen, P., Viik, J., ... & Nygård, C. H. (2008). Perceived mental stress and reactions in heart rate variability—a pilot study among employees of an electronics company.International Journal of Occupational Safety and Ergonomics, 14(3), 275-283.
28. Park, B. (2009). Psychophysiology as a tool for HCI research: promises and pitfalls. In Human-Computer Interaction. New Trends (pp. 141-148). Springer Berlin Heidelberg..
29. Pfaff, M. S. (2012). Negative affect reduces team awareness the effects of mood and stress on computer-mediated team communication. Human Factors: The Journal of the Human Factors and Ergonomics Society, 54(4), 560-571.
30. Picard, R. W., Vyzas, E., & Healey, J. (2001). Toward machine emotional intelligence: Analysis of affective physiological state. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 23(10), 1175-1191.
31. Ritz, T., Steptoe, A., DeWilde, S., & Costa, M. (2000). Emotions and stress increase respiratory resistance in asthma. Psychosomatic Medicine, 62(3), 401-412.
32. Rottenberg, J., Wilhelm, F. H., Gross, J. J., & Gotlib, I. H. (2002). Respiratory sinus arrhythmia as a predictor of outcome in major depressive disorder.Journal of affective disorders, 71(1), 265-272..
33. Seaward, B. L. (1999). Managing stress: Principles and strategies for health and wellbeing. Jones & Bartlett Pub.
34. Sharma, N., & Gedeon, T. (2012). Objective measures, sensors and computational techniques for stress recognition and classification: A survey.Computer methods and programs in biomedicine, 108(3), 1287-1301..

35. Shimmer (2016). Shimmer3 GSR + Unit. Retrieved from http://www.shimmersensing.com/shop/shimmer3-wireless-gsr-sensor
36. Singh, R. R., Conjeti, S., & Banerjee, R. (2012, February). Biosignal based on-road stress monitoring for automotive drivers. In Communications (NCC), 2012 National Conference on (pp. 1-5). IEEE.
37. Statistic Brain (2015). Stress Statistic. Retrieved from http://www.statisticbrain.com/stress-statistics
38. Stemmler, G., Heldmann, M., Pauls, C. A., & Scherer, T. (2001). Constraints for emotion specificity in fear and anger: The context counts.Psychophysiology, 38(02), 275-291.
39. Sun, F. T., Kuo, C., Cheng, H. T., Buthpitiya, S., Collins, P., & Griss, M. (2010). Activity-aware mental stress detection using physiological sensors. In Mobile computing, applications, and services (pp. 211-230). Springer Berlin Heidelberg.
40. Taelman, J., Vandeput, S., Spaepen, A., & Van Huffel, S. (2009). Influence of mental stress on heart rate and heart rate variability. In 4th European conference of the international federation for medical and biological engineering(pp. 1366-1369). Springer Berlin Heidelberg.
41. Tsigos, C., & Chrousos, G. P. (1994). Physiology of the hypothalamic-pituitary-adrenal axis in health and dysregulation in psychiatric and autoimmune disorders. Endocrinology and metabolism clinics of North America, 23(3), 451-466.
42. Ulstein, I., Wyller, T. B., & Engedal, K. (2007). High score on the Relative Stress Scale, a marker of possible psychiatric disorder in family carers of patients with dementia. International journal of geriatric psychiatry, 22(3), 195-202..
43. Watkins, L. L., Grossman, P., Krishnan, R., & Blumenthal, J. A. (1999). Anxiety reduces baroreflex cardiac control in older adults with major depression. Psychosomatic Medicine, 61(3), 334-340.