MBEC2014
6th European Conference of the International Federation for Medical and Biological Engineering

MBEC2014
Towards new horizons in biomedical engineering
7-11 September 2014
Dubrovnik, Croatia

Automatic prediction of falls via Heart Rate Variability and data-mining in hypertensive patients: the SHARE project experience

P. Melillo[1], A. Jovic[4], N. De Luca[4], S. Morgan[5], L. Pecchia[3,6]

[1] **SHARE Project,** Italy
[2] **University of Zagreb,** Croatia
[3] **University of Naples,** Italy
[4] **University of Nottingham,** The United Kingdom
[5] **University of Warwick,** The United Kingdom

paolomelillo85@gmail.com

Corresponding author: Paolo Melillo. paolo.melillo85@gmail.com

INTRODUCTION

PROGETTO SHARE

- **Falls represent a big problem:**
  - 28-35% of community-dwelling older people (over 64 yrs) fall each year;
  - about 28 per cent of falls result in injuries requiring some form of medical treatment;
  - The mean and median costs for a fall are about 9,000 and 11,000 euro;

- **Several studies identify different risk factors and methods for risk assessment:**
  - inertial-sensor-based fall risk measures;
  - functional mobility tests;
  - and other clinical assessment tools.

- Limited evidence about depressed Heart Rate Variability as a risk factor for falls

Corresponding author: Paolo Melillo. paolo.melillo85@gmail.com

GOALS

PROGETTO SHARE

- **Goal of the SHARE Project** is to develop a system to automatically assess the risk of cardiovascular events and falls
- **In this study, we presents the SHARE experience:**
  - **for automatic identification of fallers**;
  - adopting analysis of Heart Rate Variability (HRV);
  - using an ad hoc database of ECG holter signals from hypertensive patients;
  - using data-mining methods.

Corresponding author: Paolo Melillo. paolo.melillo85@gmail.com

SAMPLE DATA and HRV

- **Ad hoc database of hypertensive patients:**
  - **168 subject (60 female and 108 male) aged 55 and over**
    - 47 experienced a falls within three month from the recording
    - 101 free of recorded falls

- **HRV linear and non-linear analysis**
  - concurrent analysis of 30-min segments

Table 2 Selected frequency domain measures of HRV

Variable	Units	Description	Frequency range
		Analysis of short-term recordings (5 min)	
5 min total power	ms^2	The variance of NN intervals over the temporal segment	approximately $\leq 0{\cdot}4$ Hz
VLF	ms^2	Power in very low frequency range	$\leq 0{\cdot}04$ Hz
LF	ms^2	Power in low frequency range	0·04–0·15 Hz
LF norm	n.u.	LF power in normalised units LF/(Total Power–VLF) × 100	
HF	ms^2	Power in high frequency range	0·15–0·4 Hz
HF norm	n.u.	HF power in normalised units HF/(Total Power–VLF) × 100	
LF/HF		Ratio LF [ms^2]/HF [ms^2]	
		Analysis of entire 24 h	
Total power	ms^2	Variance of all NN intervals	approximately $\leq 0{\cdot}4$ Hz
ULF	ms^2	Power in the ultra low frequency range	$\leq 0{\cdot}003$ Hz
VLF	ms^2	Power in the very low frequency range	0·003–0·04 Hz
LF	ms^2	Power in the low frequency range	0·04–0·15 Hz
HF	ms^2	Power in the high frequency range	0·15–0·4 Hz
α		Slope of the linear interpolation of the spectrum in a log-log scale	approximately $\leq 0{\cdot}04$ Hz

Table 1 Selected time-domain measures of HRV

Variable	Units	Description
		Statistical measures
SDNN	ms	Standard deviation of all NN intervals.
SDANN	ms	Standard deviation of the averages of NN intervals in all 5 min segments of the entire recording.
RMSSD	ms	The square root of the mean of the sum of the squares of differences between adjacent NN intervals.
SDNN index	ms	Mean of the standard deviations of all NN intervals for all 5 min segments of the entire recording.
SDSD	ms	Standard deviation of differences between adjacent NN intervals.
NN50 count		Number of pairs of adjacent NN intervals differing by more than 50 ms in the entire recording. Three variants are possible counting all such NN intervals pairs or only pairs in which the first or the second interval is longer.
pNN50	%	NN50 count divided by the total number of all NN intervals.
		Geometric measures
HRV triangular index		Total number of all NN intervals divided by the height of the histogram of all NN intervals measured on a discrete scale with bins of 7·8125 ms (1/128 s). (Details in Fig. 2)
TINN	ms	Baseline width of the minimum square difference triangular interpolation of the highest peak of the histogram of all NN intervals (Details in Fig. 2.)
Differential index	ms	Difference between the widths of the histogram of differences between adjacent NN intervals measured at selected heights (e.g. at the levels of 1000 and 10 000 samples)[21].
Logarithmic index		Coefficient φ of the negative exponential curve $k \cdot e^{-\varphi t}$ which is the best approximation of the histogram of absolute differences between adjacent NN intervals[22].

Nonlinear	SD1, SD2	[ms]	The standard deviation of the Poincaré plot perpendicular to (SD1) and along (SD2) the line-of-identity
	ApEn		Approximate entropy
	SampEn		Sample entropy
	D_2		Correlation dimension
	DFA		Detrended fluctuation analysis:
	α_1		Short term fluctuation slope
	α_2		Long term fluctuation slope
	RPA		Recurrence plot analysis:
	Lmean	[beats]	Mean line length
	Lmax	[beats]	Maximum line length
	REC	[%]	Recurrence rate
	DET	[%]	Determinism
	ShanEn		Shannon entropy

Corresponding author: Paolo Melillo. paolo.melillo85@gmail.com

DATA-MINING METHODS

PROGETTO SHARE

- **Data-mining methods** adopted to extract intelligible rules
  - C4.5
  - CART
  - Naïve Bayes
  - Lift chart
  - RIPPER
- **Odd Ratio** computation to evaluate the rules as risk factors

- **Data-mining methods** adopted to develop automatic classifiers:
  - Random Forest
  - Rotation Forest
  - AdaBoost
  - MultiBoost
  - RUSBoost
  - RUSBoost and Principal Component Analysis
- **Subject-based ROC curve analysis and 10-fold person-independent crossvalidation** to estimate the performances of automatic classifiers

Corresponding author: Paolo Melillo. paolo.melillo85@gmail.com

STATISTICAL ANALYSIS

PROGETTO SHARE

10 HRV features were found to be significantly different between non-fallers and fallers

Features	units	Non-fallers Median	Non-fallers 25th	Non-fallers 75th	Fallers Median	Fallers 25th	Fallers 75th	GEE p-value
RR_{MIN}	ms	246.10	203.10	515.60	398.40	210.90	640.60	<0.05
VLF_{WE}	ms^2	3.2E3	1.0E3	10.3E3	2.3E3	0.8E3	7.8E3	<0.05
TP_{WE}	ms^2	5.6E3	1.9E3	18.4E3	3.5E3	1.4E3	13.2E3	<0.05
LF_{LO}	ms^2	10E3	4E3	17.0E3	8.0E3	3.0E3	15.0E3	<0.01
TP_{LO}	ms^2	37E3	20E3	60.0E3	29.0E3	18.0E3	52.0E3	<0.05
l_{MAX}	beats	436.00	177.00	927.00	588.00	262.00	1158.00	<0.01
l_{MEAN}	beats	21.72	12.59	35.66	24.39	14.78	39.99	<0.01
DIV	$beats^{-1}$	2.30E-3	1.10E-3	5.60E-3	1.70E-3	0.9E-3	3.80E-3	<0.01
DET	-	1.00	0.98	1.00	1.00	0.99	1.00	<0.01
ShanEn	-	3.84	3.32	4.28	3.94	3.49	4.34	<0.01

RR_{MIN}: duration of the shorter beat to beat time interval
VLF_{WE}: very low frequency bandwidth (0-0.04 Hz) power, calculated with the Welch periodogram
LF_{LO}: low frequency bandwidth (0.04-0.15 Hz) power, calculated with the Lomb-Scargle periodogram
TP_{WE}: total power (0-0.4 Hz) calculated with the Welch periodogram
TP_{LO}: total power (0-0.4 Hz) calculated with the Lomb-Scargle periodogram
l_{MAX}: maximal length of lines of the Recurrence Plot
l_{MEAN}: mean length of lines of the Recurrence Plot
DIV: divergence parameter, Recurrence Plot
DET: determinism parameter, Recurrence Plot
ShanEn: Shannon Entropy

Corresponding author: Paolo Melillo. paolo.melillo85@gmail.com

Rules associated with significant falling odds

HRV pattern	OR (CI95%, p)
#1: '$\alpha_2 \leq 0.947$' & '$pNN50 \leq 26.7$' & '$RR_{MAX} \geq 2265.6$'	5.12 (1.42-18.41, <0.01)*
#2: '$l_{MAX} \geq 2179$' & '$SDNN_i \leq 40.4$'	3.66 (0.79-17.02, 0.08)
#3: '$pVLF_{LO} > 15.1$' & '$l_{MAX} > 2177$' & '$REC \leq 0.514$'	11.16 (1.21-102.6, <0.01)*
#1 or #2	4.98 (1.80-13.79, <0.01)*
#1 or #3	7.09 (2.31-21.76, <0.01)*
#2 or #3	2.72 (0.65-11.36, <0.01)*
#1 or #2 or #3	4.32 (1.61-11.56, <0.01)*

* Statistically significant with p<0.01

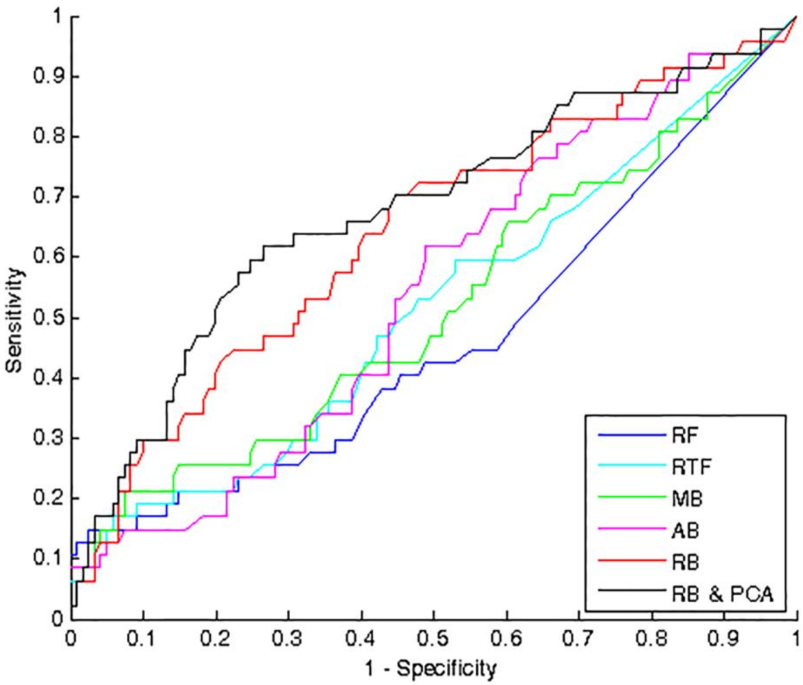
RR_{MAX}:	duration of the longer beat to beat time interval
l_{MAX}:	maximal length of lines of the Recurrence Plot
α_2:	long-term fluctuation slope in detrended fluctuation analysis
pNN50:	percentage of differences between adjacent NN intervals that are longer than 50 ms [11]
$SDNN_i$:	mean of the standard deviations of NN intervals in all 5-min segments [11]
$pVLF_{LO}$:	very low frequency bandwidth (0-0.04 Hz) power spectrum, calculated with Lomb-Scargle periodogram, percentage
REC:	recurrence rate in Recurrence Plot

Corresponding author: Paolo Melillo. paolo.melillo85@gmail.com

CLASSIFICATION

PROGETTO SHARE

Performance of the automatic classifiers



	AUC %	ACC %	SEN %	SPE %
RF	46.3	67.3	21.3	85.1
RTF	51.5	67.9	21.3	86.0
AB	51.7	68.5	25.5	85.1
MB	54.1	63.7	17.0	81.8
RB	63.9	69.0	40.4	80.2
RB PCA	67.6	72.0	51.1	80.2

RF: Random Forest

RTF:Rotation Forest

AB: AdaBoost

MB: MultiBoost

RB: RUSBoost

RB & PCA: RB enhanced with Principal Component Analysis

Melillo et al, MBEC, under revision

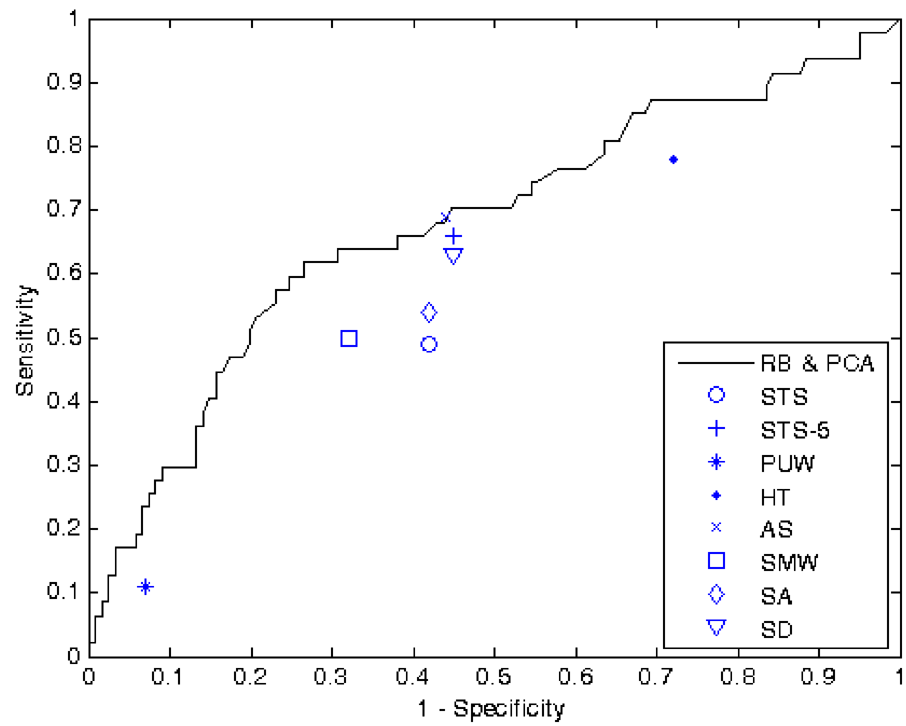
Corresponding author: Paolo Melillo. paolo.melillo85@gmail.com

DISCUSSIONS

PROGETTO SHARE

- **Good results in automatic identification of fallers**
  - Comparison with functional mobility tests (Tiedemann, Age ageing 2008);

Test	Relative risk (95% CI)
Sit-to-stand once (STS)	1.3 (0.8, 2.1)
Sit-to-stand five times (STS-5)	2.0 (1.3, 3.0)
Pick-up-weight test (PUW)	1.5 (0.8, 2.6)
Half-turn test (HT)	1.3 (0.8, 2.0)
Alternate-step test (AS)]	2.3 (1.4, 3.5)
Six-metre walk (SMW)	1.8 (1.2, 2.6)
Stair ascent (SA)	1.4 (1.0, 2.1)
Stair descent (SD)	1.7 (1.2, 2.6)
Abnormal HRV (best rule)	2.5 (1.5, 4.2)



Abnormal HRV was associated with a risk of falling 2.5 times higher

Corresponding author: Paolo Melillo. paolo.melillo85@gmail.com

CONCLUSIONS

- **We developed a system to automatically identify fallers** among hypertensive patients**:**
  - completely automatic;
  - using HRV analysis (only ECG monitoring required);
  - based on data-mining methods
- **Limits of this study:**
  - Dataset not specifically designed for falls (no information about other risk factors, only hypertensive patients, self-report of falls)
  - Small sample size (large Confidence Interval)
- **Further developments:**
  - Larger dataset
  - Longer follow-up period
  - New / other HRV indexes (i.e. point process time-frequency analysis)
  - Other non-invasive measurement (e.g. accelerometry, posture)
  - Integration in a web application (SHARE project web portal)

Corresponding author: Paolo Melillo. paolo.melillo85@gmail.com

Thank you!

Dr. Paolo Melillo

PI SHARE Project
Italy
paolomelillo85@gmail.com

Corresponding author: Paolo Melillo. paolo.melillo85@gmail.com