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 vascular events by Heart Rate Variability analysis in hypertensive patients

P. Melillo[1], P. Scala[1], N. De Luca[2], L. Pecchia[1,3]

[1] **SHARE Project,** Italy
[2] **University of Naples,** Italy
[3] **University of Warwick,** The United Kingdom

paolomelillo85@gmail.com

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

INTRODUCTION

PROGETTO SHARE

- **Vascular events**
  - include acute coronary, cerebrovascular, and peripheral vascular events
  - are the leading cause of premature death and disability in the developed countries;

- **Several studies proposed different risk factors for future vascular events:**
  - **anamnestic data;**
  - **echocardiography test** (Intima Media Thickness and Left Ventricular Mass Index);
  - **and other instrumental measures** (blood test, …).

- Limited **positive predictive value** of the previously identified risk factors

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

- **In this study, we presents classifiers to:**
  - **predict future vascular events** (within one year from recordings);
  - 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

PROGETTO SHARE

- **Ad hoc database of hypertensive patients:**
  - **142 subject aged 55 and over** (1 year follow-up)
    - 17 experienced a major vascular event
    - 125 free of vascular event

- **HRV linear and non-linear analysis**

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 ≤0·4 Hz
VLF	ms^2	Power in very low frequency range	≤0·04 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 ≤0·4 Hz
ULF	ms^2	Power in the ultra low frequency range	≤0·003 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 ≤0·04 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)[101]
Logarithmic index		Coefficient α of the negative exponential curve $k \cdot e^{-\alpha t}$ which is the best approximation of the histogram of absolute differences between adjacent NN intervals[102]

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

FEATURE SELECTION AND CLASSIFICATION

PROGETTO SHARE

- **Long-term recording** (concurrent analysis of all 30-minute segments)
- **Principal Component Analysis** to extract the most informative features
- **RUSBoost,** hybrid classification method (Undersampling and Boosting) to handle unbalanced dataset
- **subject-based ROC curve analysis and 10-fold person-independent crossvalidation** to estimate performance**:**
  - accuracy (ACC);
  - sensitivity (SEN);
  - specificity (SPE);
  - area under the curve (AUC).

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

FEATURE SELECTION AND CLASSIFICATION

PROGETTO SHARE

- **Short-term recording** (analysis of a 30-minute randomly chosen segment)
- **Oversamplig (SMOTE)** to handle small and unbalanced dataset
- **Comparison of different data-miming approach**:
  - Naïve Bayes classifier (NB);
  - Classification tree C4.5;
  - Random Forest (RF);
  - AdaBoost (AB);
  - Support Vector Machine (SVM);
  - Multilayer perceptron (MLP).
- **ROC curve and 10-fold crossvalidation** to estimate performance**:**
  - area under the curve (AUC);
  - accuracy (ACC);
  - sensitivity (SEN);
  - specificity (SPE).

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

RESULTS

CLINICAL FEATURES OF THE STUDY SAMPLE

PROGETTO SHARE

Measures	Low-risk subjects	High-risk subjects
Age (years)	71.4±7	74.1±6.5
Sex (female)	41 (33.6)	8 (47.1)
Family history of hypertension	41 (33.6)	7 (41.2)
Family history of stroke	10 (8.2)	3 (17.6)
Smoking	35 (28.7)	5 (29.4)
Diabetes	18 (14.8)	3 (17.6)
Diastolic Blood Pressure (mmHg)	76.3±9.1	73.5±8.4
Systolic Blood Pressure (mmHg)	136.6±19.5	141.7±23.5
Total Cholesterol (mg/dl)	175.7±35.1	182.9±42.7
IMT (mm)	2.3±0.7	2.4±1.1
LVMi (g/m2)	130.1±26.1	140.2±25.1
EF ()	59.3±10.9	57.8±13

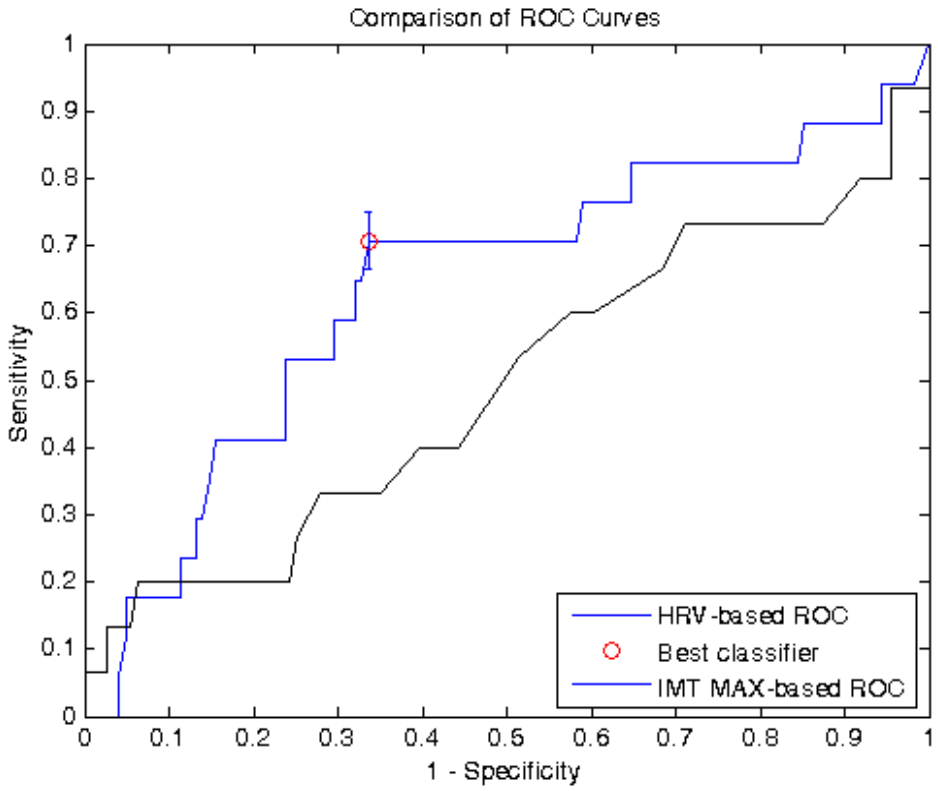
No significant differences in the baseline clinical features

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

CLASSIFICATION PERFORMANCE

PROGETTO SHARE

Long-term recording analysis



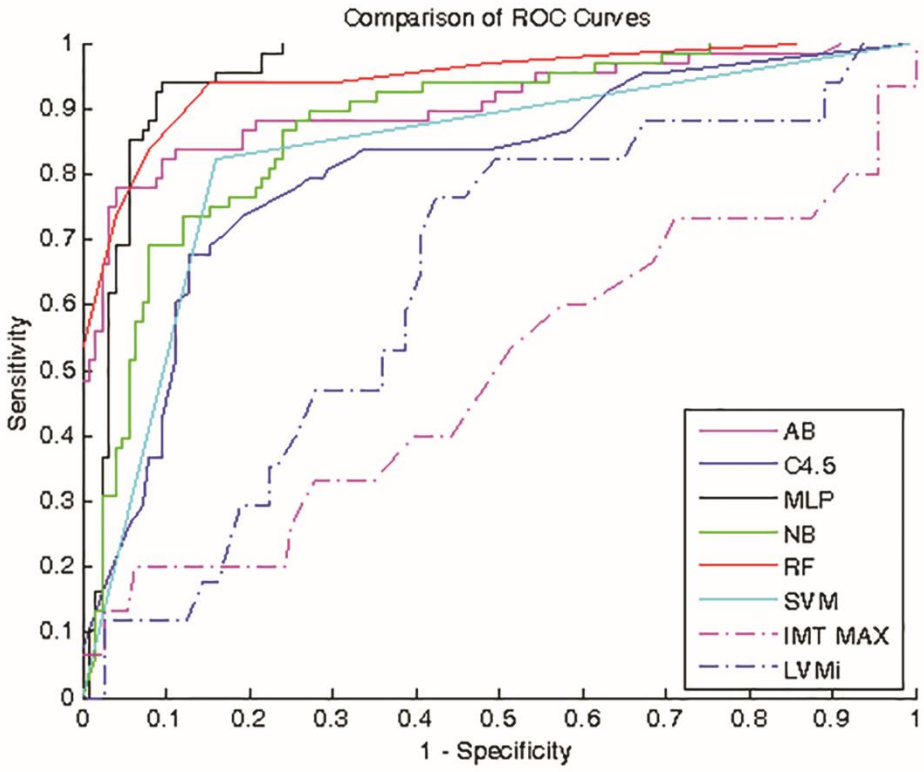
	AUC %	ACC %	SEN %	SPE %
HRV	64.0	67.0	70.6	66.4
IMT	49.0	57.9	40.0	60.3

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

CLASSIFICATION PERFORMANCE

PROGETTO SHARE

Short-term recording analysis



	AUC %	ACC %	SEN %	SPE %
MLP	95.4	90.2	94.1	88.0
RF	94.6	89.1	83.8	92.0
AB	90.9	87.0	77.9	92.0
NF	87.5	78.2	88.2	72.8
SVM	83.2	83.4	82.4	84.0
C4.5	80.8	78.8	60.3	88.8
LVMi	63.5	69.5	41.2	73.9
IMT	49.0	57.9	40.0	60.3

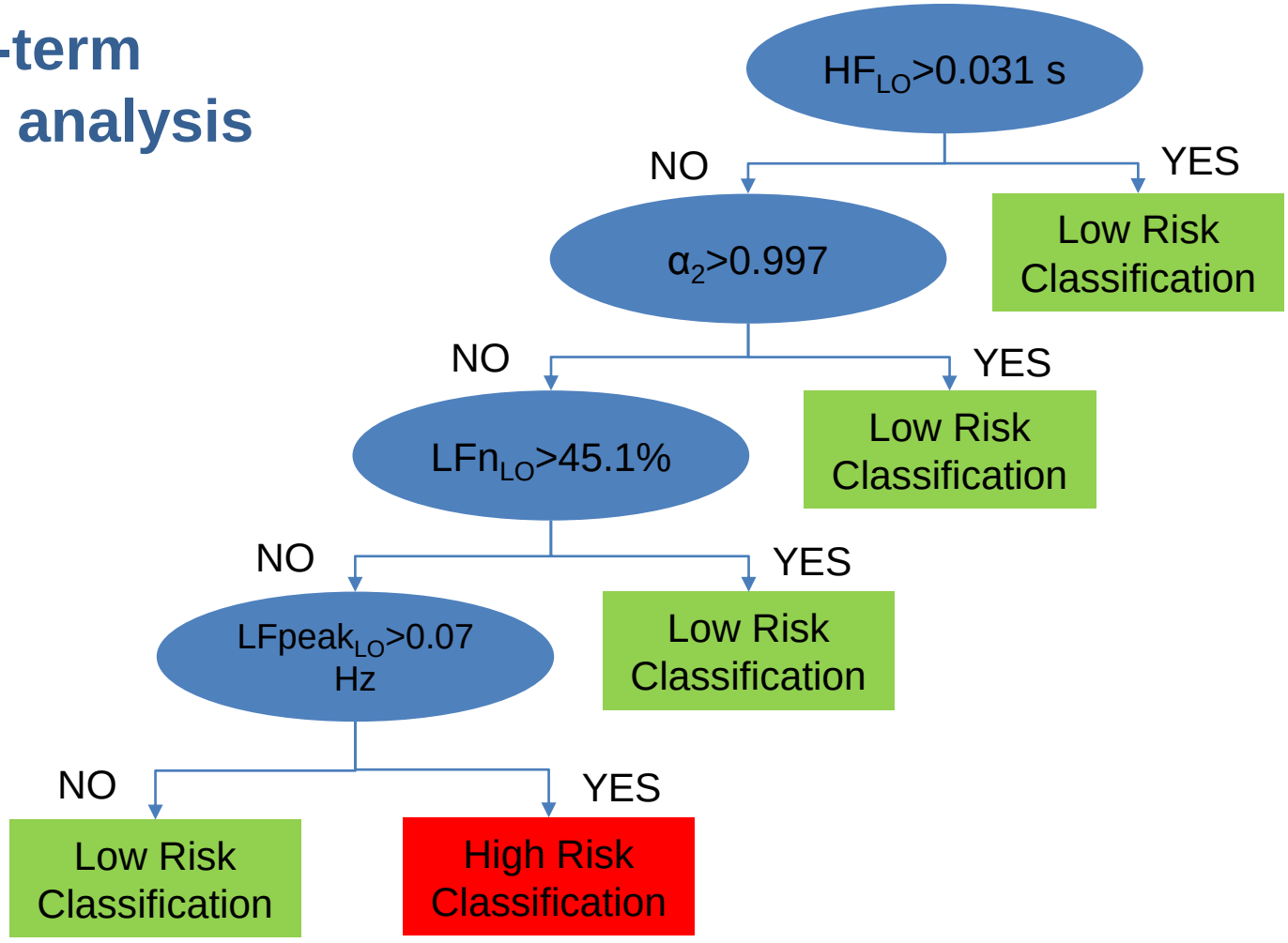
Melillo et al, Plos One, under revision

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

CLASSIFICATION MODEL

PROGETTO SHARE

Short-term recording analysis



Depressed HRV associated with High-Risk classification

Melillo et al, Plos One, under revision

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

DISCUSSIONS

PROGETTO SHARE

- **Good results in automatic risk assessment of future vascular events**
  - Most previous studies focused on risk factors and not on predictive models;
  - Few study focusing on HRV prediction of cardiac mortality in patients after acute events (acute coronary syndrome, acute myocardial infarction)
  - LVM and IMT are considered as powerful predictors of vascular events;
  - HRV-based classifiers showed better prognostic capacity compared with LVM and IMT

- **Limits of this study:**
  - Small sample size (no independent dataset for model selection evaluation)
  - Short follow-up length (twelve months)

- **Singh A and Guttag JV (2011) A comparison of non-symmetric entropy-based classification trees and support vector machine for cardiovascular risk stratification. Engineering in Medicine and Biology Society,EMBC, 2011 Annual International Conference of the IEEE. pp. 79-82.**

CONCLUSIONS

PROGETTO SHARE

- **We developed a risk assessment system for future vascular events** within 12 months from the recording
  - completely automatic;
  - using HRV analysis;
  - based on data-mining methods including intelligible model (i.e. classification tree / if-then rules).

- **Further developments:**
  - Larger dataset
  - Longer follow-up period
  - New / other HRV indexes (i.e. point process time-frequency analysis)
  - Other non-invasive measurement
  - 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