To What Extent Can We Shorten HRV Analysis in Wearable Sensing? A Case Study on Mental Stress Detection.

R. Castaldo[1], L. Montesinos[1], P. Melillo[2], S. Massaro[3] and L.Pecchia[1]

[1] University of Warwick, School of Engineering, Coventry, UK
[2] Second University of Naples, Multidisciplinary Department of Medical, Surgical and Dental Sciences, Naples, Italy
[3] University of Warwick, Warwick Business School – Behavioural Science, Coventry, UK

Abstract— **Mental stress is one of the first causes of cognitive dysfunctions, cardiovascular disorders and depression. In addition, it reduces performances, on the work place and in daily life. The diffusion of wearable sensors (embedded in smart-watches, phones, etc.) has opened up the potential to assess mental stress detection through ultra-short term Heart Rate Variability (HRV) analysis (i.e., less than 5 min). Although informative analyses of features coming from short HRV (i.e., 5 min) have already been performed, the reliability of ultra-short HRV remains unclear. This study aims to tackle this gap by departing from a systematic review of the existing literature and investigating, in healthy subjects, the associations between acute mental stress and short/ultra-short term HRV features in time, frequency, and non-linear domains. Building on these findings, three experiments were carried out to empirically assess the usefulness of HRV for mental stress detection using ultra-short term analysis and wearable devices. Experiment 1 detected mental stress in a real life situation by exploring to which extent HRV excerpts can be shortened without losing their ability to detect mental stress. This allowed us to advance a method to explore to what extent ultra-short HRV features can be considered as good surrogates of 5 min HRV features. Experiment 2 and 3 sought to develop automatic classifiers to detect mental stress through 2 min HRV excerpts, by using a Stroop Color Word Test (CWT) and a highly paced video game, which are two common laboratory-based stressors.**

Results from experiment 1 demonstrated that 7 ultra-short HRV features can be considered as good surrogates of short HRV features in detecting mental stress in real life. By leveraging these 7 features, experiment 2 and 3 offered an automatic classifier detecting mental stress with ultra-short features (2min), achieving sensitivity, specificity and accuracy rate above 60%.

Keywords— **Mental stress, HRV, real-life stressor, CWT, video game**

I. Introduction

Mental stress has been widely investigated in various fields due to its critical effects on daily routine. Mental stress can decrease attentional resources, impair working memory and memory retrieval, and overload cognitive systems. Heart Rate Variability (HRV) is currently one of the most used methods for assessing acute mental stress. Previous studies on mental stress detection have extensively focused on long and short term HRV, but only a few studies have investigated mental stress, and its effects, through ultra-short term HRV analysis thus far [1].

Missing is a model to automatically detect mental stress using ultra-short HRV features [1]. To tackle this gap, we began our overarching research project (Fig.1) by systematically exploring the existing literature to appreciate homogenously designed studies, in order to provide reliable information about the trends and the pivot values of HRV features during mental stress. We reported our main findings in Table 1 [1].

Table 1 Pooled HRV features (adapted from [1])

HRV Features	MD	CI95%	Trend
MeanNN*	-142.2	(-168.9; -115.47)	↓↓
StdNN*	-7.627	(-12.20; -2.97)	↓↓
RMSSD*	-12.03	(-16.78; -7.28)	↓↓
pNN50*	-7.98	(-14.52; -1.45)	↓↓
HF*#	-359.7	(-559.20; -160.25)	↓↓
LF/HF*#	0.61	(0.14; 1.08)	↑↑
D_2*	-0.35	(-0.46; -0.23)	↓↓

↓↓ (↑↑): significantly lower (higher) under stress (p<0.05); MD: mean difference; CI95%: confidence interval at 95%; *p<0.05; # Random Effects Model (otherwise, Fixed Effects Model)

Leveraging this evidence, three experiments were designed in order to develop an automatic classifier for the detection of mental stress using ultra-short term HRV analysis. The first experiment was based on a real life stressor, a verbal academic examination. The full description of this experiment can be found in [2]. To sum it up, its primary goal was to propose a suitable method for assessing reliable ultra-short excerpt lengths required to detect mental stress. Yet, due to the low number of subjects enrolled, it was no possible to develop a fully powered, reliable and robust model for the automatic detection of mental stress. Two further experiments (experiment 2 and 3) were then conducted in lab-settings, by using the Stroop Colour Word Test (CWT) and a highly paced video game as stressors. Experiment 2 and 3 aimed to train, validate, and test an automatic model to detect mental stress using HRV features from 2-min excerpts.

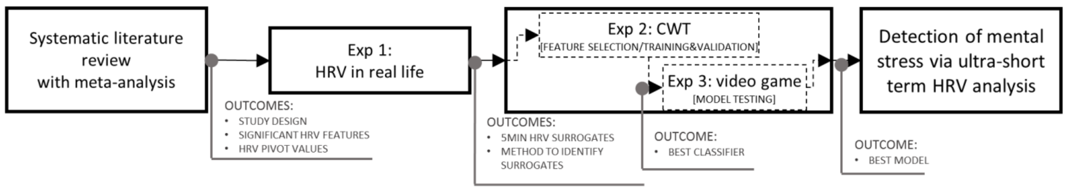

Fig. 1 Study workflow
Exp.: Experiment; CWT: Colour Word Test

The present work shows the results of these experiments and compares how shortened HRV analysis and investigating in-lab stressors (rather than real life ones) can affect the performance of an automatic classifier for mental stress detection.

II. Methods and materials

A. Short vs ultra-short in real life (Experiment 1)

a) Data collection

ECG signals from 42 students were acquired at the School of Biomedical Engineering of the University Federico II. The study was approved by the local Ethic Committee as described in [3]. Signals were acquired over two different days: the first recording was acquired during a university verbal examination (stress phase), while the second was taken in controlled resting condition (rest phase). BioPatch™ M3 devices (Zephyr, Annapolis, USA) were used to capture ECG signals.

b) HRV Analysis

As described in [3], the RR interval time series were extracted from ECG records using PhysioNet's tools [4]. The fraction of total RR intervals labelled as normal-to-normal (NN) intervals was computed as NN/RR ratio. Records with NN/RR ratio below the 90% threshold were excluded. The HRV analysis was performed using Kubios software [5]. Time and frequency measures were analysed according to international guidelines [6], whereas nonlinear measures were analysed as described in [3]. 23 HRV features were extracted in 5 min, 3 min, 2 min, 1 min and 30 sec excerpts and subsequently analysed. However, not all the HRV features were computable in ultra-short time excerpts. LF, LF/HF ratio, total power, and approximate entropy (ApEn) values were excluded for length below 2 min.

c) Statistical analysis

Median, standard deviation, 25th and 75th percentiles were computed to describe the distribution of HRV features for rest and stress phases for 5 min, 3 min, 2 min, 1 min and 30 sec. The non-parametric Wilcoxon Signed-Rank Test was used to investigate the statistical significance ($p<0.05$) of HRV features variation between stress and rest conditions, for each excerpt length (i.e. 5 min, 3 min, 2 min, 1 min and 30 sec). The Spearman's rank correlation (rho) was used to investigate to what extent an ultra-short HRV feature was a good surrogate of the equivalent short feature. Spearman's rank correlation (instead of Pearson's) was used, as HRV features are not normally distributed. In this study, we assumed that each ultra-short HRV feature was a good surrogate of the correspondent short one, if and only if:

- the feature maintained the same behaviour between rest and stress at each length if: the Wilcoxon's test p-value was less than 0.05 between conditions at each timescale; the ultra-short HRV feature trend was changing between conditions coherently with the short HRV feature one.
- the ultra-short HRV feature was highly and significantly correlated (i.e. rho>0.7 and $p_{rho}<0.05$), with the corresponding short feature, across all the time scales in both conditions.

B. Acute mental stress detection in lab settings (Experiment 2 & 3)

a) Dataset acquired during CWT stressor

128 healthy volunteers (49 females; age: 25±3.85 years) with no history of heart disease, systemic hypertension or other disease potentially influencing HRV were enrolled. They were not obese and did not consumed medication, drugs or alcohol in the 12 hours preceding the experiment. They were exanimated under standard conditions: a quiet room, at a comfortable temperature, during the same hour of the morning, and minimizing physical motions and other stimuli possibly affecting HRV. Volunteers were instructed to sit comfortably on an armchair and not move unless necessary. Continuous ECG recording was performed during rest and stress session. The rest session was recorded for 5 minutes during which the subjects were asked simple questions, regarding age, weight and height in order to induce them to speak. Before the stress session started, a brief introduction to the test was explained through demonstrative videos. Finally, the CWT was administered using a big screen for 5 nominal minutes. The study received approval from the Biomedical and Scientific Research Ethics Committee of the University of Warwick (ref. REGO-2014-656).

b) Dataset acquired during video game stressor

42 healthy volunteers (12 females; age: 24±4.5 years) with no history of heart disease, or other disease potentially influencing HRV were examined in the Behavioural Science Laboratory of the Warwick Business School. They were not obese and did not assume medication, drugs or alcohol in the 12 hours preceding the experiment. All subjects were right-handed with normal vision. None of the subjects was expert in the task. All subjects reported daily computer usage and were skilled at operating mouse and keyboard. One of the researchers gave instructions to the participants to help the familiarising with the game. Subjects were exanimated under standard conditions (see: II.B.a). During the rest session, the signal was recorded for 5 minutes in which the subjects were asked some general questions. The stressor task was a shooter popular videogame containing fast-paced content (i.e. war scenes, gun fighting), ranked as adequate for subjects aged 16 and above. During the game, one of the researchers gave instructions and the participants were invited to talk with the virtual leader. BioPatch™ M3 devices (Zephyr, Annapolis, USA) were used to record the ECG for both experiments. The study received approval from the Biomedical and Scientific Research Ethics Committee of the University of Warwick (ref.REGO-2014-656 AMO1).

c) HRV Analysis

Recorded ECGs were segmented and the final 2 minutes of both rest and stress sessions, for both experiments, were used for the ultra-short HRV analysis. The HRV analysis was performed as described in II.A.b. and [7].

d) Statistics and Performance Measurements

Median, standard deviation, 25th and 75th percentiles were calculated to describe the distribution of HRV features at 2 min. The non-parametric Wilcoxon Signed-Rank Test was used to investigate the statistical significances between the stressor session and baseline. HRV features extracted from the CWT were used to train and validate an automatic classifier to detect acute mental stress; HRV features extracted from the videogame task (experiment 3) were used to test the model. The CWT dataset was split in two folders: folder 1 (40% of subjects) was used for feature selection; folder 2 (60% of subjects) was used to train and validate the classifiers. The entire video game dataset was used to test the model. The features selection was two-staged: we performed the relevance analysis and the redundancy analysis as described in [8]. Training of the machine-learning models (including the algorithm parameter tuning) was performed using a 10-fold person-independent cross-validation approach.

Binary classification performance measures were adopted according to the standards reported in [3]. Five different machine-learning methods were used to train, validate and test the classifiers (SVM, MLP, IBK, C4.5 and LDA); the model was chosen as the classifier achieving the highest Area under the Curve (AUC), which is a reliable estimator of both sensitivity and specificity rates. The model was then tested on the videogame dataset (i.e., 42 subjects).

III. Results

A. Short vs ultra-short in real life

From the analysis of experiment 1, few HRV features showed to be reliable in 1 min and 30 sec HRV excerpts compared to those in 5 min HRV excerpts. Therefore, only the 2-min HRV features that showed to be good surrogates of short HRV features (5 min) for the detection of mental stress were considered in experiment 2 and 3.

7 ultra-short HRV features showed to be good surrogates of short HRV features to detect acute mental stress at 2 min (Fig. 2), as they displayed consistency across all the excerpt lengths (i.e. from 5 min to 2 min). Moreover, from the HRV features pooled in the initial systematic review (see Table 1), we found 4 ultra-short HRV features (MeanNN, StdNN, HF and LF/HF) showing similar trends.

B. Acute mental stress detection in lab settings

HRV feature selection performed on the CWT dataset in folder 1 (52 subjects) gave various combination of features, but in line with the results of experiment 1, only the combination containing the HRV features that showed to be good surrogates of ultra-short term was retained. MeanNN, LF and RPlmean were selected in order to develop a robust classifier using HRV features at 2 min.

The Support Vector machine (SVM) outperformed other data-mining methods achieving 70%, 60 %, 64% of sensitivity, specificity and accuracy, respectively, during the train and validation process. The SVM was then tested on the dataset of experiment 3 achieving 69 %, 60 % and 64% of sensitivity, specificity and accuracy, respectively. These performance resulted far below than those achieved with real-life stressors reported in [2, 3]. Yet, due to the design of the experiments, it was not possible to appreciate if this is due to the shortening of the HRV excerpts or to the stressor (i.e., real-life vs lab).

IV. Conclusions

Currently, short term HRV analysis is a reliable option to detect mental stress in healthy subjects. However, several studies are currently attempting to embed HRV analysis in wearable or portable devices by reducing the length of HRV excerpts. Therefore, there is the need to investigate HRV analysis in ultra-short time (less than 5 min). Results reported in this paper showed that 7 ultra-short HRV features can be

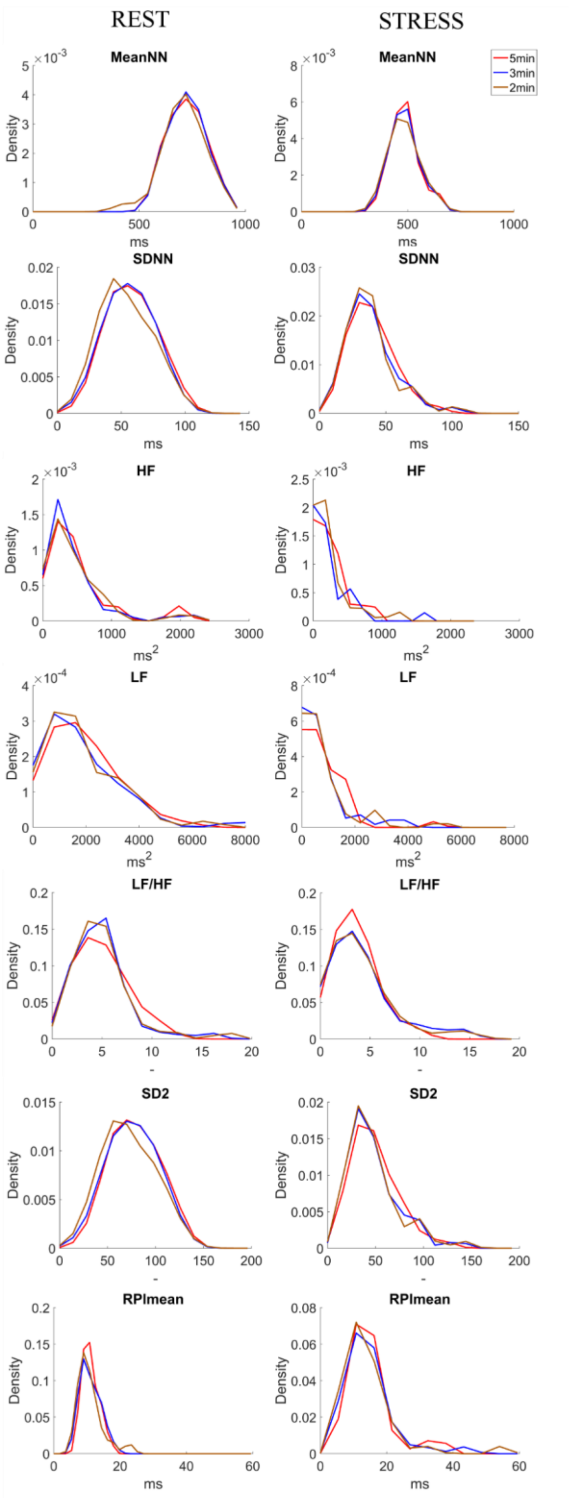

Fig. 2 Probability distribution of inter-group HRV features during rest (left) and stress (right) elicited by real stressor (i.e., verbal examination)

used in ultra-short term analysis to detect mental stress in real life. This is a noteworthy result for both theory and practice. However, as explained throughout in this paper, reduction in the length of HRV excerpts may affect their statistical significance in investigating mental stress. Yet, the results reported in this paper show a significant performance drop in the detection of mental stress if compared to previously published studies. We call for further research in investigating the opportunities of shortening HRV as well as in translating in-lab results to real life situations.

Acknowledgment

R.C. thanks University Of Warwick Departmental S'Ships/Burs's. S.M. acknowledge funding from the Global Research Priority in Behavioural Science of the University of Warwick.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] R. Castaldo, P. Melillo, U. Bracale, M. Caserta, M. Triassi, and L. Pecchia, "Acute mental stress assessment via short term HRV analysis in healthy adults: A systematic review with meta-analysis," *Biomedical Signal Processing and Control*, vol. 18, pp. 370-377, 2015.

[2] R. Castaldo, L. Montesinos, P. Melillo, and L. Pecchia, "Ultra-short term HRV Features as Surrogate of Short term HRV. A Case Study on Mental Stress Detection in Real Life," *IEEE journal of biomedical and health informatics* (under review) 2017.

[3] P. Melillo, M. Bracale, and L. Pecchia, "Nonlinear Heart Rate Variability features for real-life stress detection. Case study: students under stress due to university examination," *Biomedical engineering online*, vol. 10, p. 1, 2011.

[4] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark*, et al.*, "Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, pp. e215-e220, 2000.

[5] M. P. Tarvainen and J.-P. Niskanen, "Kubios HRV User's Guide," *Biosignal Analysis and Medical Imaging Group (BSAMIG), Department of Physics University of Kuopio*, 2013.

[6] T. Force, "Heart rate variability guidelines: Standards of measurement, physiological interpretation, and clinical use," *European Heart Journal*, vol. 17, pp. 354-381, 1996.

[7] S. Massaro and L. Pecchia, "Heart Rate Variability (HRV) Analysis A Methodology for Organizational Neuroscience," *Organizational Research Methods*, p. 1094428116681072, 2016.

[8] R. Castaldo, P. Melillo, R. Izzo, N. D. Luca, and L. Pecchia, "Fall Prediction in Hypertensive Patients via Short-Term HRV Analysis," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, pp. 399-406, 2017.