Biomedical Signal Processing and Control 18 (2015) 370–377

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Review

Acute mental stress assessment via short term HRV analysis in healthy adults: A systematic review with meta-analysis

R. Castaldo [a], P. Melillo [b], U. Bracale [c], M. Caserta [a,c], M. Triassi [c], L. Pecchia [a,*]

[a] School of Engineering, University of Warwick, Coventry CV4 7AL, UK
[b] Multidisciplinary Department of Medical, Surgical and Dental Sciences, Second University of Naples, Italy
[c] Department of Public Health, University Federico II Naples, Naples, Italy

ARTICLE INFO

Article history:
Received 28 November 2014
Received in revised form 6 February 2015
Accepted 18 February 2015
Available online 7 March 2015

Keywords:
Acute mental stress
Heart Rate Variability
Systematic review with meta-analysis

ABSTRACT

Mental stress reduces performances, on the work place and in daily life, and is one of the first causes of cognitive dysfunctions, cardiovascular disorders and depression. This study systematically reviewed existing literature investigating, in healthy subjects, the associations between acute mental stress and short term Heart Rate Variability (HRV) measures in time, frequency and non-linear domain. The goal of this study was to provide reliable information about the trends and the pivot values of HRV measures during mental stress. A systematic review and meta-analysis of the evidence was conducted, performing an exhaustive research of electronic repositories and linear researching references of papers responding to the inclusion criteria. After removing duplicates and not pertinent papers, journal papers describing well-designed studies that analyzed rigorously HRV were included if analyzed the same population of healthy subjects at rest and during mental stress. 12 papers were shortlisted, enrolling overall 758 volunteers and investigating 22 different HRV measures, 9 of which reported by at least 2 studies and therefore meta-analyzed in this review. Four measures in time and non-linear domains, associated with a normal degree of HRV variations resulted significantly depressed during stress. The power of HRV fluctuations at high frequencies was significantly depressed during stress, while the ratio between low and high frequency resulted significantly increased, suggesting a sympathetic activation and a parasympathetic withdrawal during acute mental stress. Finally, among the 15 non-linear measures extracted, only 2 were reported by at least 2 studies, therefore pooled, and only one resulted significantly depressed, suggesting a reduced chaotic behaviour during mental stress. HRV resulted significantly depressed during mental stress, showing a reduced variability and less chaotic behaviour. The pooled frequency domain measures demonstrated a significant autonomic balance shift during acute mental stress towards the sympathetic activation and the parasympathetic withdrawal. Pivot values for the pooled mean differences of HRV measures are provided. Further studies investigating HRV non-linear measures during mental stress are still required. However, the method proposed to transform and then meta-analyze the HRV measures can be applied to other fields where HRV proved to be clinically significant.

Contents

1. Introduction 371
2. Methods and materials 371
   2.1. Search strategy 371
   2.2. Inclusion and exclusion criteria for paper selection 371
   2.3. Paper shortlisting, data extraction and outcomes of interest 372
   2.4. HRV measures 372
   2.5. Statistical analysis and software tools 372

* Corresponding author at: School of Engineering, University of Warwick, Coventry CV4 7AL, UK. Tel.: +44 024 765 73383; mobile: +44 07505716490.
E-mail addresses: R.Castaldo@warwick.ac.uk (R. Castaldo), paolo.melillo@unina2.it (P. Melillo), umberto.bracale@unina.it (U. Bracale), triassi@unina.it (M. Triassi), l.pecchia@warwick.ac.uk (L. Pecchia).

http://dx.doi.org/10.1016/j.bspc.2015.02.012

3. Results 372
3.1. Characteristics of the included studies 372
3.2. Trends of the HRV measures 373
3.2.1. Time domain features 373
3.2.2. Frequencies domain features 373
3.2.3. Non-linear features 373
3.3. Pooled pivot values of HRV measures mean differences 373
3.4. Mathematical modelling of HRV measures during mental stress 374
4. Discussion 375
5. Conclusion 376
References 376

1. Introduction

Mental stress has been investigated in various fields due to its destructive effects on the daily routine. Stress may cause cognitive dysfunctions, cardiovascular diseases, depression [1,2] and may lead to death and illness [3]. There is, in fact, an average of 50% of employees, which suffer of 'work stress' [4]. Moreover, stress reduces performances in daily life [5] and particularly in the work places [1,6]. This may be particularly relevant for those jobs that expose the workers or other persons to risky situations. This is the case of surgeons performing long surgeries, in which the loss of attention or concentration may cause severe effects on the patient, or pilots flying on long distances, whose stress may be dangerous for them and the passengers [7].

Mental stress influences the Autonomous Nervous System (ANS), which controls our capability to react to external stimuli [3,5]. Therefore, the acute stress may be evaluated with non-invasive biomarker measurements, which are considered reliable estimators of ANS statuses. This is the case of Heart Rate Variability (HRV), which is considered a reliable means to indirectly observe ANS [8], also in real life settings. HRV refers to the variations of both instantaneous heart rate and the series of inter-times between consecutive peaks of the R-wave of the ECG (RR series) [9]. This variation is under the control of the ANS, which through the parasympathetic and the sympathetic branches, is responsible for adjusting the HRV in response to external or internal physical or emotional stimuli. A normal subject shows a good degree of variation of the heart rate, reflecting a good capability to react to those stimuli [9].

Several studies investigated in the last years how HRV could be used to assess acute or chronic mental stress, in healthy subjects or in specific groups of patients, of different ages. These studies collected and analyzed HRV with different methods and tools: long or short term registrations were utilized; HRV was extracted from ECG or SpO2 or via other means; a wide range of devices was employed (i.e. from commercial smart phones to CE marked medical devices for ECG monitoring, including some advanced biomedical amplifiers); HRV was analyzed with commercial software tools, sometimes well validated ones; HRV features in time and/or frequency were extracted and both linear and/or non-linear domains were explored; different statistical tests, pattern recognition or data mining methods were used to explore results [4,9,10].

The goal of this study was to review homogenously designed studies, in order to provide reliable information about the trends and the pivot values of HRV measures during mental stress. Therefore, this reviews investigated systematically studies published in peer-reviewed journals and focusing on the associations between acute mental stress and HRV, using short term HRV excerpts, extracted from ECGs, in healthy subjects above 18 years old.

2. Methods and materials

2.1. Search strategy

Relevant studies on the detection of acute mental stress through the analysis of the HRV were identified and selected by searching on PubMed and OvidSP databases.

Pertinent articles were searched using Boolean combinations of the following keywords or their equivalent Medical Subject Heading (MeSH) terms: Heart Rate Variability, HRV, mental stress, psychological stress, and emotional assessment.

The following criteria were utilized to limit the research: papers published in the last 10 years (since 2004), study on adult humans (i.e. not animals), not children, not cancer, not pregnancy.

2.2. Inclusion and exclusion criteria for paper selection

After a first screening of the titles and abstracts, studies published before November 2014 were considered suitable for this review if they met all of the following criteria:

- Paper published on scientific journals with peer review;
- Paper focused mainly on mental stress investigation using HRV;
- The experiments were conducted with a robust design, well described, including repeated measures in the same group of healthy subjects at rest and during stress section. Studies were considered robust if the acquisition of the HRV was adequately standardized (i.e. stress and rest sessions in the moment of the day to minimize the circadian effect and with the subjects in the same position).
- The study was not focusing on chronic stress, and only acute stressors where used, as defined in [11].
- The subjects were humans beings over 18 years old;

Studies were excluded if they:

- focused on chronic stress;
- utilized HRV analysis on excerpts were longer than 10 min;
- enrolled professional athletes and observed them during sport training sessions;
- investigated pain perception;
- reported HRV measures with not sufficient quality (i.e. no unit measures stated, not appropriate statistical descriptors);
- did not report inclusion/exclusion criteria for sample selection.

Two authors (RC and PM) independently assessed the methodological quality of the papers, given the inclusion/exclusion criteria. A third author (LP) arbitrated when necessary.

2.3. Paper shortlisting, data extraction and outcomes of interest

Following the research strategy described above, all the titles responding to the chosen keywords were identified. After excluding repetitions (titles indexed both in PubMed and OvidSp), studies were shortlisted according to inclusions/exclusions criteria, investigating titles, abstracts and full-papers.

HRV measures in time and frequency domains, both linear and non-linear, were extracted by the papers included in this review [1,5,6,12–20]. Further details about HRV methodological conventions adopted in this review can be found in two open access papers [21,22].

The changes of HRV measures during stress and rest sessions were investigated observing their trends. A trend was represented with arrows pointed up, if the mean value of a HRV measure was increasing during the stress section, with respect to the value of the same measure registered during the resting session. An arrow pointed down was used if the mean value of the HRV measure was decreasing. Two arrows were utilized for significant changes (p-values <0.05). Since papers reported only parameters of measure distributions (i.e. means and standard deviations) the p-values were calculated with a t-test and not with parametric tests, which in some cases would have been more appropriated.

2.4. HRV measures

Regarding linear HRV measures in time and frequency domains, the recommendations of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology guidelines [9] were followed. Therefore, the power spectrum density measures were considered if reported in normalized units or ms^2, for the low frequency bands (LF, 0.04–0.15 Hz) and the high frequency bands (HF, 0.15–0.4 Hz). When papers shortlisted for meta-analysis reported features in a non-conventional way, these measures were excluded from the pooling or converted in ms^2, if possible. For instance, three studies [16,17,19] reported frequency domain features log-transformed (LF_{Log} and HF_{Log}). These were untransformed according to [23], using the following formula:

$$\text{mean LF}(\text{ms}^2) = \exp(\text{mean } LF_{Log})$$

$$\text{std LF}(\text{ms}^2) = \text{std } LF_{Log} \times [\exp(\text{mean } LF_{Log})]^2$$

One study reported frequency measures in ms^2/Hz [14] and therefore its results were not reported in this review.

2.5. Statistical analysis and software tools

HRV measures were pooled if reported in more than one paper. Standard methods for systematic reviews with meta-analyses were employed in this paper to pool the HRV measures [24]: mean difference (MD) with 95% confidence intervals (95%CI) and p-values (p). Random or fixed effect models were used according to study heterogeneity measured with a Q statistic test. A significant Q statistic is indicative of dissimilar effect sizes across studies and to complement the Q test, we also calculated the I^2 statistic, which provides an index of the degree of heterogeneity across studies. In particular, I^2 indicates the percentage of the total variability in effect sizes due to between-studies variability, and not due to sampling error within studies [25]. Percentages of around 25% (I^2 = 25), 50% (I^2 = 50), and 75% (I^2 = 75) may be interpreted as low, medium, and high heterogeneity, respectively [25]. Differences and 95%CI were considered significant if the p-value was less than 0.05. The Open Meta-analyst Software tool was utilized to compute these statistics.

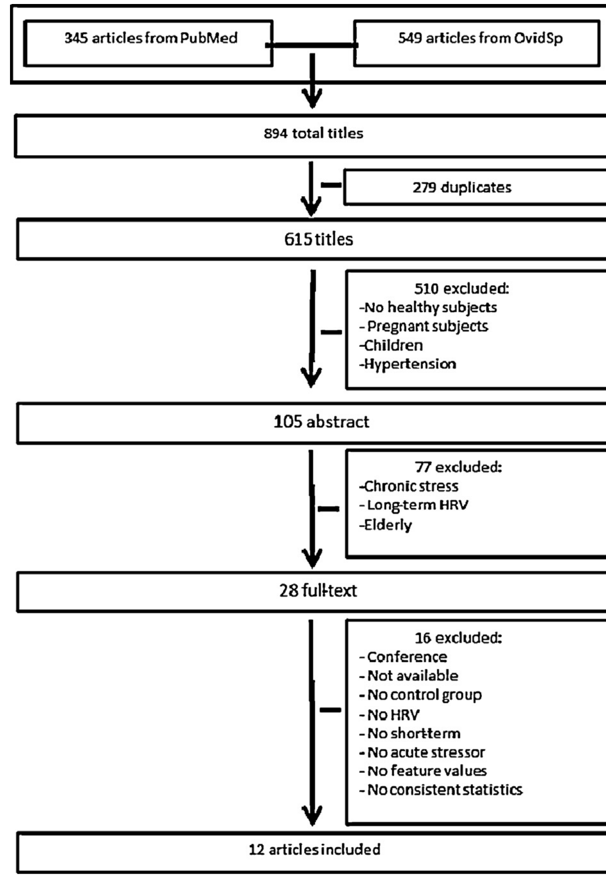

Fig. 1. Flow chart of literature search: titles, abstract and full-papers included/excluded.

3. Results

According to the search strategy described above, 894 titles were identified, 345 in PubMed and 549 in OvidSp. After removing 279 duplicates, 615 titles were considered. Of these, 510 were excluded after reading the abstracts as they did not meet the inclusion criteria. From the remaining 105 abstracts, 77 were removed due to the exclusion criteria. Finally, 28 full-texts were analyzed and among these, 16 were excluded due to inclusion/exclusion criteria. Therefore, 12 studies were finally considered appropriate for inclusion in this systematic review. A flow chart of the literature search results is shown in Fig. 1.

3.1. Characteristics of the included studies

The 12 studies enrolled from 12 to 399 subject each, for a cumulative population of 785 subjects, which was finally pooled in this review. Details about population, HRV analysis reported in each study, HRV length and statistical methods employed to explore significant variations are reported in Table 1.

Only three studies used non-parametric statistic tests (NPST), which are specifically designed to investigate significant difference of features non-normally distributed [26]. Nine studies [1,6,12–14,16,17,19,20] used ANOVA, ANCOVA or t-student test to investigate how HRV measures changed before and after the stress section. Of these, Traina et al., Papousek et al., and Lackener et al. [12,13,15] converted frequencies domain HRV features with a log transform, while the remaining applied directly standard statistical significance tests to frequency domain measures, which are not normally distributed, being asymmetric distributions of positive numbers [21,22].

Table 1
Characteristics of studies.

No. of paper	Author/year	Subjects (total/women)	HRV analysis	HRV length (min)	Significance tests
1	Hjortskov 2004 [6]	12/12	Time and frequency	3	ANOVA
2	Kofman 2006 [1]	30/–	Frequency	10	*t*-Tests, ANOVA
3	Vuksanovic 2007 [12]	23/13	Frequency and non-linear	5	*t*-Tests
4	Li 2009 [13]	399/209	Time and frequency	0.5	ANOVA
5	Schubert 2009 [14]	50/28	Time, frequency and non-linear	3	ANCOVA
6	Tharion 2009 [15]	18/9	Time and frequency	5	Non-parametric statistical tests (Wilcoxon signed rank test)
7	Papousek 2010 [17]	65/53	Time and frequency	3	ANOVA
8	Lackner 2011 [16]	20/20	Time and frequency	5	ANOVA
9	Melillo 2011 [5]	42/–	Non-linear	5	Non-parametric statistical tests (Wilcoxon signed rank test)
10	Taelman 2011 [18]	43/22	Time and frequency	6	Non-parametric statistical tests (Wilcoxon and Friedman tests)
11	Traina 2011 [19]	13/6	Frequency	5	*t*-Test
12	Visnovcova 2014 [20]	70/39	Time and frequency	6	*t*-Test, non-parametric statistical tests (not specified)

The selected studies reported ECG recordings at rest (resting session, RS) and during induced mental stress sessions (stress session, StS). Only one study [15] did not report ECG recordings during the StS, but the ECG was registered on the day of a university examination, which was assumed to be the stressing event. This may generate heterogeneity, as the response to stress may change during or after StS [27].

Mental stress was induced asking volunteers to perform, under controlled circumstances, one or more of the following tasks: Computer Work Task (CWT), Stroop Colour Word Task (SCWT), Arithmetic task (AT), Game task (GT), public Speech task (ST), Academic examination (AE), other Physical-Mental Task (PMT). A brief description of each mental task is given in Table 2. Additionally, in one study [18] were asked to perform also a physical task. HRV measures registered during those physical tasks were not included in this review since producing a different kind of response than mental stress tasks [28].

3.2. Trends of the HRV measures

Tables 3–5 reported the trends and the values of HRV measure in RS and StS in time, frequency and non-linear domains respectively. Additionally, these tables reported the number of subjects enrolled in each study and the relative weight of each study during the pooling.

3.2.1. Time domain features

Time domain HRV measures reported by the papers shortlisted for this meta-analysis included: the mean of R-R intervals (RR), a standard deviation of an R-R interval (SDRR), the square root of mean squared difference of successive R-Rs (RMSSD), and a proportion of NN50 divided by total number of NN that differ more than 50 ms (pNN50) as shown in Table 3.

In all the studies there was consensus that the mean RR, pNN50 and RMSSD decreased during StS, although some studies did not demonstrated a significant reduction of these features. A decreased SDRR was observed in the majority of studies [15,18,20] with a high level of significance. Only one study reported contradictory results [14]. In fact Schubert et al. [14] reported a discordant increase of SDRR, which the authors justifies as due to a slow respiration rate and a relative reduction in ventilation caused by the speech task used in to induce stress in this study [14].

3.2.2. Frequencies domain features

Frequency domain HRV measures extracted from the shortlisted papers included: the low frequency power (LF), the high frequency power (HF) and the LF/HF ratio as shown in Table 4.

Regarding LF, among the 8 papers reporting this measure, 5 studies agreed that LF increased (3 with statistical significance), while 3 reported an opposite trend. Also this case, the findings by Tharion et al. [15] were not consistent with the general trend, probably because in this study the stress section was recorded in the day of the examination and not during the examination session. In addition, Hjortskov et al. [6] also showed controversial results, which in this case might be explained considering that this was the only study in which an introduction session was ran before the registration of the RS and StS. Finally, Taelman et al. [18] also reported a decreased LF during StS, which might be also due to the different protocol adopted, which consisted in a physical tasks before the metal stress session.

Regarding HF, there is consensus, among the different studies, that this measure decreased during acute mental stress. Only one study indicated a controversial trend, which however was not statistical significant [12].

Regarding LF/HF ratio, there was a general consensus (5 studies out of 7) that it increased during StS. The remaining 2 studies [12,15] reported an opposite trend, which however is not supported by statistical significance.

3.2.3. Non-linear features

Non-linear HRV measures extracted from the shortlisted papers included: the Shannon and the Sample Entropy (respectively ShEn and SampEn), the largest Lyapunov exponent (LLE), the correlation dimension (D_2), short- and long term fluctuation slope (α_1 and α_2), the standard deviation of the Poincare' plot perpendicular and along to the line of identity (respectively SD1 and SD2), recurrence plot determinism (DET) and recurrence rate (REC), maximal length of lines (l_{max}) and mean length of lines (l_{mean}). Finally, the Approximate Entropy (En) computed with the threshold r set to $0.2\times$ SDNN, to the value maximizing the entropy and to the value computed by Chon et al. [29], and reported respectively as En (0.2), En(r_{max}) and En(chon) as shown in Table 5.

Only two features were investigated by more than one study: D_2 [5,14] was reported as consistently reduced during StS (academic examination in Melillo et al. [5] and arithmetical task in Shubert et al. [14]); α_1 [5,12] which was reported with significantly opposite trends.

3.3. Pooled pivot values of HRV measures mean differences

From the included studies, 8 HRV measures were pooled in this systematic meta-analysis because reported at least in two papers: 4 measures in the time domain (RR, SDRR, pNN50%, rMSSD), 3 in the frequency domain (LF, HF, LF/HF) and 2 in the non-linear domain

Table 2
Description of study design.

No. of paper	Author/year	Stressor	Description
1	Hjortskov 2004 [6]	Computer Work Task (CWT)	The task consisted of keying in random numbers with the dominant hand using the numerical part of the keyboard.
2	Kofman 2006 [1]	Stroop Colour Word Task (SCWT)	The task required naming the ink colour in which colours' names were printed on a screed being either congruent or incongruent with respect to the colour words.
3	Vuksanovic 2007 [12]	Arithmetic aloud task (AT)	The task required subtracting 17 from 1000. Subjects were asked to answer as accurate as possible. They told the results after 5 s.
4	Li 2009 [13]	Game task (GT)	The instructions for video game task 'Breakout' have been standardized via video. The subjects will lie supine on a hospital bed and a monitor 25 inch colour TV was placed 2 m away.
5	Schubert 2009 [14]	Speech task (ST)	This task involved preparing and presenting a speech in response to one of two situations. Subjects had 3 min to prepare and 3 min to present their speech.
6	Tharion 2009 [15]	Academic examination (AE)	The first recording was done on the day of the academic examination and the second during holiday.
7	Papousek 2010 [17]	Academic examination (AE)	Participants were told to not move legs and hands while speaking, to read out the questions loud before answering it.
8	Lackner 2011 [16]	Arithmetical Task (AT)	The task consisted of two triplet of one-digit numbers, which were to be added or subtracted.
9	Melillo 2011 [5]	Academic Examination (AE)	The first record was performed during verbal examination and the second one was performed after holiday.
10	Taelman 2011 [18]	Physical-Mental Task (PMT)	The task consisted of continuous mental calculation of five operations with a two or three digit number, which had to be performed without verbal stimulation. Participants used the mouse cursor to indicate the correct answer choosing between three alternatives. The participants would have been rewarded with a movie ticket
11	Traina 2011 [19]	Arithmetical Task (AT)	The task required to subtracting the number 17 starting from the number 986.
12	Visnovcova 2014 [20]	Stroop Colour Word Test (SCWT) Arithmetic Test (AT)	The task required to read the colours (green, yellow, orange, red, blue, purple) on the words displayed on a screen, which were congruent or incongruent with the written word. (Results not reported in this review) The task required to calculate three-digit numbers displayed in different random places on the screen into one digit numbers. Subsequently, participants decided that the final result was even or odd by pushing the keyboard arrow.

(α_1, D_2). The relative pooling weight of each study was reported in Tables 3–5, respectively for time, frequency and non-linear domain. The results of the pooling were reported in Table 6, where also the trend of the pooled HRV measures is reported.

3.4. Mathematical modelling of HRV measures during mental stress.

Two studies [5,19] proposed a mathematical model to automatically detect mental stress.

Melillo et al. [5] proposed a model based on Linear Discriminant Analysis (LDA), employing three HRV non-linear measures: SD1, SD2 and En(0.2). The proposed model achieved accuracy, sensitivity and specificity respectively of 90%, 86% and 95% in automatically detecting subjects under stress. These performances were achieved testing the developed classifier using a 10-fold cross validation technique, based on subjects' exclusions.

Traina et al. [19] studied the Pearson correlation between frequency domain measures (high and low components of the power spectrum) before and after the stressing session, demonstrating

Table 3
Time domain features.

Time domain features	Authors	Features trend	Weight of meta-analysis	*N*	Stress		Rest	
					Mean	SD	Mean	SD
Mean RR (ms)	Lackner 2011 [16]	↓	1.82%	20	765.31	314.3	837.1	324.5
	Papousek 2010 [17]	↓↓	11.65%	65	617.92	210.0	819.7	244.1
	Schubert 2009 [14]	↓↓	9.25%	50	686.49	240.8	808.6	206.2
	Taelman 2011 [18]	↓↓	19.98%	43	755.44	134.5	863.5	147.1
	Tharion 2009 [15]	↓↓	12.83%	18	777.40	114.3	867.3	114.0
	Visnovcova 2014 [20]	↓↓	41.22%	70	675.99	120.7	847.8	130.4
	Vuksanovic 2007 [12]	↓	3.25%	23	740.74	263.2	806.5	249.5
SDRR (ms)	Schubert 2009 [14]	↑↑	3.50%	50	96.00	86.6	33.20	23.83
	Taelman 2011 [18]	↓↓	37.56%	43	35.40	16.4	46.73	19.48
	Tharion 2009 [15]	↓↓	8.90%	18	52.40	21.7	74.20	25.93
	Visnovcova 2014 [20]	↓↓	50.05%	70	48.98	17.9	56.23	21.67
RMSSD (ms)	Li 2009 [13]	↓↓	26.85%	105	55.50	29.6	68.40	37.70
	Li 2009 [13]	↓↓	14.61%	84	57.20	36.9	74.20	44.90
	Taelman 2011 [18]	↓↓	54.38%	43	19.39	13.8	28.74	16.58
	Tharion 2009 [15]	↓	4.17%	18	49.99	31.07	74.03	39.64
pNN50 (%)	Taelman 2011 [18]	↓	78.43%	43	26.82	16.10	31.83	18.73
	Tharion 2009 [15]	↓↓	21.57%	18	20.57	19.04	39.37	23.79

↓↓ (↑↑): significantly lower (higher) under stress ($p < .05$).
↓ (↑): lower (higher) during stress section ($p > .05$).

Table 4
Frequency domain features.

Frequency domain features	Authors	Feat. trend	Weight of meta-analysis	N	Stress		Rest	
					Mean	SD	Mean	SD
LF (ms^2)	Hjortskov 2004 [6]	↓↓	8.52%	12	1391	1028	1664	808
	Lackner 2011 [16]	↑	10.16%	20	1339	1205	812.4	649.9
	Papousek 2010 [17]	↑	14.33%	65	1644	987.0	997.3	598.6
	Taelman 2011 [18]	↓↓	14.83%	43	466.9	460.2	868.4	641.0
	Tharion 2009 [15]	↓↓	5.74%	18	1192	723.6	2155	2157
	Traina 2011 [19]	↑↑	15.36%	13	1241	312.0	511.0	114.5
	Visnovcova 2014 [20]	↑↑	15.69%	70	607.9	457.7	454.9	380.6
	Vuksanovic 2007 [12]	↑↑	15.37%	23	464.0	356.1	387.6	260.3
HF (ms^2)	Hjortskov 2004 [6]	↓↓	5.45%	12	1131	718	1776	1092
	Li female 2009 [13]	↓↓	8.39%	84	1200	1454	2001	2066
	Li male 2009 [13]	↓↓	11.82%	105	1072	1058	1677	1751
	Papousek 2010 [17]	↓	18.14%	65	668.5	401.5	1097	665.1
	Taelman 2011 [18]	↓↓	15.62%	43	552.5	428.0	1005	782.6
	Tharion 2009 [15]	↓↓	1.54%	18	1691	2096	2892	2622
	Traina 2011 [19]	↓	17.91%	13	252.1	263.6	273.1	246.2
	Visnovcova 2014 [20]	↓↓	11.50%	70	445.9	1082	639.1	1337
	Vuksanovic 2007 [12]	↑	9.66%	23	665.1	925.1	595.9	714.4
LF/HF (–)	Hjortskov 2004 [6]	↑	15.12%	12	1.57	1.09	1.16	0.74
	Papousek 2010 [17]	↑	22.90%	65	1.11	0.59	0.01	0.80
	Schubert 2009 [14]	↑	19.88%	50	1.80	1.20	1.50	1.11
	Tharion 2009 [15]	↓	12.74%	18	1.30	0.80	1.40	1.80
	Traina 2011 [19]	↑↑	4.61%	13	5.81	2.88	2.57	2.10
	Vuksanovic 2007 [12]	↓	4.61%	23	1.07	3.83	1.08	2.92
	Kofman 2006 [1]	↑	20.14%	30	1.48	1.02	0.97	0.68

↓↓ (↑↑): significantly lower (higher) under stress ($p < .05$).
↓ (↑): lower (higher) during stress section ($p > .05$).

that those correlations were significant. However, this is partially arguable as the Pearson correlation would eventually lay on the assumption that the HRV measures are normally distributed, while HRV frequency measures are not. This paper did not develop a predictive model and did not validate the correlation with a cross validation techniques.

4. Discussion

This paper presented the results of a systematic literature review with meta-analysis of the articles investigating how short-HRV measures changed during induced acute mental stress. In this review, 12 studies were shortlisted and changes during mental stress of 22 HRV measures were systematically reported. Finally, 9 HRV measures were pooled.

The results demonstrated that 4 HRV measures (RR, RMSSD, pNN50 and D_2) decreased during stress [13–15,17,18,20].

The majority of studies [6,13,15,17–20] agreed that SDRR and HF decreased during stress, while LF/HF and LF increased during stress [1,6,12,14,16,17,19,20]. Regarding SDRR, only one study [14] out of the 4 [15,18,20] considering this measure, reported an increasing value during stress. This may be due to the fact that [14] analyzed HRV excerpts of 3 min, while [15,18,20] analyzed 5 min HRV excerpts.

Regarding LF, 3 studies [6,15,18] out of 8 [12,16,17,19,20] considering this measure, reported a decreasing value during stress. However, these studies adopted study designs significantly different from the others. Differently from the other, the first two studies used physical movements during the stress section: in Hjortskov et al. [6], participants used the dominant hand to digit randomly

Table 5
Non-linear features.

Non-linear domain features	Authors	Features trend	Weight of meta-analysis	N	Stress		Rest	
					Mean	SD	Mean	SD
α_1 (–)	Melillo 2011 [5]	↓↓	49.41%	42	1.05	0.44	1.41	0.16
	Vuksanovic 2007 [12]	↑↑	50.59%	23	0.97	0.19	0.85	0.19
D_2 (–)	Melillo 2011 [5]	↓↓	45.89%	42	1.65	1.28	2.83	1.09
	Schubert 2009 [14]	↓↓	54.11%	50	3.2	0.33	3.5	0.27
SD1 (ms)	Melillo 2011 [5]	↓↓	–	42	0.02	0.01	0.02	0.01
SD2 (ms)	Melillo 2011 [5]	↓↓	–	42	0.05	0.02	0.08	0.02
En(0.2) (–)	Melillo 2011 [5]	↓↓	–	42	0.99	0.24	1.09	0.13
En(r_{chon}) (–)	Melillo 2011 [5]	↓↓	–	42	0.98	0.24	1.11	0.11
En(r_{max}) (–)	Melillo 2011 [5]	↓	–	42	1.09	0.17	1.12	0.10
α_2 (–)	Melillo 2011 [5]	↓	–	42	0.76	0.13	0.78	0.18
ShEn (–)	Melillo 2011 [5]	↑↑	–	42	3.42	0.39	3.17	0.23
DET (%)	Melillo 2011 [5]	↑	–	42	98.75	1.28	98.61	0.86
REC (%)	Melillo 2011 [5]	↑↑	–	42	42.24	12.05	33.46	6.27
l_{mean} (beats)	Melillo 2011 [5]	↑↑	–	42	14.88	6.77	11.09	2.48
l_{max} (beats)	Melillo 2011 [5]	↓↓	–	42	213.4	136.5	286.7	111.2
LLE	Vuksanovic 2007 [12]	↑	–	23	0.06	0.019	0.06	0.019
SampEn (–)	Vuksanovic 2007 [12]	↓↓	–	23	1.65	0.06	1.77	0.04

↓↓ (↑↑): significantly lower (higher) under stress ($p < .05$).
↓ (↑): lower (higher) during stress section ($p > .05$).

Table 6
Pooled HRV measures.

Outcome	Subjects	Heterogeneity (I^2, p)	Model	Units	MD	CI95%	p-value	Trend
Time								
RR	289	33%, 0.17	Fixed	ms	−**142.2**	(−**168.9**; −**115.47**)	**<0.01**	↓↓
SDRR	181	92%, <0.01	Fixed	ms	−**7.627**	(−**12.20**; −**2.97**)	**<0.01**	↓↓
RMSSD	250	0%, 0.50	Fixed	ms	−**12.03**	(−**16.78**; −**7.28**)	**<0.01**	↓↓
pNN50	61	65%, 0.09	Fixed	–	−**7.98**	(−**14.52**; −**1.45**)	**<0.05**	↓↓
Frequency								
LF	264	91%, <0.01	Random	ms^2	156.1	(157.6; 469.8)	0.33	↑
HF	433	62%, <0.01	Random	ms^2	−**359.7**	(−**559.20**; −**160.25**)	**<0.01**	↓↓
LF/HF	211	75%, <0.01	Random	–	**0.61**	(**0.14**; **1.08**)	**<0.01**	↑↑
Non-linear								
α_1	65	96%, <0.01	Random	–	−0.12	(−0.59; 0.35)	0.63	↓
D_2	92	91%, <0.01	Fixed	–	−**0.35**	(−**0.46**; −**0.23**)	**<0.01**	↓↓

↓↓ (↑↑): significantly lower (higher) under stress ($p < .05$).
↓ (↑): lower (higher) during stress section ($p > .05$).
MD: mean difference; CI95%: confidence interval at 95%; DL: dimensionless.
Bold values indicate significant changes.

numbers on the left part of a keyboard; in Taelman et al. [18], participants used the mouse cursor to indicate the correct answer choosing between three alternatives. Differently from the other protocols, the use of the hand activated a cortical area that is not triggered by the designs adopted by the remaining studies, where there is no physical activity. These results were also consisted with the findings reported in Yu et al. [30] with arithmetic test, in which the participants were required to use the keyboard. The other study reporting a decreased value of LF during the stress session was Tharion et al. [15], which however measured the physiological response to the mental stress during the day of the University examination, and not during the examination, as done in [5,17], which also used University examination as stressor. At this regard, there is consensus that the reaction to a stressing situation is composed by several phases [27], each implying a different response of the ANS and therefore a different HRV modulation. Finally, only one HRV measure, α_1, was reported by two studies [5,12] achieving completely opposite results, and both with statistical significance. No clear explanation for such heterogeneous results can be inferred by the two papers.

Regarding the meta-analysis, the pooled values of 7 HRV measures (RR, SDRR, RMSSD, pNN50, D_2, HF and LF/HF) out of the 9 meta-analyzed changed significantly during mental stress. The pooled values of those 7 measures should be regarded as they are more reliable than those presented by each paper and should be considered as possible pivot values for future studies. For instance, the increase of the polled LF/HF was statistically significant while it was not statistically significant in 6 out of the 7 papers reporting this HRV measure. Since the p-value is strongly dependent by the sample size, a possible explanation is that the number of volunteers enrolled in each of those 6 studies was too small compared with the LF/HF mean differences measured, which, in turn, were too small compared with the standard deviations measured in each group during stress and rest [24]. Regarding the LF, the pooled results would have become statistical significant if these three studies employing a different design (as discussed above [18,6] moving hands and [15] measuring stress in the day of examination and not during it) were excluded. In fact, removing these 3 studies the pooled mean difference for LF during stress was 286.56 ms^2 (CI95% 183.89–389.23, $p < 0.01$, fixed effect model), calculated on 164 subjects with a very small heterogeneity (I^2 94%, $p < 0.001$). Finally, it was hard to discuss the results about α_1, since the two studies calculating this HRV measures adopted similar protocols and comparable sample sizes. Melillo et al. [5] enrolled about the double of subjects, but the SD measured in the stressing session was too high, affecting the weight of the study. Therefore, future researches investigating non-linear HRV features during mental stress are still needed.

The decreases of RR, RMSSD, pNN50, SDRR reflected a depressed HRV during stress. This is consistently confirmed by pooled values of HRV frequency measures, among which HF was proved to decrease significantly, reflecting a decreased HRV variability. Moreover, the observed decrease of D_2 during stressing sections can be interpreted as a reduced complexity of the HRV, reflecting a lower adaptability and fitness of the cardiac pacemaker and a functional restriction of the participating cardiovascular elements [14].

Finally, the frequency domain measures consistently supported the idea that during stress there is a general depression of HRV with a relative displacement of the vago-sympathetic balance, during which sympathetic activation relatively overcome parasympathetic once. In fact, the LF accounting for an activation of both parasympathetic and sympathetic system, increased, while the HF, which is associated to the parasympathetic system activation, decreased. This result was confirmed by the increase of LF/HF. These outcomes confirmed the inducted shift of the ANS balance towards the sympathetic activation and the parasympathetic withdrawal during acute mental stress [20]. This phenomenon was explained through the theory of the fight or flight response, which support the idea that there is an inhibition of the vago and a prevalence of the sympathetic system during a stressing situation [27]. This result could change if the stressing session was measured after (and not during) the stressing event, as demonstrated in [15] and consistently with the phases of the stress as described in [27].

5. Conclusion

In conclusion, these review proved that there was consensus about the HRV measures changing consistently during mental stress. Particularly, the results of the pooling of the HRV measures, provided pivot values for at least 7 HRV measures that changed significantly during stressing sections. These significant changes confirmed previous results about the inducted shift of the ANS balance towards the sympathetic activation and the parasympathetic withdrawal during acute mental stress. However, a huge heterogeneity among studies investigating physiological response to mental stress was observed in this review. Moreover, this review proved that future studies are needed to confirm the behaviour of non-linear HRV measures during stress.

Finally, future studies are recommended to: define the study design (i.e. length of HRV measures) and the study protocol (i.e. definition of stressor, avoid physical activity if not required) according to the best available evidence; to analyses the HRV features (i.e. using measurement units according to international guidelines); to use statistical tests consistently with the HRV measure distribution (i.e. check if HRV measures are normally distributed of not); to check HRV stationary if HRV measures are longer than 5 min; to measure stressing session according to the goal of the study (i.e. during, before or after the stressing session).

References

[1] O. Kofman, N. Meiran, E. Greenberg, M. Balas, H. Cohen, Enhanced performance on executive functions associated with examination stress: evidence from task-switching and Stroop paradigms, Cognit. Emot. 20 (2006) 577–595.
[2] O.V. Crowley, P.S. McKinley, M.M. Burg, J.E. Schwartz, C.D. Ryff, M. Weinstein, et al., The interactive effect of change in perceived stress and trait anxiety on vagal recovery from cognitive challenge, Int. J. Psychophysiol. 82 (December) (2011) 225–232.
[3] R.L. Pan, J.K. Li, A noninvasive parametric evaluation of stress effects on global cardiovascular function, Cardiovasc. Eng. 7 (June) (2007) 74–80.
[4] T. Chandola, A. Heraclides, M. Kumari, Psychophysiological biomarkers of workplace stressors, Neurosci. Biobehav. Rev. 35 (September) (2010) 51–57.
[5] P. Melillo, M. Bracale, L. Pecchia, Nonlinear Heart Rate Variability features for real-life stress detection. Case study: students under stress due to university examination, Biomed. Eng. Online 10 (2011) 1–13.
[6] N. Hjortskov, D. Rissén, A. Blangsted, N. Fallentin, U. Lundberg, K. Søgaard, The effect of mental stress on heart rate variability and blood pressure during computer work, Eur. J. Appl. Physiol. 92 (2004) 84–89.
[7] R.R. Singh, S. Conjeti, R. Banerjee, A comparative evaluation of neural network classifiers for stress level analysis of automotive drivers using physiological signals, Biomed. Signal Process. Control 8 (2013) 740–754.
[8] A.H. Marques, M.N. Silverman, E.M. Sternberg, Evaluation of stress systems by applying noninvasive methodologies: measurements of neuroimmune biomarkers in the sweat, heart rate variability and salivary cortisol, Neuroimmunomodulation 17 (2010) 205–208.
[9] T. Force, Heart rate variability guidelines: standards of measurement, physiological interpretation, and clinical use, Eur. Heart J. 17 (1996) 354–381.
[10] U.R. Acharya, K.P. Joseph, N. Kannathal, C.M. Lim, J.S. Suri, Heart rate variability: a review, Med. Biol. Eng. Comput. 44 (2006) 1031–1051.
[11] S.C. Segerstrom, G.E. Miller, Psychological stress and the human immune system: a meta-analytic study of 30 years of inquiry, Psychol. Bull. 130 (2004) 601–630.
[12] V. Vuksanović, V. Gal, Heart rate variability in mental stress aloud, Med. Eng. Phys. 29 (2007) 344–349.
[13] Z. Li, H. Snieder, S. Su, X. Ding, J.F. Thayer, F.A. Treiber, et al., A longitudinal study in youth of heart rate variability at rest and in response to stress, Int. J. Psychophysiol. 73 (September) (2009) 212–217.
[14] C. Schubert, et al., Effects of stress on heart rate complexity – a comparison between short-term and chronic stress, Biol. Psychol. 80 (2009) 325–332.
[15] E. Tharion, S. Parthasarathy, N. Neelakantan, Short-term heart rate variability measures in students during examinations, Natl. Med. J. India 22 (Mar-Apr (2)) (2009) 63–66.
[16] H.K. Lackner, I. Papousek, J.J. Batzel, A. Roessler, H. Scharfetter, H. Hinghofer-Szalkay, Phase synchronization of hemodynamic variables and respiration during mental challenge, Int. J. Psychophysiol. 79 (March) (2011) 401–409.
[17] I. Papousek, K. Nauschnegg, M. Paechter, H.K. Lackner, N. Goswami, G. Schulter, Trait and state positive affect and cardiovascular recovery from experimental academic stress, Biol. Psychol. 83 (February) (2010) 108–115.
[18] J. Taelman, S. Vandeput, E. Vlemincx, A. Spaepen, S. Van Huffel, Instantaneous changes in heart rate regulation due to mental load in simulated office work, Eur. J. Appl. Physiol. 111 (July) (2011) 1497–1505.
[19] C.A. Traina, M. Galullo, F.G. Russo, Effects of anxiety due to mental stress on heart rate variability in healthy subjects, Minerva Psichiatr. 52 (2011) 227–231.
[20] Z. Visnovcova, M. Mestanik, M. Javorka, D. Mokra, M. Gala, A. Jurko, et al., Complexity and time asymmetry of heart rate variability are altered in acute mental stress, Physiol. Meas. 35 (2014) 1319.
[21] P. Melillo, N. De Luca, M. Bracale, L. Pecchia, Classification tree for risk assessment in patients suffering from congestive heart failure via long-term heart rate variability, IEEE J. Biomed. Health Inf. 17 (May) (2013) 727–733.
[22] P. Melillo, R. Izzo, N. De Luca, L. Pecchia, Heart rate variability and target organ damage in hypertensive patients, BMC Cardiovasc. Disord. 12 (2012) 105.
[23] J.M. Bland, D.G. Altman, Transformations, means, and confidence intervals, BMJ: Br. Med. J. 312 (1996) 1079.
[24] A.J. Sutton, K.R. Abrams, D.R. Jones, D.R. Jones, T.A. Sheldon, F. Song, Methods for Meta-analysis in Medical Research, John Wiley, 2000.
[25] J.A. Chalmers, D.S. Quintana, J. Maree, A. Abbott, A.H. Kemp, Anxiety disorders are associated with reduced heart rate variability: a meta-analysis, Front. Psychiatry 5 (2014).
[26] S. Siegel, Nonparametric Statistics for the Behavioral Sciences, 1956.
[27] H.S. Bracha, Freeze, flight, fight, fright, faint: adaptationist perspectives on the acute stress response spectrum, CNS Spectr. 9 (2004) 679–685.
[28] S.A. Paritala, Effects of Physical and Mental Tasks on Heart Rate Variability, Kakatiya University, India, 2009.
[29] K. Chon, C.G. Scully, S. Lu, Approximate entropy for all signals, IEEE Eng. Med. Biol. Mag. 28 (November–December) (2009) 18–23.
[30] X. Yu, J. Zhang, Estimating the cortex and autonomic nervous activity during a mental arithmetic task, Biomed. Signal Process. Control 7 (2012) 303–308.