35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Comparing Stress Markers Across Various Cohorts in a Mobile Setting

Shreyans Gandhi[1], Jawahar Jain[2], Maryam Shojaei Baghini[1] and Soumyo Mukherji[1]

Abstract— **There exist multiple markers for measuring psychological stress, with varying specificities and sensitivities. However, in a real-life setting, there is limited data on how robust these methods may be especially in a relatively mobile context where the signal fidelity maybe limited. Thus any large scale data to inform how these methods perform, using commonly available sensors, based on both context and cohort characterization, can greatly add to our knowledge of their respective utility in real-life settings. This paper presents a study of 253 subjects which provides crucial data for analysing various stress markers in a mobile setting. We also provide early analysis results.**

Index Terms— **Mental Stress, Heart rate variability; Time-Domain HRV; Frequency-Domain HRV; HR; pNN50; SDNN; RMSSD; LF/HF**

I. INTRODUCTION

It is well known that through a variety of biological and behavioural mechanisms, chronic stress can cause a range of physiological and psychological problems over time such as hypertension, diabetes, cardiovascular disease, gastrointestinal problems, mood disorders, substance abuse etc. To manage and/or monitor such disorders, it is essential to have effective methods to measure stress. Several methods for measuring stress have been proposed including Heart Rate Variability [1], salivary cortisol, questionnaire [2], facial expressions, pupil diameter, voice analysis [3], skin conductance and skin temperature. Though there has been a lot of research on methods to quantify stress, the relative preference of those methods is still debatable. Many of these methods exhibit limited sensitivity and convergent validity. The most widely used methods are based on Heart Rate Variability. Both time domain and frequency domain approaches are widely used but limited agreement exists as to superiority of a particular approach when deployed in real-life ambulatory settings.

The most prevalent are pNN50, SDNN and RMSSD in the time domain and LF-HF method in the frequency domain [1]. Time-domain methods are computationally simple and robust under artifacts and measurement errors but are believed to be less sensitive [4] as well as lack the ability to discriminate between complex physiological etiologies to HRV such as relative contribution of various components of autonomic nervous systems (sympathetic and parasympathetic) [4]. Spectral analysis via LF/HF has been shown to be more sensitive and descriptive but has extremely low resilience to artifacts, especially in the HF part of the spectrum [5] and requires longer measurement epochs which hampers their applicability to acute forms of stress. There are also several contradictory results such as in [1], [6], [7] where it is established that LF/HF directly correlates to Stress, while in [8] no particular correlation was observed between such a measure and the short-term stress.

In the time domain methods, since proposed in 1984 by Ewing [9], pNNx (x=50) has been the dominant model. However recent results [10] show lower pNNx (x<20) to be a better marker, which is unfortunately again contradicted by Hutchinson [11]. In general, there does not appear to be a consensus on a single method for quantifying HRV in the context of stress, especially in a real-life setting. This is partly due to the fact that different methods may be more applicable for different problems. as well as lack of adequate sized data, accurately collected in a controlled setting, but arguably also due to the lack of a reference benchmark for stress. These contradictory results urge one to examine the factors on which HRV-based stress measurements depend on and in which way these factors affect the inferred results. Finally, the accuracy of the methods also depends on genotypical and phenotypical attributes since these factors inform personal adaptation to chronic stress.

We especially want to note that historically stress research has been conducted in a medical setting where hospital quality equipment was used to collect bio-data, typically ECG. The analysis results do not generalize to real-life and mobile settings, where commonly available, relatively inexpensive, sensors are used, since there is a large amount of noise in ECG data including dropped packets; settings in which noise resilience, instead of sensitivity, becomes the key advantage of a stress model. We present protocol and outlines of a study on 253 subjects whose stress was measured in a carefully controlled setting via analysis of biomarkers collected by sensitive but readily available mobile sensors and a powerful, real-time, data aggregation and analysis platform. We believe that our trial design is instructive for such endeavours for data collection - a key weakness hampering current state of the art - and our experimental analysis adds to the knowledge of the factors on which the HRV based stress models depends on, making it possible to compare these markers for a variety of real-life and experimental settings. Due to space constraints, this paper only presents broad outlines of the study undertaken, and early analysis results in comparing the popular methods for quantifying HRV to study psychophysiological response to stress.

A. Time Domain Based Markers

The most widely used time domain based HRV markers are SDNN, SDANN, pNN50, pNN20 and RMSSD. following is a short description of them:

1) SDNN: SDNN is the standard deviation of all normal NN intervals (intervals measured between consecutive sinus beats.) The normal range of SDNN depends on the record length. Generally SDNN is calculated as the mean of all 5-minute standard deviations of NN intervals during a 24-hour period.

2) pNNx: pNNx is the ratio of the consecutive NN intervals differing by more than "x" milliseconds, divided by the total number of NN intervals. pNNx is derived from NN50, first introduced by Ewing et al. [9], and later converted to pNN50 by Bigger et al. [12].

3) RMSSD: RMSSD is equal to the square root of the mean squared difference of successive NNs. It is also a widely used HRV marker [13].

B. Frequency Domain Based Markers

The most widely used frequency domain based HRV marker is LF/HF [14]. This marker is derived from the spectral analysis of the

[1]S. Gandhi, M. S. Baghini and S. Mukherji are with Indian Institute of Technology Bombay, Mumbai, India shreyans@iitb.ac.in; mshojaei@ee.iitb.ac.in; mukherji@iitb.ac.in

[2]Jawahar Jain was with Fujitsu Laboratories of America, Sunnyvale, CA USA

heart rate, where LF is the low frequency region ranging from 0.04-0.15Hz and HF ranging from 0.15-0.4Hz. The power distribution or the frequencies allotted for the LF and HF bands are not necessarily same. The variations in the bands are caused by the autonomic modulation of heart period [15].

II. PROTOCOL

A. Subject Population

This study was approved by the Birla Hospital and Research Center Ethics committee at Satna, India. Informed consent was obtained from all subjects participating in the study. The target population for this study was males belonging to healthy, (pre)diabetic, (pre)hypertensive, or in a mixed syndrome category. No subjects with other form of active diseases were specifically targeted. A sample of 253 subjects (average age $\pm$ SD= 48$\pm$11.59, average BMI $\pm$ SD = 25$\pm$4.16). The sample was distributed such that the study was not centered or evaluated on any age group, to study stress across a wide variety of population. All subjects were compensated in form of free blood tests as well as full body checkup whose detailed report was provided to each of them.

B. Measurements

The study targeted determining co-relation of stress with key markers of wellness, especially key cardiometabolic riskmarkers (hypertension, blood glucose, lipids, obesity, etc.). This paper only details relevant stress quantification. The subjects were grouped in different categories based upon an initial screening by the hospital physicians which was confirmed and refined by subsequent medical tests (pathology testing and blood pressure measurements). Per protocol, subjects were called fasting at 9AM for a blood draw (fasting blood glucose, glycosylated hemoglobin, and lipids). This phase was followed by clinical examination that included a detailed medical history along with resting heart rate and resting, average, blood pressure by automated oscillometric method. In the subsequent phase, an array of anthropomorphic measurements were made to quantify known markers of obesity. This included BMI, abdominal, waist and hip circumference as well as subcutaneous fat measurement via four point skin fold calipers. Also, an advanced Omron smart weighing scale [16] was used to determine fat and muscle composition, including visceral fat, using bio-electrical impedance.

C. Mobile Psychophysiological Monitoring

Following, the subjects were wired with the Zephyr BioHarness chest strap [17], synchronized to our experimental mobile healthcare platform named "Sprout". The strap containing the dry electrodes is worn on the chest and the embedded sensor transmits data using bluetooth to the sprout. Sprout collects, timestamp and stores the data. Any device connected to the sprouts' wifi network can view the data transmitted in real time. The following biomarkers were collected: EKG, Respiration, Galvanic skin response (GSR), Skin temperature, Posture and Movement via accelerometry. All sensor data were wirelessly transmitted to the mobile platform. The subjects' monitoring using this platform was divided into three phases - normal, stress and relaxation.

The normal (baseline) phase was of 40 minutes duration and consisted of the subject sitting comfortably on a couch with newspapers and magazines to read. The stress phase of the monitoring was of 20-minute duration. During this period, a reasonably difficult analytical test was given to the subjects, which had to be solved in the stipulated duration. The score would yield a purported financial reward of free medical examination including stress/relaxation profile.[1] An examiner would intermittently and professionally but curtly remind the remaining time. A stopwatch was placed in front of the subject, a continual reminder of the dwindling time.The method of stress induction was similar to the methods used by [7]. The relaxation phase was of 10 minutes duration in which subjects asked to practice deep breathing for 5 minutes followed by meditation for 5 minutes.

III. DATA DRIVEN HEALTH CARE PLATFORM

Our data-driven healthcare platform (Sprout platform) has been developed for real-time biomarker data collection and analysis. It comprises three components: a mobile data aggregation and analysis device; a software stack for real-time recording, synchronization, sanitization and analysis of multiple wired or wireless bio-sensor data streams; and a collection of sophisticated applications for real-time data visualization which enables active monitoring of protocol compliance.

A. Sprout Mobile Device

The Sprout hardware is based on a powerful 600MHz ARM Cortex A8 CPU with 512MB of RAM and utilizes microSD memory cards for data storage. It was designed for continuous, interrupted recordings and can connect to biosensors via 3 Wireless protocols (specifically Bluetooth 2.1, 802.11bg WiFi and TI low-power SimpliciTI RF) as well as three USB 2.0 host ports and 3 analog ports. A total of 15 off-the-shelf sensors (including ECG, SpO2, BP, breathing, pFlow, Galvanic skin Response, temperature, weight-scale, posture and accelerometry sensors) have been interfaced to it.

The Sprout software stack is built on top of Linux 3.0. It provides abstractions for integrating hardware sensors as well as virtual (meta) sensors. It simplifies the creation of meaningful analytics by relieving the developer from low-level hardware details and simplifies the task of data cleanup and synchronization. Analytics metasensors continuously run on the mobile device and are stored for further analysis or visualization. The platform facilitates starting / stopping meta-sensors at real-time without affecting the rest of the system. The device runs web services through which a user can observe the recorded datasets at real-time via web applications as well as native iOS and Android applications.

IV. RESULTS

We present our results on Table I and Figure 1. We report means and standard deviations for four biomarkers, namely pNN50, SDNN, RMSSD, LF/HF. Results are presented on the complete sample size, as well as across different cohorts, based on health conditions. Specifically the cohorts (253 subjects) details are: Normal (54), Diabetic-Hypertensive (41), Diabetic-Prehypertensive (24), Only-Diabetic (29), Only-Hypertensive (25), Only-Prediabetic (26), Only-Prehypertensive (33), Prediabetic-Hypertensive (11) and Prediabetic-Prehypertensive (10).

As is evident from the data presented, different markers may perform better for different stress / relaxation phases and across different cohorts. There is no single method that is able to clearly quantify acute stress across all cohorts and conditions.

Figure 1 gives ROC (Receiver Operator Characteristic) graphs which show how sensitivity varies with the threshold (minimum change, per subject, from normal phase for the event to be tagged

[1] different versions of tests were at hand to suit a wide dispersion in the education profiles of the participants.

	Cohort	pNN50		SDNN		RMSSD		LF/HF	
		μ	σ	μ	σ	μ	σ	μ	σ
NORMAL PHASE	ALL	0.08	0.12	188.8	375.1	25.52	15.33	10.41	23.81
	N	0.11	0.135	114.628	109.254	28.545	15.696	2.085	2.218
	DH	0.063	0.101	252.438	19.119	25.969	14.528	4.289	7.037
	DPH	0.057	0.123	116.016	124.352	23.05	17.093	2.538	1.986
	OD	0.049	0.082	143.93	104.727	22.233	13.285	2.074	1.025
	OH	0.085	0.102	172.091	142.251	28.864	14.526	2.068	1.675
	OPD	0.074	0.153	103.203	86.281	25.683	21.087	80.13	9.853
	OPH	0.061	0.075	126.71	98.396	23.626	11.953	1.956	0.943
	PDH	0.065	0.113	143.006	108.377	22.93	16.841	4.589	7.776
	PDPH	0.058	0.055	75.969	29.27	23.46	10.177	2.29	0.918
STRESS PHASE	ALL	0.07	0.12	124.08	237.35	24	17.43	10.32	26.18
	N	0.1	0.133	92.344	81.691	27.612	17.321	1.045	0.797
	DH	0.067	0.128	136.646	19.119	24.617	20.596	2.046	1.976
	DPH	0.057	0.13	82.763	87.746	21.632	19.677	1.87	1.542
	OD	0.052	0.096	110.958	108.142	20.431	14.801	1.685	0.924
	OH	0.079	0.125	103.312	84.727	24.065	16.41	3.893	6.237
	OPD	0.058	0.141	82.723	77.641	23.129	20.731	87.516	12.495
	OPH	0.069	0.088	120.962	89.724	25.074	14.128	1.24	1.211
	PDH	0.025	0.035	82.723	78.334	17.851	10.195	1.561	0.865
	PDPH	0.043	0.067	134.873	211.457	20.826	10.95	1.093	0.437
DEEP BREATHING PHASE	ALL	0.11	0.14	114.37	120.77	29.27	18.89	9.55	24.9
	N	0.154	0.157	105.974	92.448	33.417	18.197	1.100	0.742
	DH	0.108	0.146	93.276	91.792	29.646	21.584	1.806	1.436
	DPH	0.128	0.173	117.012	109.135	31.798	22.801	1.247	1.187
	OD	0.063	0.090	90.507	89.649	23.203	13.657	1.623	1.209
	OH	0.076	0.104	138.678	124.643	24.084	14.145	1.656	1.357
	OPD	0.084	0.167	132.042	231.522	27.310	22.022	82.761	14.845
	OPH	0.108	0.121	144.915	116.943	29.905	17.679	1.241	0.724
	PDH	0.155	0.204	125.745	114.522	35.765	22.974	1.413	0.925
	PDPH	0.050	0.035	68.484	15.569	23.394	7.461	1.483	0.673
MEDITATION PHASE	ALL	0.09	0.13	139.36	378.5	26.78	17.36	9.31	23.1
	N	0.131	0.141	86.914	78.188	29.995	15.5	1.426	1.253
	DH	0.082	0.119	266.631	819.937	26.638	17.328	2.536	3.413
	DPH	0.058	0.125	73.678	63.394	22.573	18.594	1.577	1.021
	OD	0.055	0.099	109.214	160.215	23.518	14.402	1.413	0.987
	OH	0.087	0.127	145.57	157.761	26.145	15.901	1.965	1.781
	OPD	0.084	0.182	92.411	87.305	25.793	21.623	77.479	11.27
	OPH	0.086	0.108	191.614	458.668	27.474	17.055	1.509	1.041
	PDH	0.163	0.21	154.172	127.031	34.155	27.625	1.687	1.475
	PDPH	0.043	0.037	75.798	30.959	22.338	7.089	1.719	0.899

TABLE I

MEDICAL CONDITION BASED SEGREGATION OF SUBJECTS AND PERFORMANCE OF VARIOUS STRESS MARKERS IN NORMAL, STRESS, DEEP BREATHING AND MEDITATION PHASES. LEGEND: N = NORMAL, D = DIABETIC, H = HYPERTENSIVE, P* = PRE-*, O*=ONLY-*.

as a stress event for a given marker). The ROC graphs can be used to measure the allowed threshold (and hence tolerable noise level) for a desired sensitivity. Our data shows that such thresholds indeed differ between various markers. It is clear that the popular frequency domain method LF/HF does not offer enough noise resilience or sensitivity in the mobile context. LF/HF does not perform as well to quantify stress as even a small amount of noise can significantly corrupt the FFT results. Further, our results suggest that in absence of sustained acute stress, the sympathovagal balance cannot be easily studied by frequency spectrum analysis especially when the measurement windows are very large (20 minutes stress measurement in our case). Time domain methods appear more consistent though the most popular method, pNN50 is not robust. A cutoff of 50ms appears excessive and can lead to a situation where either the marker is unable to capture any stress since there are no corresponding data points, or arbitrary (becoming a very sensitive marker for stress owing to the fact that even a modest increase in the samples will have a disproportional effect on corresponding analysis.) Time domain methods of SDNN and RMSSD appear more stable markers, effective across a variety of cohorts (with the SDNN method having better sensitivity at higher thresholds), and suitable for real-life application especially in a mobile context. To develop a more robust marker, real time filtering and signal stitching should be considered. A study on the sensitivity of differnet markers towards real-life noise and distortions can be helpful. Present methods of HRV measurement are based on the disease quantifying markers. A new measure of HRV, which is more robust and sensitive for stress measurement can also be explored.

V. CONCLUSION

There is a paucity of large scale trial data comparing effectiveness of academically established stress markers. Further, there is minimal information about their robustness in a mobile setting using wear-

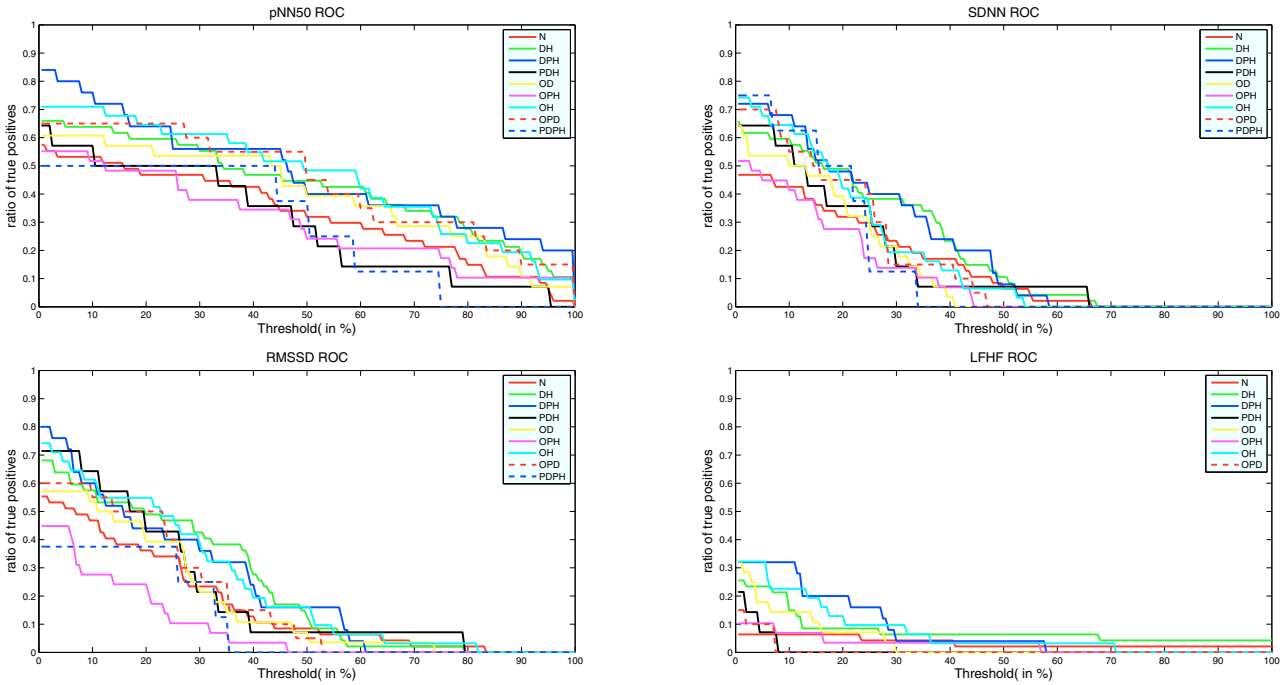

Fig. 1. ROC graphs for pNN50, SDNN, RMSSD and LF/HF.

able and affordable sensors. This paper presented a study of 253 subjects which provides crucial data to address above shortcomings, and overview of analysis results comparing the most popular time-domain and frequency-domain Heart Rate Variability biomarkers.As is evident from the data presented, different markers may perform better for different stress / relaxation phases and across different cohorts. There is no single method that is able to clearly quantify acute stress across all cohorts and conditions. Time domain methods were found to be more robust and effective than frequency domain approaches. A better understanding of how popular stress models perform based on both context as well as cohort characterization should have a significant impact in increasing their utility in real-life settings. We hope that this comparison will help characterize the relative performance of these markers and act as a reference point for further developing more robust stress markers. Such a marker would combine robustness with its capability to capture physiological characteristic of diverse disease traits.

REFERENCES

[1] L. Salahuddin, J. Cho, M. Jeong, and D. Kim, "Ultra short term analysis of heart rate variability for monitoring mental stress in mobile settings," in *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*. IEEE, 2007, pp. 4656–4659.

[2] A. Hogh, Å. Hansen, E. Mikkelsen, and R. Persson, "Exposure to negative acts at work, psychological stress reactions and physiological stress response," *Journal of Psychosomatic Research*, 2012.

[3] X. Yao, T. Jitsuhiro, C. Miyajima, N. Kitaoka, and K. Takeda, "Physical characteristics of vocal folds during speech under stress," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*. IEEE, 2012, pp. 4609–4612.

[4] D. Dutt and S. Krishnan, "Computer processing of heart rate variability signals for detection of patient status in cardiac care units," *CURRENT SCIENCE-BANGALORE-*, vol. 78, no. 7, pp. 864–868, 2000.

[5] C. Peters, R. Vullings, M. Rooijakkers, J. Bergmans, S. Oei, and P. Wijn, "A continuous wavelet transform-based method for time-frequency analysis of artefact-corrected heart rate variability data," *Physiological Measurement*, vol. 32, no. 10, p. 1517, 2011.

[6] M. Pagani, N. Montano, A. Porta, A. Malliani, F. Abboud, C. Birkett, and V. Somers, "Relationship between spectral components of cardiovascular variabilities and direct measures of muscle sympathetic nerve activity in humans," *Circulation*, vol. 95, no. 6, pp. 1441–1448, 1997.

[7] U. Nater, R. La Marca, L. Florin, A. Moses, W. Langhans, M. Koller, and U. Ehlert, "Stress-induced changes in human salivary alpha-amylase activity–associations with adrenergic activity." *Psychoneuroendocrinology*, 2006.

[8] C. Schubert, M. Lambertz, R. Nelesen, W. Bardwell, J. Choi, and J. Dimsdale, "Effects of stress on heart rate complexityâĂŤa comparison between short-term and chronic stress," *Biological psychology*, vol. 80, no. 3, pp. 325–332, 2009.

[9] D. Ewing, J. Neilson, and P. Travis, "New method for assessing cardiac parasympathetic activity using 24 hour electrocardiograms." *British heart journal*, vol. 52, no. 4, pp. 396–402, 1984.

[10] J. Mietus, C. Peng, I. Henry, R. Goldsmith, and A. Goldberger, "The pnnx files: re-examining a widely used heart rate variability measure," *Heart*, vol. 88, no. 4, pp. 378–380, 2002.

[11] T. Hutchinson, "Statistics and graphs for heart rate variability: pnn50 or pnn20?" *Physiological Measurement*, vol. 24, no. 3, pp. N9–N14, 2003.

[12] J. Bigger, R. Kleiger, J. Fleiss, L. Rolnitzky, R. Steinman, and J. Miller, "Components of heart rate variability measured during healing of acute myocardial infarction," *The American journal of cardiology*, vol. 61, no. 4, pp. 208–215, 1988.

[13] J. Nolan, P. Batin, R. Andrews, S. Lindsay, P. Brooksby, M. Mullen, W. Baig, A. Flapan, A. Cowley, R. Prescott *et al.*, "Prospective study of heart rate variability and mortality in chronic heart failure: results of the united kingdom heart failure evaluation and assessment of risk trial (uk-heart)," *Circulation*, vol. 98, no. 15, pp. 1510–1516, 1998.

[14] L. Salahuddin, M. Jeong, D. Kim, S. Lim, K. Won, and J. Woo, "Dependence of heart rate variability on stress factors of stress response inventory," in *e-Health Networking, Application and Services, 2007 9th International Conference on*. IEEE, 2007, pp. 236–239.

[15] H. Tsuji, M. Larson, F. Venditti, E. Manders, J. Evans, C. Feldman, and D. Levy, "Impact of reduced heart rate variability on risk for cardiac events: the framingham heart study," *Circulation*, vol. 94, no. 11, pp. 2850–2855, 1996.

[16] Omron, http://www.omronwebstore.com/detail/OMR+HBF%2D516B, 2012, [Online; accessed 15-Dec-2012].

[17] Z. Inc., "Zephyr bioharness," http://www.zephyr-technology.com/products/bioharness_bt, 2012, [Online; accessed 15-Dec-2012].