Measuring psychosocial stress with heart rate variability-based methods in different health and age groups

Seipäjärvi, Santtu M.

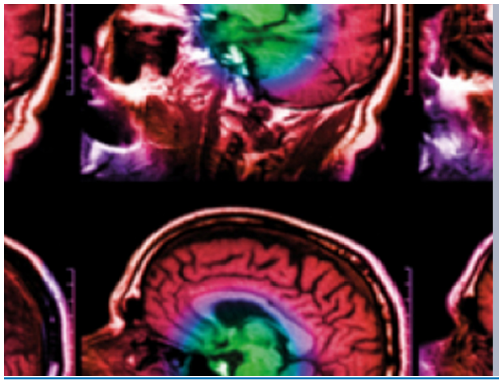
2022-05-31

Seipäjärvi , S M , Tuomola , A , Juurakko , J , Rottensteiner , M , Rissanen , A-P E , Kurkela , J L O , Kujala , U M , Laukkanen , J A & Wikgren , J 2022 , ' Measuring psychosocial stress with heart rate variability-based methods in different health and age groups ' , Physiological Measurement , vol. 43 , no. 5 , 055002 . https://doi.org/10.1088/1361-6579/ac6b7c

http://hdl.handle.net/10138/345830
https://doi.org/10.1088/1361-6579/ac6b7c

PAPER • OPEN ACCESS

Measuring psychosocial stress with heart rate variability-based methods in different health and age groups



IOP Publishing Physiol. Meas. 43 (2022) 055002 https://doi.org/10.1088/1361-6579/ac6b7c

Physiological Measurement

IPEM Institute of Physics and Engineering in Medicine

OPEN ACCESS

RECEIVED
18 October 2021

REVISED
14 March 2022

ACCEPTED FOR PUBLICATION
28 April 2022

PUBLISHED
25 May 2022

PAPER

Measuring psychosocial stress with heart rate variability-based methods in different health and age groups

Santtu M Seipäjärvi[1,*], Anniina Tuomola[1], Joona Juurakko[1], Mirva Rottensteiner[2,3], Antti-Pekka E Rissanen[2,4,5], Jari L O Kurkela[1], Urho M Kujala[3], Jari A Laukkanen[2] and Jan Wikgren[1]

[1] Centre for Interdisciplinary Brain Research, Department of Psychology, University of Jyväskylä, Jyväskylä, Finland
[2] Central Finland Health Care District, Jyväskylä, Finland
[3] Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, Finland
[4] Department of Sports and Exercise Medicine, Clinicum, University of Helsinki, Helsinki, Finland
[5] HULA—Helsinki Sports and Exercise Medicine Clinic, Foundation for Sports and Exercise Medicine, Helsinki, Finland
[*] Author to whom any correspondence should be addressed.

E-mail: santtu.m.seipajarvi@jyu.fi

Keywords: cortisol, heart rate variability, perceived stress, wellbeing, stress

Supplementary material for this article is available online

Abstract

Objective. Autonomic nervous system function and thereby bodily stress and recovery reactions may be assessed by wearable devices measuring heart rate (HR) and its variability (HRV). So far, the validity of HRV-based stress assessments has been mainly studied in healthy populations. In this study, we determined how psychosocial stress affects physiological and psychological stress responses in both young (18–30 years) and middle-aged (45–64 years) healthy individuals as well as in patients with arterial hypertension and/or either prior evidence of prediabetes or type 2 diabetes. We also studied how an HRV-based stress index (Relax-Stress Intensity, RSI) relates to perceived stress (PS) and cortisol (CRT) responses during psychosocial stress. *Approach.* A total of 197 participants were divided into three groups: (1) healthy young (HY, $N = 63$), (2) healthy middle-aged (HM, $N = 61$) and (3) patients with cardiometabolic risk factors (Pts, $N = 73$, 32–65 years). The participants underwent a group version of Trier Social Stress Test (TSST-G). HR, HRV (quantified as root mean square of successive differences of R–R intervals, RMSSD), RSI, PS, and salivary CRT were measured regularly during TSST-G and a subsequent recovery period. *Main results.* All groups showed significant stress reactions during TSST-G as indicated by significant responses of HR, RMSSD, RSI, PS, and salivary CRT. Between-group differences were also observed in all measures. Correlation and regression analyses implied RSI being the strongest predictor of CRT response, while HR was more closely associated with PS. *Significance.* The HRV-based stress index mirrors responses of CRT, which is an independent marker for physiological stress, around TSST-G. Thus, the HRV-based stress index may be used to quantify physiological responses to psychosocial stress across various health and age groups.

1. Introduction

Whether a certain environmental demand is perceived as psychologically stressful or not depends on individual's evaluation of the potential threat and the availability of personal resources (Lazarus and Folkman 1984, Cohen *et al* 2016). Biologically, stress can be defined as perturbations of physiological systems, such as autonomic nervous (ANS) or endocrine systems, maintaining homeostasis (Cohen *et al* 2016). The counterpart of stress is the ability to recover from these perturbations. This ability is compromised in the case of chronically elevated stress, which is associated with an increased risk of several diseases such as cardiovascular disease (Kivimäki and Steptoe 2018), type 2 diabetes (Nyberg *et al* 2014), and mental health problems (Madsen *et al* 2017). Prolonged exposure to psychological stress is thought to mediate its negative effects by changing the reactivity of the stress systems, which

might again serve as a potential marker in the prevention of various negative health outcomes (Cole *et al* 1999, Chrousos 2009, de Rooij 2013).

The activity of the sympathetic part of ANS increases under stressful circumstances. This is seen as increased activity in the sympatho-adrenomedullary (SAM) system and hypothalamus-pituitary-adrenal (HPA) axis (Padgett and Glaser 2003). The SAM system is responsible for 'fight or flight' responses regulating the activity of cardiovascular and rapid metabolic (adrenaline, noradrenaline) processes preparing the body for the immediate threat (Padgett and Glaser 2003). The HPA axis reacts to stress by synthesizing and secreting glucocorticoids like cortisol (CRT) (Bozovic *et al* 2013). Phenomena related to both the SAM system and the HPA axis can be quantified by measuring heart rate variability (HRV) and salivary CRT levels, respectively. The return of physiological responses to their basal state is often interpreted as a physiological system recovering to its pre-stressor or 'normal' state in which the activity of the parasympathetic part of ANS plays a vital role (Cole *et al* 1999, Mezzacappa *et al* 2001). As with approaches used in sports (Buchheit 2014), the intensity and the recovery of the psychosocial stress responses could offer valuable information about the state of the body and identify health risks (Thayer *et al* 2009, Weber *et al* 2010).

Despite known negative health outcomes and economical burden of stress-related adverse effects costs to organizations and society (Hassard *et al* 2018), the objective and cost-efficient means to measure stress and recovery are lacking. Due to the complexity of the physiological phenomena, there are no unequivocal methods to measure stress in everyday life contexts. Measuring heart rate (HR) and HRV (e.g. root mean square of successive differences of R–R intervals, RMSSD) offers a practical tool to record physiological signals of stress (Taelman *et al* 2009, Melillo *et al* 2011). However, there are multiple challenges related to this approach since HR and HRV responses to psychosocial stress are highly dependent on individual factors such as health status (Koskinen *et al* 2009, Assoumou *et al* 2010), age (O'Brien *et al* 1986), sex (Umetani *et al* 1998), fitness level (Rimmele *et al* 2007, Mücke *et al* 2018), psychological appraisal skills (Gaab *et al* 2005), genes (Boomsma *et al* 1990, Kupper *et al* 2004), respiration rate (Schipke *et al* 1999), and recovery status from earlier exercises (Mourot *et al* 2004). Similar difficulties are faced when endocrine stress responses are measured (Kudielka and Kirschbaum 2003, Otte *et al* 2005). In addition, the relationship between perceived psychological and physiological stress has not been consistent between studies (Campbell and Ehlert 2012). Wearable technology used for self-monitoring of wellbeing may offer cost-effective tools to quantify stress and help to prevent the negative outcomes of excess stress and inefficient recovery. However, little is known about the validity of wearable technologies to measure physiological responses to psychosocial stress in different health and age populations.

The purpose of this study was twofold: first, we aimed to determine how psychosocial stress affects physiological and psychological responses in individuals with different ages and clinically relevant health conditions such as arterial hypertension and impaired glucose metabolism. Second, we aimed to study how an HRV-based stress index, provided by a wearable technology, is related with other commonly used stress variables (HR, RMSSD), perceived stress (PS), and CRT responses during and around psychosocial stress. To induce psychosocial stress, a group version (TSST-G; von Dawans *et al* 2011) of a commonly used Trier Social Stress Test (TSST; Kirschbaum *et al* 1992) was used since previous studies have shown these protocols to reliably induce cardiovascular, endocrine, and psychological stress responses in various participant populations (Rimmele *et al* 2007, von Dawans *et al* 2011, Klaperski *et al* 2014).

2. Methods

Data for this study were collected as part of a research collaboration entitled 'Heart rate variability analytics to support behavioural interventions for chronic disease prevention and management' (HealthBeat) in Jyväskylä, Finland. The study protocol conformed to the Declaration of Helsinki and was approved by the ethics committee of Central Finland Hospital District, Jyväskylä, Finland (Dnro 23U/2018). Each participant gave written informed consent before participation in the study.

2.1. Participants

The HealthBeat study consisted of two separate populations: healthy participants and patients with cardiometabolic risk factors. The healthy participants were recruited via online advertisements and email lists. The patients were recruited via online advertisements, public advertisements on local noticeboards, and by contacting the local health care providers who informed their patients about the research collaboration. The recruitment process and the workflow of this study are described in figure S2 (available online at stacks.iop.org/PMEA/43/055002/mmedia) in supplementary materials and the participant characteristics are presented in tables 2–3. After completing the study protocol the participants received a movie ticket and were offered an optional Firstbeat lifestyle assessment report with a 30 min feedback session.

Table 1. Common exclusion criteria concerning both healthy and patient participants.

Chronic cardiac condition (e.g. chronic atrial fibrillation, heart failure, ischaemic heart disease, pacemaker, significant or non-specified valvular disease)
Left bundle branch block
Pregnancy
Psychotic disorder or some other unstable psychiatric disorder
Specific medications: β-blockers, insulin, serotonin and noradrenaline reuptake inhibitors, tricyclic antidepressants
Substance abuse
Symptomatic/unstable disorder of thyroid gland

2.1.1. Healthy participants

Recruitment and preparticipation screening of the healthy participants was conducted by the researchers of the University of Jyväskylä. The inclusion criterion was age between 18–30 and 45–64 years. The exclusion criteria included the criteria common for all participants in the HealthBeat study (table 1) and any chronic neurological disease. After recruitment process a total of 148 healthy individuals aged between 18–30 and 45–64 years participated in the study. After concerning additional self-reported information (especially medication) obtained from participants, a total of 24 participants reporting disease or medication were excluded from the population of the healthy participants (See figure S2 for details). As a result, 124 healthy participants (88 females) were included in the final analysis. Metabolic syndrome risk factor (0–5) indicating the number of risk factors for the individual participant was determined according to the criteria of International Diabetes Federation (Alberti *et al* 2006). For healthy participants, the risk factor was set as 0 unless the participant had body mass index (BMI) $\geqslant 30\,\mathrm{kg\,m^{-2}}$, after which the risk factor was set as 1.

2.1.2. Patients with cardiometabolic risk factors

Regarding the patient participants of this study, the inclusion criteria were (1) age between 18 and 64 years, (2) BMI $<40\,\mathrm{kg\,m^{-2}}$, (3) either previous evidence of prediabetes (i.e. impaired fasting glucose and/or impaired glucose tolerance) or type 2 diabetes diagnosed no more than five years ago, and/or diagnosed arterial hypertension, and (4) overall physical function not preventing the participant from safely performing the experiments including cardiopulmonary exercise testing (CPET). The exclusion criteria of the patients included the criteria common for all participants in the HealthBeat study (table 1) as well as anemia, breastfeeding, cancer, chronic obstructive pulmonary disease, cerebrovascular disease, clinically significant hypertension-mediated organ damage, diagnosed diabetes-related microvascular disease (i.e. nephropathy, neuropathy, retinopathy), obstructive sleep apnoea requiring continuous positive airway pressure treatment, secondary hypertension, a significant deficit in overall physical function, and symptomatic/unstable asthma. Patients potentially eligible for participating in the study went through preparticipation health screening conducted by a physician and a nurse from Central Finland Health Care District. The preparticipation health screening consisted of a interpretation of individual's medical history, clinical status, resting blood pressure, resting 12-lead electrocardiography (ECG), and weight and height measurements. The antecubital venous blood samples were drawn after an overnight fast in an accredited laboratory (FimLab Laboratoriot Ltd, Jyväskylä, Finland) complemented the health screening, and included assessment of blood count, lipid profile, glycemic control, electrolyte balance, and renal function. Overall, the preparticipation health screening of the patients focused on evaluating individual's signs or symptoms and/or known cardiovascular, metabolic, or renal disease, and the current level and type of physical activity. Amount of metabolic syndrome risk factors was determined based on the preparticipation health screening according to the International Diabetes Federation criteria (Alberti *et al* 2006).

According to the preparticipation health screening, 87 patients met the inclusion criteria and were invited to CPET. Of those, 73 participants (56 females) participated in the psychosocial stress test (See figure S2 for details).

2.2. Psychosocial stress test: procedure

Three to four participants at a time participated in a two-hour experimental session starting at either 2 p.m. or 4 p.m. These fixed afternoon onset times were selected to control for the effects of the circadian rhythm on the measured physiological variables.

2.2.1. Before the test

Participants were instructed to start the HRV measurement on the day preceding the stress test. Participants were advised to avoid physical stress and alcohol for 24 h and smoking for two hours before the experimental

Table 2. Characteristics of the participants included in final analysis.

Group	*N*	Sex (F/M)	Age (years)		BMI (kg m^{-2})		Estimated/measured $V^{\cdot}O_{2peak}$ (ml min^{-1} kg^{-1})	
Healthy young (HY)	63	45/18	26 ± 3	26 (23−27)	23.3 ± 2.8	23.0 (21.7−25.0)	45.0 ± 5.7	44.4 (41.2−49.1)
Healthy middle-aged (HM)	61	32/29	52 ± 5	51 (48−56)	26.3 ± 4.2	25.0 (23.4−28.3)	35.4 ± 7.5	34.7 (29.0−41.0)
Patients (Pts)	73	56/17	53 ± 8	55 (50−59)	28.5 ± 4.6	27.4 (24.8−31.4)	30.8 ± 6.3	30.1 (26.3−34.2)

Values for Age, BMI, and Estimated/measured $V^{\cdot}O_{2peak}$ are presented as mean values ± SD and median (IQR). Estimation of peak pulmonary O_2 oxygen uptake ($V^{\cdot}O_{2peak}$) for the healthy participants was produced by a commercial technology based on HR and body acceleration data collected during a self-paced walk. Patients' $V^{\cdot}O_{2peak}$ was measured directly during CPET.

Table 3. Cardiometabolic and -vascular risk profile of the patient participants ($N = 73$).

Blood samples	
Total cholesterol (mmol l^{-1})	4.9 ± 0.9
LDL cholesterol (mmol l^{-1})	3.0 ± 0.9
HDL cholesterol (mmol l^{-1})	1.5 (1.25–1.80)
Triglycerides (mmol l^{-1})	1.1 (0.80–1.85)
Fasting glucose (mmol l^{-1})	5.8 (5.30–6.30)
HbA_{1c} (mmol mol^{-1})	38.0 (35.0–41.5)
Blood pressure	
Resting systolic blood pressure (mmHg)	134 ± 13
Resting diastolic blood pressure (mmHg)	83 ± 7
Cardiometabolic and—vascular conditions	
Arterial hypertension	62 (85%)
Prediabetes[a]	16 (22%)
Type 2 diabetes	16 (22%)
Metabolic syndrome[b]	51 (70%)
Cardiometabolic and—vascular medications	
ACE or ARB	53 (73%)
Calcium channel blockers	17 (23%)
Diuretics	10 (14%)
Statins	12 (16%)
Tablet treatment for diabetes	14 (19%)
Smoking	
Yes	5 (7%)
No	68 (93%)

Values are presented as mean values ± SD for normally distributed continuous variables, median (IQR) for nonnormally distributed continuous variables, and *N* (%) for categorical variables.
ACE, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker.
[a] Evidence of impaired fasting glucose (6.1–6.9 mmol l^{-1}) previously and/or in this study, and/or previous evidence of impaired glucose tolerance, but no type 2 diabetes.
[b] As defined by the International Diabetes Federation.

session. The participants were informed that they would participate in a stress test but no specific details about the upcoming test were revealed beforehand.

2.2.2. Arrival

The test procedure started in a lobby of a university building, where the experiment leader met the participants one by one and instructed them to avoid any unnecessary talking to other participants and to follow the upcoming instructions. During this short briefing, the participants reported their education and occupation. This information was later utilized in determining the work assignment the participant was applying for in a mock job interview. After personal instructions, the participants were guided to an experimental room, where everyone was seated and the actual experimental procedure started. The experimental room was equipped with a fake video camera.

2.2.3. Psychosocial stress test

A group version (TSST-G) (von Dawans *et al* 2011) of the commonly used Trier Social Stress Test (TSST; Kirschbaum *et al* 1992) was used to study psychobiological responses to stress. The experimental procedure (figure 1) went as follows:

(i) An initiation period, during which the participants sat down for 20–25 min, listened to instructions for the upcoming experimental session, and answered questionnaires on their background information and physical activity habits. Participants sat next to each other separated by removable walls so that they could not see each other. Participants were facing in the same direction in which a fake two-person committee panel, consisting of research team members, would later be seated.

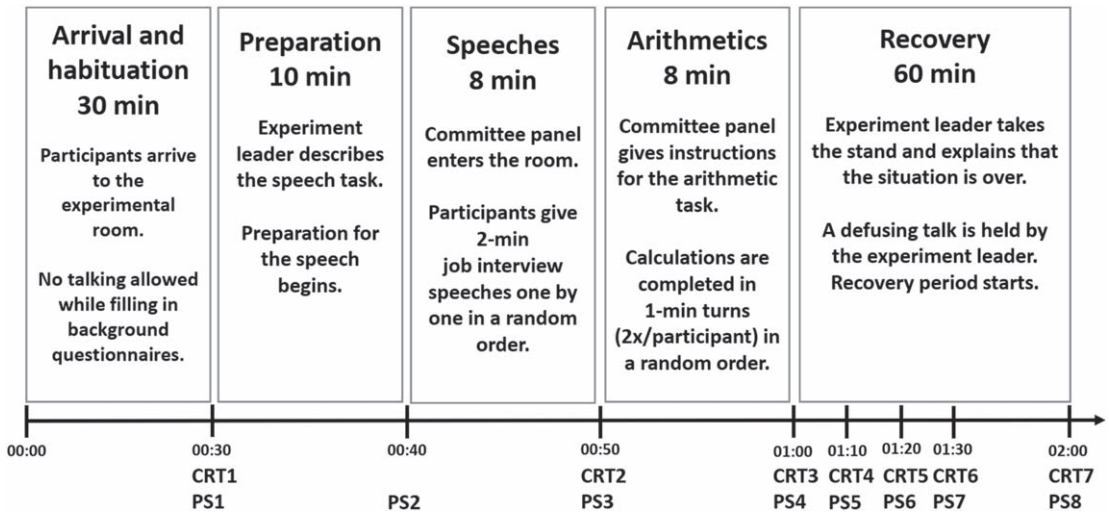


Figure 1. Design and timeline (h:min) of the stress test protocol. CRT1-CRT7 refer to salivary cortisol samples and PS1–PS8 refer to perceived stress questionnaires.

(ii) A 10 min anticipation period during which the participants prepared to give a 2 min speech in a mock job interview and answer possible questions made by the committee panel. To increase engagement to the task, the jobs applied were individually matched based on the short interview made at arrival.

(iii) Job interview speeches in front of the committee panel from all the participants in a pseudorandomized order (10 min). The committee panel did not provide any feedback or encouragement during the speeches. If the participant finished the speech too early or kept quiet, the panel reacted to it in a standardized way and advised the participant to continue.

(iv) Two 1 min arithmetic tasks, where the participants were, in an unpredictable pseudorandom order, told to start counting backwards from a random three-digit number at the steps of e.g. 23. Participants were instructed to count out loud as fast and accurately as possible and to start over from the beginning if they made a mistake.

(v) A 60 min resting period that started with a defusing talk held by the experiment leader during which it was made clear that no further tasks were involved. The participants were allowed to talk, use the mobile phones, excluding calls or work email, and use the restroom if needed. The experiment leader was the one who controlled and gave all the instructions related to the procedure, except for the protocol parts ii and iii that were led by the panel. In the case of fewer participants than the appropriate four, the committee panel was instructed to stall during the stress tasks to maintain a similar schedule and stress exposure for every session.

2.3. Psychosocial stress test: measurements

HRV data were measured continuously, and PS and salivary CRT were measured regularly during the experimental protocol. The timetable for different measurements is illustrated in figure 1.

2.3.1. HRV-based stress

Beat-to-beat R–R intervals of the heart and 3-axis acceleration signal were measured with a wearable one-lead ECG device (Firstbeat Bodyguard 2, Firstbeat Technologies Ltd, Jyväskylä, Finland) with an ECG sampling frequency of 1000 Hz and R–R interval accuracy of 1 ms, and movement sampling frequency of 12.5 Hz. The participants wore the device day and night (during sleeping) for three consecutive days, starting one day before psychosocial stress test and ending one day after the stress test. These data were also utilitized in other parts of the HealthBeat project. The participants provided their height, body mass, age, sex, self-reported physical activity level and additional background information needed for accurate calculation of HRV-based stress index.

The R–R interval and movement data were analyzed with the Firstbeat Life service (Firstbeat Technologies Ltd) to provide HRV-based stress and recovery information (Relax-Stress Intensity, RSI) and traditional parameters of cardiac autonomic activity including HR and an indicator of HRV (RMSSD). The method includes artefact detection and correction for falsely detected, missed, or premature heartbeats and movement artefacts.

RSI was analyzed by first detecting physiological state of the body by distinguishing stress and recovery reactions from physical activity and other states by utilizing R–R interval and body movement data for evaluating physiological phenomena such as respiration rate, oxygen consumption, excess post-exercise oxygen consumption, and ANS balance. Thereafter, when stress, i.e. sympathetic dominance of the body, or recovery, i.e. parasympathetic dominance of the body, were detected, also the intensity of such phenomena were analyzed with values ranging from −100 to +100. The closer the RSI values are to zero, the lower the intensity of reaction is and, accordingly, −100 means very high momentary stress and +100 extremely relaxed state in the body. Value of +100 (=maximum recovery) is reached when the parasympathetic activity of the ANS is high, the person's HR level is close to individual resting state, and HRV is large. On the contrary, RSI of −100 (=maximum stress) is reached when the sympathetic activity of the ANS is high, HR is markedly elevated from resting levels, and HRV is reduced without any physical activity-related reason. For more information about the method, the reader is referred to the supplementary materials (section 1.1) and whitepapers describing the methodology (for example Firstbeat technologies Ltd 2005, 2014a, 2014b).

2.3.2. Cortisol (CRT) and perceived stress (PS)

PS questionnaires were fulfilled and salivary samples for CRT were collected regularly during the stress test protocol (figure 1). Seven saliva samples were collected during the protocol as follows: after the habituation period (baseline), after the speech, after the arithmetic task and at 10 min, 20 min, 30 min, and 60 min during the recovery period. The samples were collected in Salivette Cortisol tubes (Sarstedt, Nürnbrecht, Germany). The tubes were later centrifuged for 3 min at 1000 × g, the clear saliva was transferred to microcentrifuge tubes and frozen at −20 °C until analysis. Cortisol values were analyzed using electrochemiluminescence immunoassay on an Immulite 2000 analyzer (Siemens, Llanberis, UK). Intra-assay and total coefficients of variation were 6.0% and 7.8%, respectively. Level of detection was 5.5 nmol l^{-1}. The PS was measured with 0–10 numeric rating scale, and the participants were instructed to evaluate their stress from the preceding 10 min period. The PS questionnaires were filled in at 30 min, 40 min, 50 min, 60 min, 70 min, 80 min, 90 min, and 120 min from the onset of the experiment.

2.4. Determination of cardiorespiratory fitness

2.4.1. Estimated $V\dot{O}_{2peak}$ of the healthy participants

$V\dot{O}_{2peak}$ of each healthy participant was estimated on a separate occasion by a self-paced walk: the participants were instructed to conduct a self-paced 30 min walk while wearing the Firstbeat Bodyguard 2 ECG device (Firstbeat Technologies Ltd, Jyväskylä, Finland) described in the previous section 2.3. The device includes technology, which provides an estimate of $V\dot{O}_{2peak}$ based on HRV and triaxial body acceleration (Smolander *et al* 2008, Smolander *et al* 2011, Robertson *et al* 2015). The participants were instructed to perform the walk on a separate occasion most suitable for them after the experimental session and to choose a steady route for walking while avoiding talking and carrying any extra load.

2.4.2. Measured $V\dot{O}_{2peak}$ of the patients

$V\dot{O}_{2peak}$ of each patient was measured during CPET, which was performed on a treadmill until individual volitional task failure according to the USAFSAM protocol (Wolthuis *et al* 1977). During CPET, $V\dot{O}_2$ was directly measured by measuring pulmonary ventilation (a volume turbine; Triple V®, Erich Jaeger, Friedberg, Germany) and gas concentrations (Oxycon Pro® Version 5.0, VIASYS Healthcare GmbH, Hoechberg, Germany) breath by breath. Measured $V\dot{O}_{2peak}$ was determined as the highest 30 s $V\dot{O}_2$ average detected during the test (American Thoracic Society and American College of Chest Physicians 2003).

2.5. Statistical analysis

Repeated measures ANOVA was used to compare the effect of time on measured stress variables for each group. All variables except PS were checked for outliers. Exclusion criteria for outliers were set to an absolute $|Z\text{-score}| \geqslant 2.68$. This was based on the definition of Tukey (1977) , according to which values located 1.5 interquartile ranges outside lower and upper quartiles are regarded as outliers; thus, as the quartiles for normal distribution are −0.67 and 0.67 standard deviations, the interquartile range thereby equals 1.34 standard deviations, and 1.5 times 1.34 equals 2.01, Z-scores less than −2.68 or greater than 2.68 are regarded as outliers. Ln-transformation was applied in case of nonnormally distributed data. Post hoc tests were Bonferroni corrected.

Mixed model ANOVA was conducted to compare the effects of time between different groups. Post hoc tests for significant main effects were Bonferroni corrected. In case of a significant interaction, simple main effects were calculated with Bonferroni correction.

Total stress responses for HRV-based data were calculated to estimate the total stress load during the tasks. The time frame used to calculate the total stress load was from the habituation to the end of the arithmetics (00:20–01:00). This period was chosen because the observed responses during the habituation period signal the presence of physiological and psychological stress. Total responses of RSI (RSI_{Total}), RMSSD ($RMSSD_{Total}$), and HR (HR_{Total}) were calculated as a sum of the average values measured in 10 min periods, and in the case of PS (PS_{Total}), as a sum of absolute PS values. RMSSD and HR values were referenced to values measured at the end of the experimental session (01:50–02:00) that were thought to serve as baseline values. This baseline value was chosen (1) to measure baseline values in a controlled situation, (2) to ensure a similar body posture, which affects cardiac hemodynamics (Tulen *et al* 1999), and (3) because the amount of PS was at its lowest during this time window.

Total stress response for cortisol (CRT_{AUCg}) was calculated using the area under the curve relative to the ground (AUCg) method (Pruessner *et al* 2003). CRT_{AUCg} was calculated starting from an interpolated value between CRT1 and CRT2 until CRT6 (CRT1/CRT2, CRT2.., CRT6). This was done because CRT responses are measurable from the saliva with a delay of approximately 10–20 min (Schlotz *et al* 2008, von Dawans *et al* 2011), and thus, this time frame was estimated to reflect the area most related to the CRT responses accumulated during the 00:20–01:00 time period.

Correlations were calculated using Spearman's correlation due to occasional violations in linearity assumptions. Correlations were calculated separately for every group after excluding outliers suspected as erroneous measurement (CRT_{AUCg} $n = 1$). ***Correlations*** *between changes in perceived stress and HRV-based stress* were calculated for two occasions including change from habituation to actual stress tasks (habituation to stress) and from arithmetics to recovery period (stress to recovery). Changes in PS were individually determined as the largest change from the habituation period to stress tasks (habituation to stress) and from arithmetics to the recovery period (stress to recovery). Absolute changes in RSI, HR, and RMSSD were then calculated individually from the same time intervals.

Regression analysis for determinants of RSI score. Outliers were checked for CRT_{AUCg}, HR_{Total}, $RMSSD_{Total}$, and RSI_{Total} responses. Exclusion criteria were set to an absolute $|\text{Z-score}| \geqslant 2.68$ (Tukey 1977). The regression model was created by excluding non-significant variables one by one.

3. Results

Average results and repeated measures ANOVA scores are presented in table 4 separately for each group and all groups combined (all). Additional data grouped by age and sex along with measures in the HRV frequency domain (LF, HF, LF/HF and total power) can be found in the supplementary materials.

3.1. Effects of the stress protocol

3.1.1. Relax-stress intensity

RSI decreased in every group (Post hoc $p < 0.001$) from habituation to preparation, speeches and arithmetic tasks (figure 2). RSI increased (Post hoc $p < 0.001$) in every group from arithmetics to the start of recovery (recovery 0–10 min). A significant main effect for group was observed ($F = 16.15, p < 0.001$). Post hoc tests showed that **Pts** had a significantly lower RSI than **HM** and **HY** ($p < 0.001$) when averaged across all points of time.

3.1.2. Heart rate

HR (figure 3) increased in every group (Post hoc $p < 0.001$) from habituation to preparation, speeches and arithmetics and decreased in every group (Post hoc $p < 0.001$) from arithmetics to the start of recovery (recovery 0–10 min). A statistically significant interaction between time and group was found $F[7.5, 670.7] = 2.2, p = 0.032$. Simple main effects showed that **Pts** group had higher HR than **HM** in every point of time ($p < 0.05$). **HY** group differed significantly from **HM** in speeches ($p = 0.049$) and from **Pts** in recovery 0–10 min ($p = 0.01$) and recovery 20–30 min ($p = 0.030$).

3.1.3. RMSSD

Every group showed lower RMSSD values for stress tasks when compared to habituation (figure 4, table 4). A significant main effect for group ($F = 41.80, p < 0.001$) was found. Post hoc tests showed a significant between-group difference ($p < 0.001$). **HY** having the highest and **Pts** having the lowest RMSSD values.

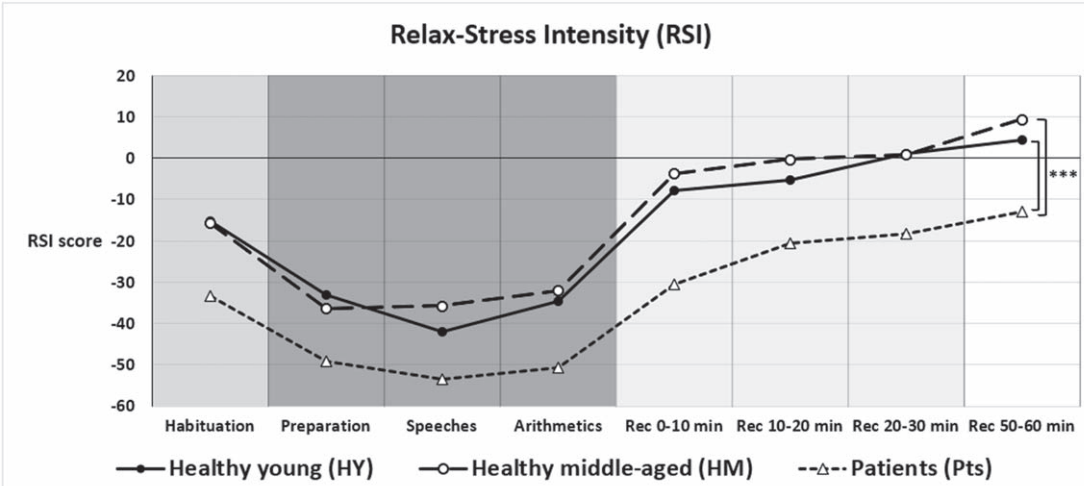
3.1.4. Cortisol

A significant increase in salivary CRT was seen in **Pts** and **HM** (figure 5, table 4) when compared to habituation. Salivary CRT decreased in every group from peak values (Recovery 0–10 min) to Recovery 80–90 min. A

Table 4. Results for each measured variable during the stress test protocol. RSI, HR, and RMSSD results are presented as 10 min averages. Sleep best values describe the highest RSI, RMSSD and the lowest HR measured 10 min average during sleep. Sleep avg. values were calculated as the highest RSI, RMSSD, and the lowest HR average from the whole sleep period during the HRV-collection. CRT measures were collected after each reported 10 min time period.

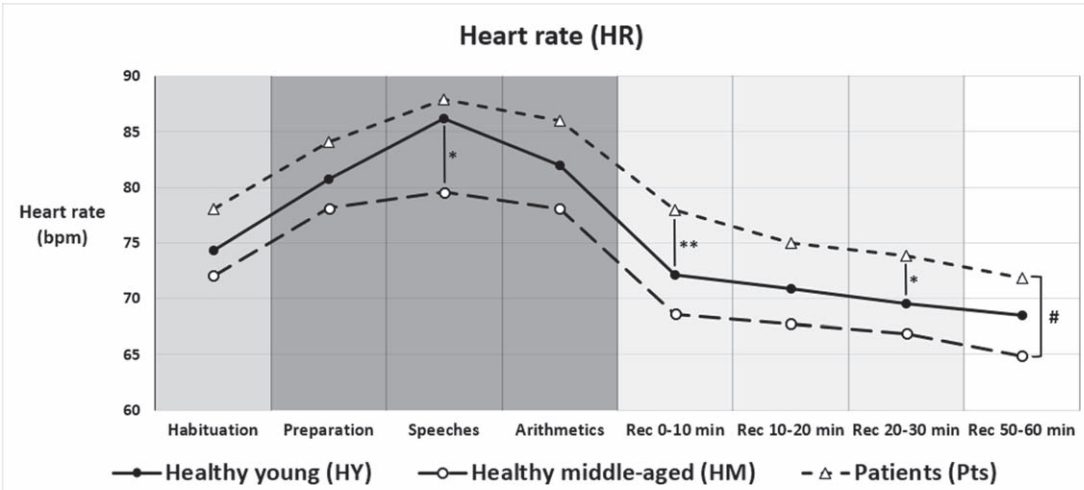
	N	Habituation 20–30 min	Preparation 30–40 min	Speeches 40–50 min	Arithmetics 50–60 min	Recovery 60–70 min	Recovery 70–80 min	Recovery 80–90 min	Recovery 110–120 min	Sleep best	Sleep avg.
HEALTHY YOUNG											
RSI	45	−15 ± 21	−33 ± 24***	−42 ± 23***	−35 ± 24***	−8 ± 22$^{\#\#\#}$	−5 ± 20	1 ± 22	4 ± 23	88 ± 11	54 ± 17
HR (bpm)	57	74 ± 11	81 ± 13***	86 ± 16***	82 ± 15***	72 ± 11$^{\#\#\#}$	71 ± 9	70 ± 10	68 ± 10	42 ± 6	54 ± 7
RMSSD (ms)	57	46 ± 23	39 ± 19	36 ± 18**	40 ± 19	48 ± 22$^{\#\#\#}$	49 ± 22	50 ± 23	54 ± 24	130 ± 54	86 ± 40
PS (0–10)	63	2.0 ± 1.6	3.7 ± 1.9***	5.8 ± 1.9***	5.0 ± 2.1***	1.8 ± 1.4$^{\#\#\#}$	1.1 ± 1.1	0.7 ± 0.8	0.6 ± 0.8		
CRT (nmol l^{-1})	56	16.6 ± 7.9		15.6 ± 6.8	17.8 ± 8.7	18.9 ± 9.4	17.1 ± 7.8	15.7 ± 7.2$^{\#\#\#}$	12.0 ± 5.0		
HEALTHY MIDDLE-AGED											
RSI	44	−16 ± 25	−36 ± 25***	−36 ± 24***	−32 ± 22***	−4 ± 23$^{\#\#\#}$	0 ± 19	1 ± 20	9 ± 25	86 ± 9	45 ± 14
HR (bpm)	56	72 ± 11	78 ± 12***	80 ± 13***	78 ± 11***	69 ± 9$^{\#\#\#}$	68 ± 9	67 ± 9	65 ± 8	46 ± 6	57 ± 7
RMSSD (ms)	54	28 ± 13	25 ± 12***	26 ± 13	27 ± 13	33 ± 13$^{\#\#\#}$	32 ± 13	33 ± 13	34 ± 14	73 ± 33	44 ± 19
PS (0–10)	61	1.4 ± 1.2	3.0 ± 1.8***	4.6 ± 2.3***	5.0 ± 2.1***	1.9 ± 1.6$^{\#\#\#}$	1.1 ± 1.2	0.8 ± 0.9	0.5 ± 0.7		
CRT (nmol l^{-1})	55	12.2 ± 5.3		11.7 ± 6.0	13.9 ± 7.6	16.4 ± 9.4**	15.1 ± 8.0	13.1 ± 6.4$^{\#\#\#}$	9.8 ± 4.5		
PATIENTS											
RSI	53	−33 ± 24	−49 ± 26***	−54 ± 23***	−51 ± 25***	−31 ± 24$^{\#\#\#}$	−21 ± 22	−18 ± 23	−13 ± 24	78 ± 20	35 ± 20
HR (bpm)	69	78 ± 11	84 ± 13***	88 ± 14***	86 ± 13***	78 ± 11$^{\#\#\#}$	75 ± 10	74 ± 10	72 ± 10	49 ± 7	60 ± 8
RMSSD (ms)	64	20 ± 11	17 ± 9**	17 ± 10*	18 ± 10	24 ± 13$^{\#\#\#}$	24 ± 12	24 ± 12	25 ± 12	64 ± 36	37 ± 20
PS (0–10)	72	1.7 ± 1.5	3.0 ± 1.7***	4.9 ± 2.2***	5.6 ± 2.3***	2.0 ± 1.7$^{\#\#\#}$	1.3 ± 1.3	0.9 ± 1.1	0.6 ± 1.0		
CRT (nmol l^{-1})	67	14.9 ± 7.0		13.6 ± 6.1	16.7 ± 8.7	18.5 ± 9.4*	17.8 ± 8.8*	15.5 ± 7.6$^{\#\#\#}$	12.7 ± 6.8		
ALL											
RSI	142	−22 ± 25	−40 ± 26	−44 ± 25	−40 ± 25	−15 ± 26	−9 ± 22	−6 ± 23	−1 ± 26	84 ± 15	44 ± 19
HR (bpm)	182	75 ± 11	81 ± 13	85 ± 15	82 ± 14	73 ± 11	72 ± 10	70 ± 10	69 ± 10	46 ± 7	57 ± 8
RMSSD (ms)	175	31 ± 20	27 ± 17	26 ± 16	28 ± 17	35 ± 19	35 ± 19	35 ± 20	37 ± 21	88 ± 51	55 ± 35
PS (0–10)	196	1.7 ± 1.5	3.2 ± 1.8	5.1 ± 2.2	5.2 ± 2.2	1.9 ± 1.5	1.2 ± 1.2	0.8 ± 0.9	0.5 ± 0.8		
CRT(nmol l^{-1})	178	14.6 ± 7.0		13.7 ± 6.4	16.2 ± 8.5	18 ± 9.4	16.7 ± 8.3	14.8 ± 7.2	11.6 ± 5.7		

Values are presented as mean values ± SD. $^{***}p < 0.001$, $^{**}p < .01$, $^{*}p < .05$ different from habituation period. $^{\#\#\#}p < 0.001$, $^{\#\#}p < 0.01$, $^{\#}p < 0.05$ first point of time to significantly differ from arithmetics period or from peak value in cortisol. Ln transformation was done for HR, RMSSD and CRT variables. Statistical analysis was not performed for ALL. Specific reasons for excluding missing participants are detailed in the supplementary figures S3–S6.



Repeated measures ANOVA - Time	*Mixed model ANOVA - Time*	*Group*	*Time*Group*
HY: F[4.45, 195.92] = 57.56, p < .001	F[4.23, 587.22] = 198.54, p < .001	F[2, 139] = 16.15, p < .001	F[8.45, 587.22] = 1.44, p = .173
HM: F[3.86, 165.90] = 73.69, p < .001			
Pts: F[3.68, 191.27] = 70.50, p < .001			

Figure 2. RSI (10 min averages) of different groups during the stress protocol. Pts had significantly lower RSI scores compared to HM and HY. $^{***}p < 0.001$. All ANOVA results are Greenhouse–Geisser corrected.



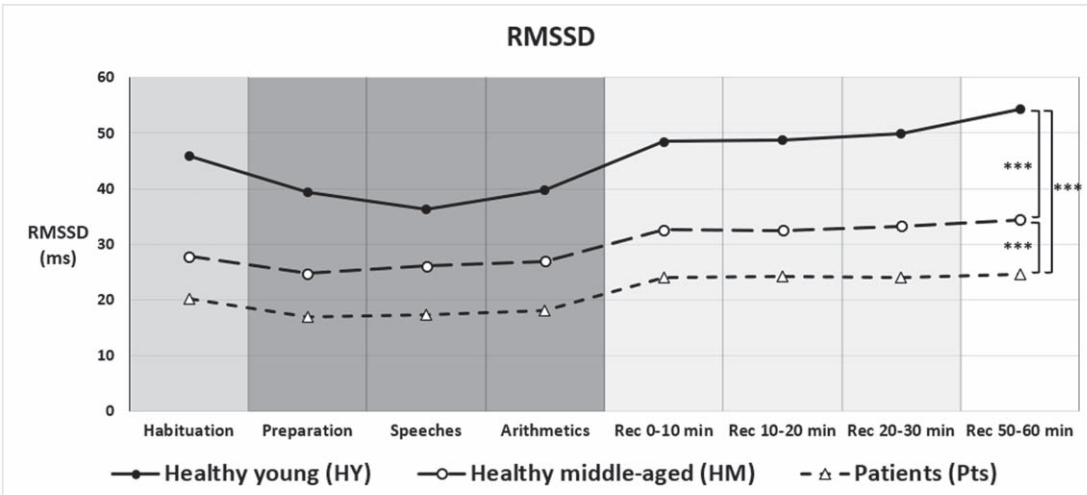
Repeated measures ANOVA - Time	*Mixed model ANOVA - Time*	*Group*	*Time*Group*
HY: F[3.87, 216.53] = 89.54, p < .001	F[3.75, 670.70] = 337.21, p < .001	F[2, 179] = 7.74, p = .001	F[7.49, 670.70] = 2.16, p = .032
HM: F[3.38, 186.04] = 129.11, p < .001			
Pts: F[3.44, 233.59] = 133.55, p < .001			

Figure 3. HR responses (10 min averages) during TSST-G stress protocol for different groups. A significant interaction between time and group was found. Tests for simple main effects showed that Pts group had higher HR than HM in every time point (# Post hoc $p < 0.05$). HY group differed significantly from HM in Speeches ($^{*}p = 0.049$) and from Pts in Recovery 0–10 min ($^{**}p = 0.01$) and in Recovery 20–30 min ($^{*}p = 0.030$). Statistical analysis was done with Ln transformed values and all ANOVA results are Greenhouse–Geisser corrected.

significant main effect for group was observed ($F = 3.62, p = 0.029$). **HM** group had lower CRT than **HY** (Post hoc $p = 0.039$).

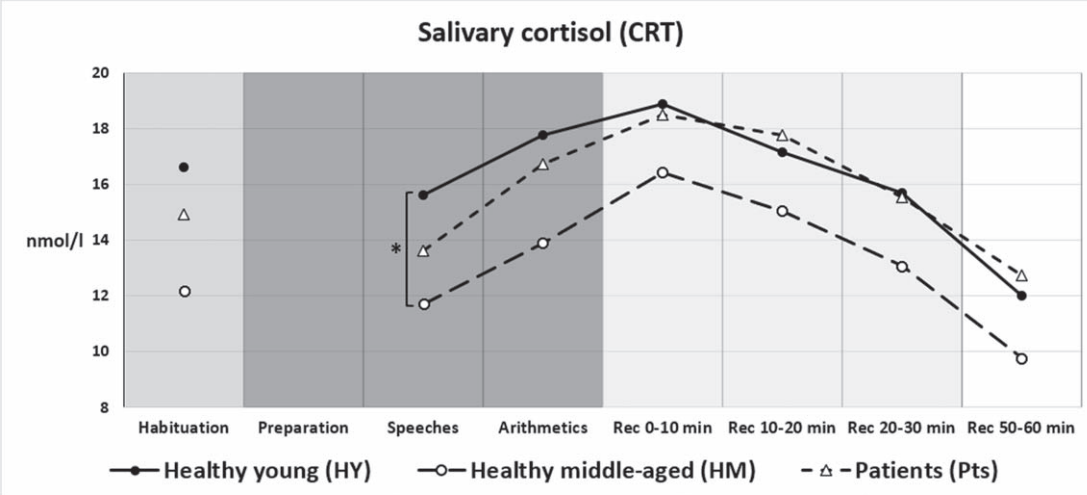
3.1.5. Perceived stress

PS increased from habituation to preparation, speeches and arithmetics (figure 6) in every group (Post hoc $p < 0.001$). PS decreased from arithmetics to recovery 0–10 min in all groups (Post hoc $p < 0.001$). A statistically significant interaction between time and group was observed: $F[6.37, 614.89] = 4.43, p < 0.001$.



Repeated measures ANOVA - Time	*Mixed model ANOVA - Time*	*Group*	*Time*Group*
HY: F[2.89, 161.73] = 18.20, p < .001	F[3.59, 616.82] = 79.82, p < .001	F[2, 172] = 41.80, p < .001	F[7.17, 616.82] = 1.11, p = .352
HM: F[3.54, 187.57] = 31.17, p < .001			
Pts: F[3.56, 224.51] = 37.26, p < .001			

Figure 4. RMSSD responses (10 min averages) during TSST-G stress protocol for different groups. A significant main effect for group was found and post hoc tests showed between-group differences in RMSSD ($^{***}p < 0.001$). Statistical analysis was done with Ln transformed values and all ANOVA results are Greenhouse–Geisser corrected.



Repeated measures ANOVA - Time	*Mixed model ANOVA - Time*	*Group*	*Time*Group*
HY: F[3.35, 184.01] = 16.46, p < .001	F[3.53, 617.52] = 47.57, p < .001	F[2, 175] = 3.62, p = .029	F[7.06, 617.52] = 1.29, p = .252
HM: F[3.70, 199.70] = 23.80, p < .001			
Pts: F[3.70, 243.89] = 15.74, p < .001			

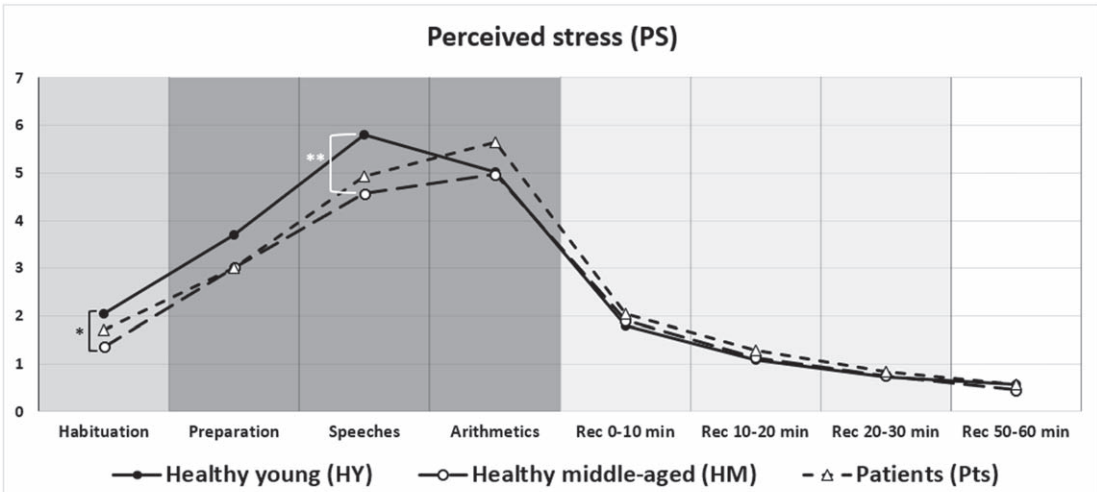
Figure 5. Salivary CRT responses were measured during TSST-G stress protocol for different groups. A significant main effect for group was observed ($F = 3.62, p = 0.029$). HM group had a lower CRT than HY group (Post hoc $p = 0.039$). $^{*}p < 0.05$. Statistical analysis was done with Ln transformed values and all ANOVA results are Greenhouse–Geisser corrected.

Simple main effects showed **HY** group having greater PS than **HM** during the habituation period ($p = 0.035$) and speeches ($p = 0.006$).

3.2. Correlations of HR, RMSSD and RSI to PS and cortisol

3.2.1. Perceived stress

Change in PS from habituation to stress task (table 5) correlated negatively with RSI and positively with HR. Changes in PS from arithmetics to recovery period correlated negatively with RSI and RMSSD and positively with HR.



Repeated measures ANOVA - Time	*Mixed model ANOVA - Time*	*Group*	*Time*Group*
HY: F[3.77, 233.67] = 199.28, p < .001	F[3.19, 614.89] = 517.73, p < .001	F[2, 193] = 1.22, p = .299	F[6.37, 614.89] = 4.43, p < .001
HM: F[2.76, 165.46] = 150.42, p < .001			
Pts: F[2.76, 196.04] = 181.41, p < .001			

Figure 6. PS (0–10 scale) was measured during TSST-G stress protocol for different groups. A statistically significant interaction between time and group ($F = 4.4, p < 0.001$) was found. HY group reported significantly greater PS during the habituation period ($^{*}p = 0.035$) and speeches ($^{**}p = 0.006$) than HM. All ANOVA results are Greenhouse–Geisser corrected.

Table 5. Correlations of HR, RMSSD and RSI to perceived stress.

	Group	ΔRSI	ΔRMSSD	ΔHR
Change in perceived stress (habituation to stress tasks)	All	**−0.157**[a]	−0.085	**0.165**[a]
	Sig. (2-tailed)	**0.045**	0.246	0.025
	N=	**163**	186	186
Change in perceived stress (arithmetics to recovery)	All	**−0.245**[b]	**−0.213**[b]	**0.268**[b]
	Sig. (2-tailed)	**0.001**	**0.004**	**<0.001**
	N=	172	185	185

All correlations were determined with Spearman's rho.

[a] Correlation is significant at the .05 level (2-tailed). Change in perceived stress refers to the largest observed change in PS.

[b] Correlation is significant at the .01 level (2-tailed).

3.2.2. Total cortisol

CRT_{AUCg} correlated negatively with RSI_{Total} in every group (table 6). **HM** groups' CRT_{AUCg} responses showed a statistically significant correlation with every stress variable. In the **Pts** group CRT_{AUCg} correlated with RSI_{Total}, $RMSSD_{Total}$, and HR_{Total}. **HY** groups' CRT_{AUCg} correlated with RSI_{Total} and PS_{Total}.

3.3. Multiple linear regression to determine factors affecting RSI

Multiple linear regression was calculated to predict the total stress load estimated by the commercial stress meter (RSI_{Total}). The prediction was done using $RMSSD_{Total}$, HR_{Total}, CRT_{AUCg}, and metabolic syndrome risk factor status (MetS risk factors) as predicting variables. Before this age, sex, PS_{Total}, $V\dot{O}_{2peak}$, and BMI were excluded from the model. A significant regression equation was found ($F[4\ 141] = 25.32, p < 0.001$), with R^2 of 0.418 (table 7, figure 7).

4. Discussion

4.1. Responses to psychosocial stress

The results of this study demonstrate that while TSST-G induced similar psychological responses in all groups, the physiological stress responses to the induced psychosocial stress differed along with health status and age. In addition, the results suggest that if stress in general is defined as PS or CRT responses, RSI, which is an individually adaptive variable, is a better predictor of stress than the commonly used RMSSD alone.

Table 6. Correlations of PS, HR, RMSSD and RSI to total cortisol response.

	Group	RSI_{Total}	$RMSSD_{Total}$	HR_{Total}	PS_{Total}
Cortisol response (AUCg)	Healthy young	**−0.349**[a]	0.055	0.076	0.279[a]
	Sig. (2-tailed)	**0.019**	0.692	0.587	0.034
	N=	45	54	54	58
	Healthy middle-aged	**−0.456**[b]	**−0.372**[b]	**0.496**[b]	**0.390**[b]
	Sig. (2-tailed)	**0.001**	**0.005**	**<0.001**	**0.002**
	N=	49	56	56	60
	Patients	**−0.365**[b]	**−0.368**[b]	**0.476**[b]	**0.028**
	Sig. (2-tailed)	**0.004**	**0.002**	**<0.001**	**0.816**
	N=	61	68	68	71
	All	**−0.408**[b]	**−0.230**[b]	**0.364**[b]	**0.233**[b]
	Sig. (2-tailed)	**<0.001**	**0.002**	**0.001**	**0.001**
	N=	155	178	178	189

All correlations were determined with Spearman's rho.

[a] Correlation is significant at the .05 level (2-tailed). RSI, RMSSD, HR, PS present a total response measured in a certain variable during the stress protocol (See methods).

[b] Correlation is significant at the .01 level (2-tailed).

The overall responses observed in PS were similar in every group. Besides higher PS reported by HY during the early parts of the stress protocol, no further between-group differences were found. These results suggest that TSST-G protocol produced similar levels of psychological stress in the examined groups. Based on the reported PS values and physiological measures, it seems evident that the participants were experiencing stress already during the habituation period. Therefore the measures during habituation may not reflect a completely stress-free or neutral state in the body, and a more reliable baseline or a reference level of the measured stress responses may reside at the end of the recovery period.

The observed HR, HRV, and CRT responses to TSST-G can be regarded as normal physiological stress responses in all groups. Between-group differences occurred in HR, HRV, and CRT responses. Differences seen in HR between healthy groups (HY and HM) support previous findings of an age-related decrease in HR reactivity (Kudielka *et al* 2004a, Wrzus *et al* 2014). Our results also show a trend of higher HR in HY compared to HM, which could again relate to higher stress reported by HY in the early parts of the protocol; however, this trend may also be partly due to a between-group difference in sex distribution as the proportion of females was slightly higher in HY (71%) compared to HM (53%). Because differences in HR and HRV also occurred between HM and the patients with similar age profiles, the differences do not seem to relate only to the effects of age but also to the effect of health status. These trends can also be observed in the ECG data measured during nocturnal sleep (table 4). Indeed, cardiometabolic risk factors have been reported to affect autonomic balance by shifting it to the direction of sympathetic dominance (Liao *et al* 1998, Koskinen *et al* 2009).

It has been previously shown that age increases the CRT response to different challenges (Kudielka *et al* 2004b, Otte *et al* 2005). In this study, HY showed the highest total CRT response (AUCg) to the stress protocol. This conflicts with the findings of Otte *et al* (2005) and might be due to higher PS reported by HY in the early parts of the protocol. This difference in CRT between HY and HM seems to occur because CRT values of young females are clearly higher than those in the middle-aged (supplementary figure S6), which is not due to unequal sex distribution between HY and HM. Instead, it is more likely caused by higher PS levels experienced by younger females (supplementary tables S4–S5) and/or higher physiological reactivity to stress. Although not statistically significant, the patients showed a trend for higher total CRT responses when compared to HM, again supporting the hypothesis of sympathetic dominance related to chronic health conditions.

4.2. HRV-based stress index (RSI) in measuring stress

The HRV-based stress index (RSI) examined in this study is calculated based on ECG, accelerometer, and background data with a potential advantage of providing a more overall view of physiological reactions than using just a single variable like HR, RMSSD, or some other HR-based variable. The results of the present study show that RSI reacts to physiological changes related to psychosocial stress and recovery. Also, the trends observed in RSI are as expected in relation to HR, RMSSD, and PS. Responses seen in frequency domain HRV parameters (See supplementary materials 2.3) are partly conflicting but HF and total power seem concordant with both RSI and previous literature, suggesting HF and total power to reflect active modulation of vagal activity (Shaffer and Ginsberg 2017). Based on our results, LF/HF ratio does not seem to act as a proper indicator of sympatho-vagal balance in a psychosocial stress situation. This is likely because LF power (figure S7),

Table 7. Results from multiple linear regressions were calculated including all participants. Different stress variables were used to predict the total RSI response. HR_{Total} was calculated as beats per minute, $RMSSD_{Total}$ was calculated as milliseconds, CRT_{AUCg} was calculated in nmol l^{-1}, and metabolic syndrome risk factors as a number indicating the amount risk factors.

Model	R^2	Adjusted R^2	ΔR^2	ΔF	ΔP	Standardized β	t	P
$DV = 21.891 - 0.5691_1 + 0.4682_2 - 0.4093_3 - 14.128_4$	0.418	0.402	0.418	25.322	<0.001			
Predictor variables:								
Heart rate $total_1$						−0.214	−2.711	0.008
RMSSD $total_2$						0.337	4.425	<0.001
Cortisol $AUCg_3$						−0.208	−3.022	0.003
MetS risk $factors_4$						−0.274	−4.068	<0.001

*Model: $F = 25.322$; df = 4 141; $p < 0.001$.

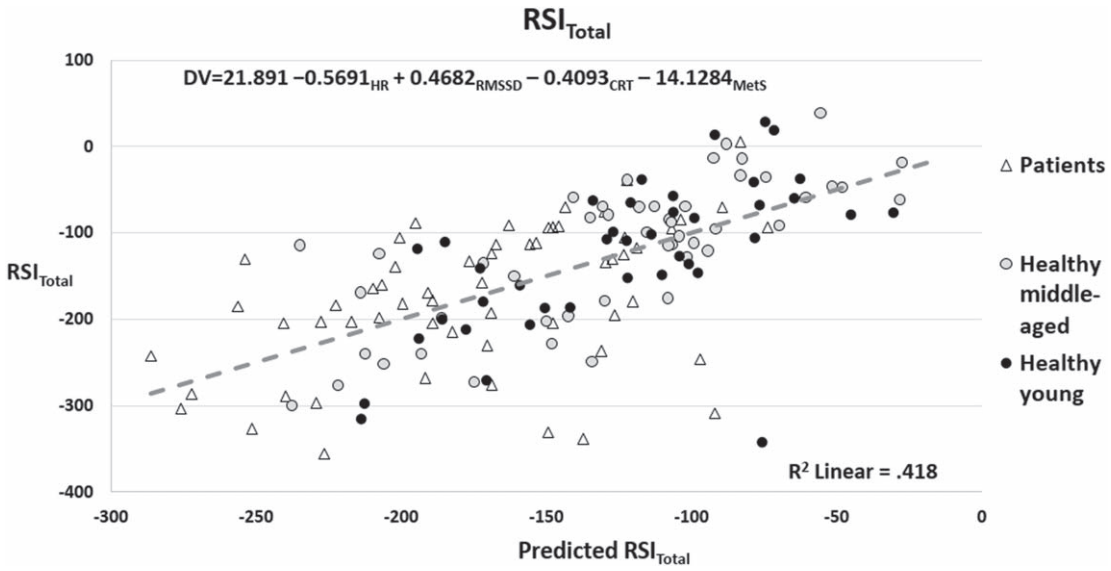


Figure 7. Visualization of the RSI_{Total} calculated and RSI_{Total} predicted using the multiple linear regression model determining the RSI score. RSI_{Total} was calculated as a sum of the average values measured in 10 min periods measured during habituation and stress tasks (40 min).

especially in our experimental setup involving short-term measures and performed at seated rest, is not solely presenting sympathetic activity but also parasympathetic activity (Shaffer *et al* 2014, Shaffer and Ginsberg 2017).

Despite the small differences in the PS scores, the healthy groups seemed to experience similar amounts of psychological stress. Even though considerable between-group differences were observed in HR and especially RMSSD, the RSI results, which estimate the intensity of the stress response based on physiology, were very similar in the healthy groups regardless of age. This is something one might expect when evaluating the capability of RSI from the perspective of psychological stress and implies that when quantifying stress, the RSI calculation successfully takes into account the individual background information (e.g. age, sex, physical fitness) that affect the physiological responses (for example O'Brien *et al* 1986, Umetani *et al* 1998). These notions are also supported by the results of multiple linear regression analyses, indicating that age, sex, $V\dot{O}_{2peak}$, and BMI did not have predicting value in the RSI model. Further, these notions are supported by somewhat similar RSI levels during sleep in the groups, whereas nocturnal recordings revealed clear differences seen in HR and especially RMSSD values. However, although similar levels of PS were observed also in the patients, their RSI score was significantly more negative compared to the healthy groups. In other words, similar amounts of reported PS led to more negative stress level estimation in the patients. It should also be noted that patients' RSI scores remained negative in the recovery period, although reported PS was minimal. Whether this is the desired outcome or not, depends on what components of stress (physiological or psychological) one is trying to measure and put emphasis on.

4.3. RSI and alternative methods in detecting stress

When trying to evaluate the potential of RSI in quantifying psychological stress and the HPA axis activity, RSI and other commonly known physiological measures were compared with each other. A total stress response for each stress variable including RSI, HR, RMSSD, PS, and CRT was calculated. Total response for CRT was calculated using AUCg (Pruessner *et al* 2003). This approach was applied also for other stress variables including RSI, HR, RMSSD, and PS to calculate the total response detected by each stress variable. This approach was chosen because stress reactions are already seen in the habituation period and therefore it does not serve as a good reference point of activity. Instead of studying peak responses occuring in various timeframes in comparison to a certain reference point, studying total responses induced by the protocol would be less affected by the chosen reference point. Also, reviewing the stressful situation as a whole might provide more meaningful results since measuring stress in real life would probably focus on estimating stress in situations lasting longer than 10 min.

When the association of the overall cortisol response (CRT_{AUCg}) was compared to different stress variables, RSI_{Total} showed the highest correlations (small to medium) with CRT_{AUCg}. However, when evaluating the association of PS to different physiological stress variables HR showed the greatest correlations (small) with PS. This finding is logical when considering easily perceived bodily signals such as elevated HR as an important factor in stress perception (Schultz and Vögele 2015).

When evaluating the factors affecting RSI, a multiple regression analysis of the RSI_{Total} score was calculated. Age, sex, and $V\dot{O}_{2peak}$ did not have predicting value in the model suggesting that RSI calculation accounts for these parameters successfully. CRT_{AUCg} and the amount of metabolic syndrome components predicted the RSI_{Total} score together with HR_{Total} and $RMSSD_{Total}$, suggesting the ability of RSI to observe physiological adjustments on a more overall scale.

The finding that metabolic syndrome risk factors serve as a predictor of the RSI_{Total} score, together with the findings showing more negative RSI results in the patients, indicate the RSI calculation to be sensitive to changes occurring in one's physiology. This finding is logical since metabolic syndrome has been associated with changes in HRV dynamics (Liao *et al* 1998, Koskinen *et al* 2009). However, a more negative 'baseline' is a finding that one should be aware of when interpreting the results of the RSI stress index. On one hand, although not serving as a clinical tool, these findings raise the question of whether RSI could differentiate 'healthy' persons from persons having cardiometabolic diseases or in risk developing such diseases. On the other hand, the results imply that when interpreting the RSI results within a shorter time frame, such as within a single workday, individuals with metabolic syndrome components will likely receive more negative estimations of the intensity of stress compared to healthy individuals. In other words, although RSI seems to be a reactive measure for changes occurring within a short time frame even for individuals with metabolic syndrome components, the feedback is likely to be more negative in individuals with metabolic syndrome components compared to healthy individuals, since the RSI baseline is interpreted as more negative in the former group. This is a relevant finding since the prevalence of cardiometabolic risk factors is substantial in the adult population (Scuteri *et al* 2014) and therefore affects the interpretation of the results of several potential users.

Problems in defining the correct baseline do not concern only RSI but also other stress meters. In practice, all HR-based stress meters share similar problems in making the distinction between stress and recovery. Measuring stress by interpreting bodily signals requires a lot of information since physiological responses like HR and HRV are known to vary due to individual factors like age (O'Brien *et al* 1986), sex (Umetani *et al* 1998), and health status (Koskinen *et al* 2009, Assoumou *et al* 2010) but also depend on body posture (Tulen *et al* 1999) and recovery state (Mourot *et al* 2004). In this study, these differences were seen in all physiological responses measured during the stress tasks but similar trends were also seen in the nocturnal data. For example, absolute RMSSD values reported in the present study were around 50% greater in HY than in HM. However, our results suggest that the RSI calculation can take into account at least some of these factors when estimating stress responses.

It has also been questioned whether using linear methods (e.g. HR and RMSSD) to analyze HRV-based data are sufficient to quantify complex changes in the state of the body (Schubert *et al* 2009). Indeed, based on correlation and regression analyses (tables 6, S2–S3) with none to moderate associations, the ability of any single variable to explain stress responses seems limited and implies that a more overall approach, including information synthesized from multiple measures, would be preferable. For instance, taking the effects of respiration (Hernando *et al* 2016) into account could provide further meaningful insights to define the physiological state of an individual. Therefore, while it seems that stress responses can be quantified with relatively straightforward analysis methods used in this study, interpreting the results is likely challenging and offers only rough estimates. Also, determining the origin of the stress response (i.e. physical versus psychological stress) is a substantial challenge when measuring stress in real life. When considering these matters, compared to HR and RMSSD, RSI offers an easier way to measure physiological stress responses in persons of different age and sex.

5. Limitations of the study

This study has its limitations. Only linear methods were used to quantify HRV responses; thus, the analyses may not optimally cover all aspects of ANS processes such as the unpredictability and complexity of a series of R–R intervals (Shaffer and Ginsberg 2017). In addition, the functioning of RSI was evaluated by comparing it to other physiological and psychological measures, while the exact RSI calculation process was out of the scope of this study (See supplementary materials 1.1). Although each of the three groups included more females than males, the sex distributions were not perfectly balanced in the groups (HY: 71% females, HM: 53% females, Pts: 77% females), which may have affected some findings on HRV (Voss *et al* 2015) and/or CRT (Kudielka *et al* 2004b). In the healthy groups (HY and HM), cardiorespiratory fitness was estimated indirectly with a commercial $V\dot{O}_{2peak}$ estimation algorithm and health status was screened with a phone interview and questionnaires. In addition, the habituation period in the present study does not reflect a relaxed baseline value for evaluating stress responses; a controlled baseline measure with no psychological stress, controlled body posture, and time of the day should be considered in the future studies. Even though the PS responses were very similar in the groups, the

timely amount of psychosocial stress experienced by a single participant might have varied due to different amount of participants in each session.

6. Conclusions

The present study used various methods, including HR, RMSSD, PS, CRT, and an HRV-based stress index (RSI), to measure physiological stress reactions during the standardized psychosocial stress test (TSST-G) in different age and health groups. Psychosocial stress induced significant physiological and psychological responses in all groups. Despite similar responses in PS, the groups showed different physiological responses. Of all the stress variables studied, the examined HRV-based stress index was most consistently associated with physiological stress responses. Meanwhile, HR followed the trends of PS, which might be due to its role as an important internal bodily signal of stress.

Overall, several approaches may be successfully used to quantify physiological responses to psychosocial stress. However, to determine the status of stress versus recovery, the HRV-based stress index used in the present study requires less prior information (e.g. baseline levels of different measures) from the user than other HR-based measurements. Therefore, the HRV-based stress index seems to be not only valid but also easy-to-apply method to measure physiological stress responses. However, the presence of cardiometabolic risk factors affects the stress estimation of the used commercial product and should therefore be acknowledged when interpreting the stress index scores.

Acknowledgments

The HealthBeat project was implemented in cooperation between the University of Jyväskylä (Department of Psychology and Faculty of Sport and Health Sciences), Central Finland Health Care District, and Firstbeat Technologies Ltd. The study was funded by Business Finland (Grant 2697/31/2018) and Firstbeat Technologies Ltd (Jyväskylä, Finland). We thank Valtteri Huttunen (Faculty of Sport and Health Sciences, JYU) for his notable assistance in conducting TSST-G protocol and Jukka Hintikka (Faculty of Sport and Health Sciences, JYU) for precise work in analyzing the salivary cortisol samples. We also want to thank Tero Myllymäki (Firstbeat Technologies Ltd) and Ilkka Korhonen (Firstbeat Technologies Ltd) for organizing the resources needed for implementing the study and Tomas Snellman (Firstbeat Technologies Ltd) and Janne Solanpää (Firstbeat Technologies Ltd) for gathering and organizing the HRV data for the use of researchers.

Conflict of interest

Firstbeat Technologies Ltd (Jyväskylä, Finland) provided funding as well as technical and equipment support for conducting the study. A family member of M R is an employee and stockowner in Firstbeat Technologies Ltd (Jyväskylä, Finland). A T was employed by Firstbeat Analytics Ltd during the preparation of the manuscript. This occurred after data analysis and interpretation. Other authors declare no conflict of interest. Conclusions made reflect the authors' opinions, not affected by personal interests.

Author contributions

M R, A-P E R, J L O K, U M K, J A L and J W conceived the idea to conduct this study and planned the study protocol; S S, A T, J J, M R, A-P E R, and J L O K performed the data collection; S S, A T, J J, J L O K, and J W conducted the data analysis; S S, A T, J J, A-P E R, U M K, and J W interpreted the data; S S prepared the draft of the manuscript; S S, A T, J J, M R, U M K, A-P E R, J L O K, and J W participated in the revision of the manuscript. All authors accepted the final version.

ORCID iDs

Santtu M Seipäjärvi https://orcid.org/0000-0003-0891-2453
Anniina Tuomola https://orcid.org/0000-0003-2508-9462
Joona Juurakko https://orcid.org/0000-0002-1171-6739
Mirva Rottensteiner https://orcid.org/0000-0003-1988-0952
Antti-Pekka E Rissanen https://orcid.org/0000-0003-2415-3979
Jari L O Kurkela https://orcid.org/0000-0002-4437-6347

Urho M Kujala https://orcid.org/0000-0002-9262-1992
Jari A Laukkanen https://orcid.org/0000-0002-3738-1586
Jan Wikgren https://orcid.org/0000-0002-3403-1396

References

Alberti K G M M, Zimmet P and Shaw J 2006 Metabolic syndrome—a new world-wide definition. A consensus statement from the international diabetes federation *Diabetic Med.* **23** 469–80
American Thoracic Society & American College of Chest Physicians 2003 ATS/ACCP statement on cardiopulmonary exercise testing *Am. J. Respiratory Crit. Care Med.* **167** 211–77
Assoumou H G N, Pichot V, Barthelemy J C, Dauphinot V, Celle S, Gosse P, Kossovsky M, Gaspoz J M and Roche F 2010 Metabolic syndrome and short-term and long-term heart rate variability in elderly free of clinical cardiovascular disease: the PROOF study *Rejuvenation Res.* **13** 653–63
Boomsma D I, van Baal G C and Orlebeke J F 1990 Genetic influences on respiratory sinus arrhythmia across different task conditions *Acta Geneticae Med. Et Gemellologiae* **39** 181–91
Bozovic D, Racic M and Ivkovic N 2013 Salivary cortisol levels as a biological marker of stress reaction *Medicinski Arhiv* **67** 33774–7
Buchheit M 2014 Monitoring training status with HR measures: do all roads lead to Rome? *Front. Physiol.* **5** 73
Campbell J and Ehlert U 2012 Acute psychosocial stress: does the emotional stress response correspond with physiological responses? *Psychoneuroendocrinology* **37** 1111–34
Chrousos G P 2009 Stress and disorders of the stress system *Nat. Rev. Endocrinol.* **5** 374–81
Cohen S, Gianaros P J and Manuck S B 2016 A stage model of stress and disease *Perspect. Psychol. Sci.* **11** 456–63
Cole C R, Blackstone E H, Pashkow F J, Snader C E and Lauer M S 1999 Heart-rate recovery immediately after exercise as a predictor of mortality *New Engl. J. Med.* **341** 1351–7
de Rooij S R 2013 Blunted cardiovascular and cortisol reactivity to acute psychological stress: a summary of results from the dutch famine birth cohort study *Int. J. Psychophysiol.* **90** 21–7
Firstbeat technologies Ltd 2005 VO_2 estimation method based on heart rate measurement (https://firstbeat.com/en/oxygen-consumption-estimation-firstbeat-white-paper/)
Firstbeat technologies Ltd 2014a Automated fitness level (VO_2max) estimation with heart rate and speed data (https://firstbeat.com/fi/aerobic-fitness-level-vo%E2%82%82max-estimation-firstbeat-white-paper/)
Firstbeat technologies Ltd 2014b Stress and recovery analysis method based on 24 h heart rate variability (https://firstbeat.com/en/stress-recovery-analysis-method-based-24-hour-heart-rate-variability-firstbeat-white-paper-2/)
Gaab J, Rohleder N, Nater U M and Ehlert U 2005 Psychological determinants of the cortisol stress response: the role of anticipatory cognitive appraisal *Psychoneuroendocrinology* **30** 599–610
Hassard J, Teoh K R H, Visockaite G, Dewe P and Cox T 2018 The cost of work-related stress to society: a systematic review *J. Occup. Health Psychol.* **23** 1–17
Hernando A, Lazaro J, Gil E, Arza A, Garzon J M, Lopez-Anton R, de la Camara C, Laguna P, Aguilo J and Bailon R 2016 Inclusion of respiratory frequency information in heart rate variability analysis for stress assessment *IEEE J. Biomed. Health Inform.* **20** 1016–25
Kirschbaum C, Bartussek D and Strasburger C J 1992 Cortisol responses to psychological stress and correlations with personality traits *Personality Individual Differences* **13** 1353–57
Kivimäki M and Steptoe A 2018 Effects of stress on the development and progression of cardiovascular disease *Nat. Rev. Cardiol.* **15** 215–29
Klaperski S, von Dawans B, Heinrichs M and Fuchs R 2014 Effects of a 12 Week endurance training program on the physiological response to psychosocial stress in men: a randomized controlled trial *J. Behav. Med.* **37** 1118–33
Koskinen T, Kähönen M, Jula A, Mattsson N, Laitinen T, Keltikangas-Järvinen L, Viikari J, Välimäki I, Rönnemaa T and Raitakari O T 2009 Metabolic syndrome and short-term heart rate variability in young adults *Diabetic Med.* **26** 354–61
Kudielka B M, Buske-Kirschbaum A, Hellhammer D H and Kirschbaum C 2004a Differential heart rate reactivity and recovery after psychosocial stress (TSST) in healthy children, younger adults, and elderly adults: the impact of age and gender *Int. J. Behav. Med.* **11** 116–21
Kudielka B M, Buske-Kirschbaum A, Hellhammer D H and Kirschbaum C 2004b HPA axis responses to laboratory psychosocial stress in healthy elderly adults, younger adults, and children: Impact of age and gender *Psychoneuroendocrinology* **29** 83–98
Kudielka B M and Kirschbaum C 2003 Awakening cortisol responses are influenced by health status and awakening time but not by menstrual cycle phase *Psychoneuroendocrinology* **28** 35–47
Kupper N H M, Willemsen G, van den Berg M, de Boer D, Posthuma D, Boomsma D I and de Geus Eco J 2004 Heritability of ambulatory heart rate variability *Circulation* **110** 2792–6
Lazarus R S and Folkman S 1984 *Stress, Appraisal, and Coping* (New York, NY: Springer)
Liao D, Sloan R P, Cascio W E, Folsom A R, Liese A D, Evans G W, Cai J and Sharrett A R 1998 Multiple metabolic syndrome is associated with lower heart rate variability. The atherosclerosis risk in communities study *Diabetes Care* **21** 2116–22
Madsen I E H *et al* 2017 Job strain as a risk factor for clinical depression: systematic review and meta-analysis with additional individual participant data *Psychol. Med.* **47** 1342–56
Melillo P, Bracale M and Pecchia L 2011 Nonlinear heart rate variability features for real-life stress detection. Case study: students under stress due to university examination *Biomed. Eng. Online* **10** 96
Mezzacappa E S, Kelsey R M, Katkin E S and Sloan R P 2001 Vagal rebound and recovery from psychological stress *Psychosomatic Med.* **63** 650–7
Mourot L, Bouhaddi M, Tordi N, Rouillon J and Regnard J 2004 Short- and long-term effects of a single bout of exercise on heart rate variability: comparison between constant and interval training exercises *Eur. J. Appl. Physiol.* **92** 508–17
Mücke M, Ludyga S, Colledge F and Gerber M 2018 Influence of regular physical activity and fitness on stress reactivity as measured with the trier social stress test protocol: a systematic review *Sports Med. (Auckland)* **48** 2607–22
Nyberg S *et al* 2014 Job strain as a risk factor for type 2 diabetes: a pooled analysis of 124 808 men and women *Diabetes Care* **37** 2268–75
O'Brien I A, O'Hare P and Corrall R J 1986 Heart rate variability in healthy subjects: effect of age and the derivation of normal ranges for tests of autonomic function *Br. Heart J.* **55** 348–54
Otte C, Hart S, Neylan T C, Marmar C R, Yaffe K and Mohr D C 2005 A meta-analysis of cortisol response to challenge in human aging: importance of gender *Psychoneuroendocrinology* **30** 80–91
Padgett D A and Glaser R 2003 How stress influences the immune response *Trends Immunol.* **24** 444–8

Pruessner J C, Kirschbaum C, Meinlschmid G and Hellhammer D H 2003 Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change *Psychoneuroendocrinology* **28** 916–31
Rimmele U, Zellweger B C, Marti B, Seiler R, Mohiyeddini C, Ehlert U and Heinrichs M 2007 Trained men show lower cortisol, heart rate and psychological responses to psychosocial stress compared with untrained men *Psychoneuroendocrinology* **32** 627–35
Robertson A, King K, Ritchie S, Gauthier A, Laurence M and Dorman S 2015 Validating the use of heart rate variability for estimating energy expenditure *Int. J. Hum. Mov. Sports Sci.* **3** 19–26
Schipke J D, Arnold G and Pelzer M 1999 Effect of respiration rate on short-term heart rate variability *J. Clin. Basic Cardiol.* **2** 92–5
Schlotz W, Kumsta R, Layes I, Entringer S, Jones A and Wüst S 2008 Covariance between psychological and endocrine responses to pharmacological challenge and psychosocial stress: a question of timing *Psychosomatic Med.* **70** 787–96
Schubert C, Lambertz M, Nelesen R A, Bardwell W, Choi J and Dimsdale J E 2009 Effects of stress on heart rate complexity—A comparison between short-term and chronic stress *Biol. psychol.* **80** 3325–32
Schulz A and Vögele C 2015 Interoception and stress *Front. Psychol.* **6** 993
Scuteri A *et al* 2014 Metabolic syndrome across Europe: different clusters of risk factors *Eur. J. Preventive Cardiol.* **22** 486–91
Shaffer F and Ginsberg J P 2017 An overview of heart rate variability metrics and norms *Front. Public Health* **5** 258
Shaffer F, McCraty R and Zerr C L 2014 A healthy heart is not a metronome: an integrative review of the heart's anatomy and heart rate variability *Front. Psychol.* **5**
Smolander J, Ajoviita M, Juuti T, Nummela A and Rusko H 2011 Estimating oxygen consumption from heart rate and heart rate variability without individual calibration *Clin. Physiol. Funct. Imaging* **31** 266–71
Smolander J, Juuti T, Kinnunen M-L, Laine K, Louhevaara V, Männikkö K and Rusko H 2008 A new heart rate variability-based method for the estimation of oxygen consumption without individual laboratory calibration: application example on postal workers *Appl. Ergon.* **39** 325–31
Taelman J, Vandeput S, Spaepen A and Van Huffel S 2009 Influence of mental stress on heart rate and heart rate variability *4th European Conference of the International Federation for Medical and Biological Engineering* ed J Vander Sloten *et al* 22 (Berlin, Heidelberg: IFMBE Proceedings) (https://doi.org/10.1007/978-3-540-89208-3_324)
Thayer J F, Yamamoto S S and Brosschot J F 2009 The relationship of autonomic imbalance, heart rate variability and cardiovascular disease risk factors *Int. J. Cardiol.* **141** 122–31
Tukey J W 1977 *Exploratory Data Analysis* (Reading: Addison-Wesley)
Tulen J H, Boomsma F and Man in 't Veld A J 1999 Cardiovascular control and plasma catecholamines during rest and mental stress: effects of posture *Clin. Sci.* **96** 567–76
Umetani K, Singer D H, McCraty R and Atkinson M 1998 Twenty-four hour time domain heart rate variability and heart rate: relations to age and gender over nine decades *J. Am. Coll. Cardiol.* **31** 593–601
von Dawans B, Kirschbaum C and Heinrichs M 2011 The trier social stress test for groups (TSST-G): a new research tool for controlled simultaneous social stress exposure in a group format *Psychoneuroendocrinology* **36** 514–22
Voss A, Schroeder R, Heitmann A, Peters A and Perz S 2015 Short-term heart rate variability—influence of gender and age in healthy subjects *PLoS One* **10** e0118308
Weber C S, Thayer J F, Rudat M, Wirtz P H, Zimmermann-Viehoff F, Thomas A, Perschel F H, Arck P C and Deter H C 2010 Low vagal tone is associated with impaired post stress recovery of cardiovascular, endocrine, and immune markers *Eur. J. Appl. Physiol.* **109** 201–11
Wolthuis R A, Froelicher V F, Fischer J and Triebwasser J H 1977 The response of healthy men to treadmill exercise *Circ. (New York, N.Y.)* **55** 153–7
Wrzus C, Müller V, Wagner G G, Lindenberger U and Riediger M 2014 Affect dynamics across the lifespan: with age, heart rate reacts less strongly, but recovers more slowly from unpleasant emotional situations *Psychol. Aging* **29** 563–76