Annals of Biomedical Engineering, Vol. 38, No. 10, October 2010 *(© 2010)* pp. 3173–3184
DOI: 10.1007/s10439-010-0081-8

Detection Performance and Risk Stratification Using a Model-Based Shape Index Characterizing Heart Rate Turbulence

JUAN PABLO MARTÍNEZ,[1,2] IWONA CYGANKIEWICZ,[3,4] DANNY SMITH,[5,6] ANTONIO BAYÉS DE LUNA,[4] PABLO LAGUNA,[1,2] and LEIF SÖRNMO[5,6]

[1]Communications Technology Group (GTC), Aragón Institute of Engineering Research (I3A), University of Zaragoza, Zaragoza, Aragon, Spain; [2]Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Zaragoza, Aragon, Spain; [3]Department of Electrocardiology, Medical University of Lodz, Lodz, Poland; [4]Institut Català de Ciències Cardiovasculars, Barcelona, Catalonia, Spain; [5]Signal Processing Group, Department of Electrical and Information Technology, Lund University, Lund, Sweden; and [6]Center of Integrative Electrocardiology at Lund University (CIEL), Lund, Sweden

(Received 8 March 2010; accepted 18 May 2010; published online 2 June 2010)

Associate Editor Ioannis A. Kakadiaris oversaw the review of this article.

Abstract—A detection–theoretic approach to quantify heart rate turbulence (HRT) following a ventricular premature beat is proposed and validated using an extended integral pulse frequency modulation (IPFM) model which accounts for HRT. The modulating signal of the extended IPFM model is projected into a three-dimensional subspace spanned by the Karhunen–Loève basis functions, characterizing HRT shape. The presence or absence of HRT is decided by means of a likelihood ratio test, the Neyman–Pearson detector, resulting in a quadratic detection statistic. Using a labeled dataset built from different interbeat interval series, detection performance is assessed and found to outperform the two widely used indices: turbulence onset (TO) and turbulence slope (TS). The ability of the proposed method to predict the risk of cardiac death is evaluated in a population of patients ($n = 90$) with ischemic cardiomyopathy and mild-to-moderate congestive heart failure. While both TS and the novel HRT index differ significantly in survivors and cardiac death patients, mortality analysis shows that the latter index exhibits much stronger association with risk of cardiac death (hazard ratio = 2.8, CI = 1.32–5.97, $p = 0.008$). It is also shown that the model-based shape indices, but not TO and TS, remain predictive of cardiac death in our population when computed from 4-h instead of 24-h ambulatory ECGs.

Keywords—Heart rate turbulence, Neyman–Pearson detection, Likelihood ratio test, Karhunen–Loève transform, Detection theory, Mortality analysis, Risk stratification, Ischemic cardiomyopathy, Congestive heart failure.

Address correspondence to Juan Pablo Martínez, Communications Technology Group (GTC), Aragón Institute of Engineering Research (I3A), University of Zaragoza, Zaragoza, Aragon, Spain. Electronic mail: jpmart@unizar.es

INTRODUCTION

Heart rate turbulence (HRT) is the name given to the typical response in heart rate due to a ventricular premature beat (VPB). This response consists of an early heart rate acceleration followed by a deceleration.[3,14] The mechanism behind this phenomenon is not completely understood, but it is considered as a baroreflex response triggered by the blood pressure drop induced by the VPB. Pharmacological studies have shown that the main role in HRT is played by the parasympathetic branch of the autonomic nervous system,[10,11,20] though a significant correlation between sympathetic burst activity and HRT parameters was recently found.[16]

HRT is commonly assessed by two parameters: turbulence onset (TO) and turbulence slope (TS), both computed from the sequence of beat-to-beat intervals (RR intervals) following VPBs. TO measures the change in RR interval length immediately after the VPB, whereas TS quantifies the speed of RR interval increase following the initial shortening.[3,14] Absence of HRT, identified by a non-negative TO or a low TS, has been shown to be a powerful risk predictor in different populations, mainly in postinfarction[1,5,6,12] and congestive heart failure (CHF) patients.[4,13]

Besides TO and TS, several other HRT indices have been presented such as turbulence timing, turbulence jump, and correlation coefficient of TS.[19] Other TS-related indices include those being adjusted with the number of averaged beats[7] and turbulence dynamics, which quantifies the correlation between TS and heart rate.[2] All these indices are empirical as they do not result from data modeling.

We have recently introduced a model-based, statistical approach to HRT characterization.[17,18] An extended integral pulse frequency modulation (IPFM) model was proposed in which the modulating signal was assumed to be the sum of the background heart rate variability (HRV) and the specific response to the VPB, i.e., the possible HRT. This response was linearly modeled with the Karhunen–Loève (KL) expansion. With this model as the starting point, the statistical detection problem of deciding whether HRT is present or not was addressed, leading to detectors which are based on the generalized likelihood ratio test (GLRT); the difference between the detectors being whether *a priori* information on mean HRT shape is employed or not. Both detectors were found to outperform TS and TO on simulated signals.

In this study, we will use the same description of HRT shape, but will explore and evaluate a different statistical approach for HRT detection, accounting for second-order statistics of the observations. The ability of the proposed HRT shape index to discriminate observations with HRT from observations without HRT is assessed on a labeled dataset made from real series of RR intervals. The new index is also validated in terms of risk stratification on patients with ischemic cardiomyopathy and mild-to-moderate CHF. The performance is compared to that obtained with TO and TS.

DATASETS

Dataset for Assessment of Detection Performance

To assess HRT detection performance, an annotated dataset is needed which includes a significant number of signals with and without HRT. Since it is not possible to annotate the RR interval series following a single VPB with respect to the presence of HRT, we have created an evaluation dataset with RR interval series obtained from Holter recordings which is described next.

The Long-Term ST Database

The Long-Term ST database[8] contains 68 two-lead and 18 three-lead 24-h sinus rhythm ambulatory records with significant ST events annotated by expert cardiologists. These records have a sampling rate of 250 Hz and 12-bit resolution. The database also provides individual QRS annotations, including automatic beat classification revised by an expert Holter technician.

Dataset Construction

Our approach relies on the assumption that Holter recordings with enough VPBs, and a clear HRT pattern as quantified by TS in the averaged tachogram, present an HRT response after every individual VPB. Thus, the RR intervals adjacent to all the VPBs in such records are included in the HRT class ($\mathcal{S}_1$). On the other hand, records with few VPBs ($\leq$15) or with an impaired HRT in the averaged tachogram were discarded due to the uncertain presence of HRT following a VPB. To avoid this uncertainty, the class of non-HRT series ($\mathcal{S}_0$) was created by including tachograms extracted from periods without any premature or ectopic beat, i.e., only containing spontaneous HRV.

The VPBs in all records were selected according to the annotations provided in the database. To avoid the effect of neighboring non-sinus beats, artifacts, and false classifications, VPBs were discarded if any of the following criteria was fulfilled[15,17]: (1) the coupling interval is $\geq$80% of the average RR in the ten sinus beats prior to the VPB, (2) the compensatory pause is $<$120% of the average RR, (3) any beat in the neighborhood is annotated as non-normal, (4) any RR interval in the neighborhood is $<$300 or $>$2000 ms, or has a difference $>$20% with respect to the reference interval defined as the average of the five intervals preceding the VPB, and (5) the difference between any two adjacent RR intervals in the neighborhood is $>$200 ms. For these criteria, the neighborhood is defined as 10 RR intervals before the coupling interval and 20 RR intervals after the compensatory pause.

After applying these criteria, 64 records remained with suitable VPBs, with only 38 records having more than 15 suitable VPBs. Out of those 38 records, only 26 showed TS $>$ 2.5 ms/beat in the averaged tachogram. All the tachograms for the 5764 individual VPBs in those 26 records were included in class $\mathcal{S}_1$. To build class $\mathcal{S}_0$, we extracted consecutive 32-beat tachograms with normal beats from the 10 patients who lacked VPBs and treated these as tachograms without HRT since they contained "background" HRV. A total of 26,577 tachograms without HRT were obtained in this way. One hundred tachograms, randomly selected from each class, are shown in Fig. 1 together with the mean tachogram.

Study Population for Risk Stratification of Heart Failure Patients

The clinical population consists of 96 patients with ischemic cardiomyopathy and mild-to-moderate CHF (II–III NYHA class) enrolled into the MUSIC (Muerte Súbita en Insuficiencia Cardiaca = Sudden Death in Heart Failure) study by one of the participating centers. Only patients with sinus rhythm were included in the study. The study protocol was approved by institutional investigation committees, and all patients signed informed consent.

A 24-h Holter recording (3 orthogonal leads, 200 Hz sampling rate) was performed at enrollment by

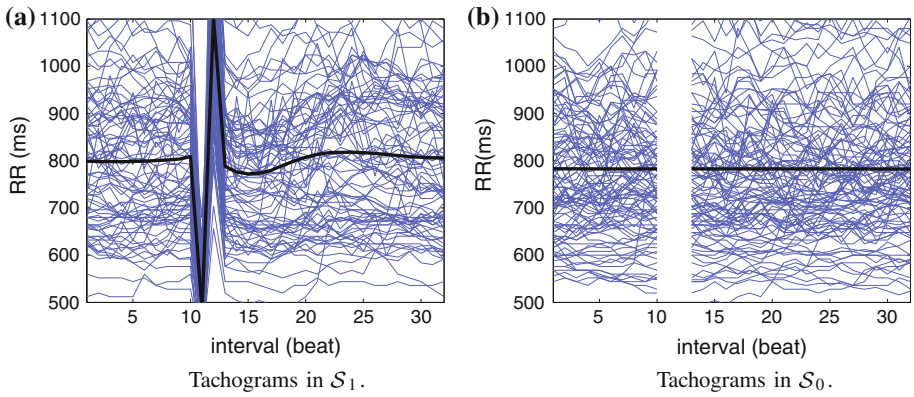

FIGURE 1. Tachograms for (a) HRT class $\mathcal{S}_1$ and (b) non-HRT class $\mathcal{S}_0$. The mean tachogram is displayed with a thick line, and the mean HRT shape can be observed in (a). Intervals 11 and 12 correspond to the coupling interval and compensatory pause in (a), and have been blanked in (b) since there are no VPBs in $\mathcal{S}_0$.

means of a SpiderView recorder (ELA Medical, Sorin Group) and analyzed by accompanying software; the results were manually verified to assure proper beat labeling. Data on RR intervals and beat class annotations were exported for analysis of HRT. VPBs unsuitable for HRT analysis were discarded, according to the criteria described in the section "Dataset for assessment of detection performance." Six patients did not have any suitable VPB and were therefore excluded from subsequent analysis.

The final study population consisted of 90 patients with mean age 64 ± 9 years, 80 males (mean age 64 ± 9 years) and 10 females (mean age 64 ± 8 years). The mean LVEF was 36 ± 10%, 72 patients belonged to NYHA class II (80%), and 45 patients were suffering from diabetes (50%). Patients were followed for a median of 44 months with endpoints defined as total mortality, cardiac mortality, and sudden death.

METHODS

The Extended IPFM Model

The extended model includes an additional feedback path which accounts for HRT as an additive signal $s(t)$, triggered by a VPB, so that the modulating signal $x(t)$ of the extended model is the sum of the background HRV $m(t)$ and the feedback signal $s(t)$, i.e., $x(t) = s(t) + m(t)$. The feedback signal was modeled as a linear combination of several basis functions, obtained from the truncated basis of the KLT learned on a training set. The modulating signal $x(t)$ following a VPB was considered as an observation of the response to the VPB. The background modulating signal $m(t)$ was treated as observation noise.

Thus, the first processing step is to estimate the modulating signal $x(t)$ from each individual post-VPB tachogram.[18] Then, the first 15 s of the estimated $x(t)$ after each VPB was evenly resampled at 2 Hz to produce an $N \times 1$ vector $\mathbf{x}$ ($N = 31$). This vector is the sum of the turbulence and background HRV components: $\mathbf{x} = \mathbf{s} + \mathbf{m}$. The modulating signals, corresponding to the tachograms shown in Fig. 1, are displayed in Fig. 2 together with the corresponding mean signals of $\mathcal{S}_1$ and $\mathcal{S}_0$. The shape of the mean HRT is evident from Fig. 2a.

The HRT term $\mathbf{s}$ is modeled as a linear combination of r basis functions

$$\mathbf{s} = \mathbf{B}\boldsymbol{\theta}_\mathbf{s}, \tag{1}$$

where $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_r]$ contains the r basis functions in its columns and $\boldsymbol{\theta}_\mathbf{s}$ is the $r \times 1$ coefficient vector describing HRT.

The basis functions $\mathbf{b}_i$ were obtained as the KLT basis for the set of vectors $\mathbf{x}_l$ belonging to the HRT class, $\mathcal{S}_1$. The percentage of each class' energy explained by the subspace, defined by the first r KLT basis functions, is shown in Fig. 3a. Figure 3b shows the distribution of the two classes in the most important KL coefficients. The three first basis functions, displayed in Fig. 4, account for 95% of the energy in the HRT class $\mathcal{S}_1$. Therefore, we will use $r = 3$ as the linear model dimension. Note that this linear model also accounts for 90% of the energy in the non-HRT class $\mathcal{S}_0$.

HRT Detection Using Mean Shape Information

Based on the extended IPFM model, the detection problem was previously formulated as one where HRT is either present (hypothesis $\mathcal{H}_1$) or absent (hypothesis $\mathcal{H}_0$) in the observed vector $\mathbf{x}$[17]:

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{x} = \mathbf{m} \\ \mathcal{H}_1 &: \mathbf{x} = \mathbf{B}\boldsymbol{\theta}_\mathbf{s} + \mathbf{m}. \end{aligned} \tag{2}$$

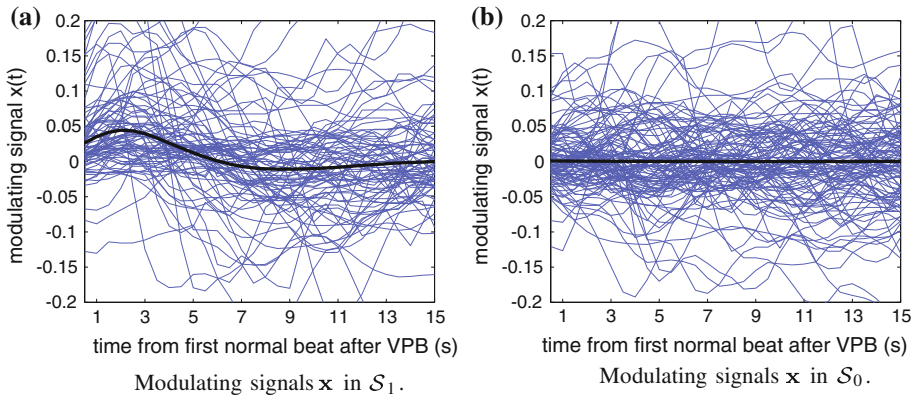

FIGURE 2. IPFM modulating signals (a) following the VPB for the HRT dataset ($\mathcal{S}_1$) and (b) the non-HRT ($\mathcal{S}_0$) dataset. The mean modulating signal is displayed with a thick line.

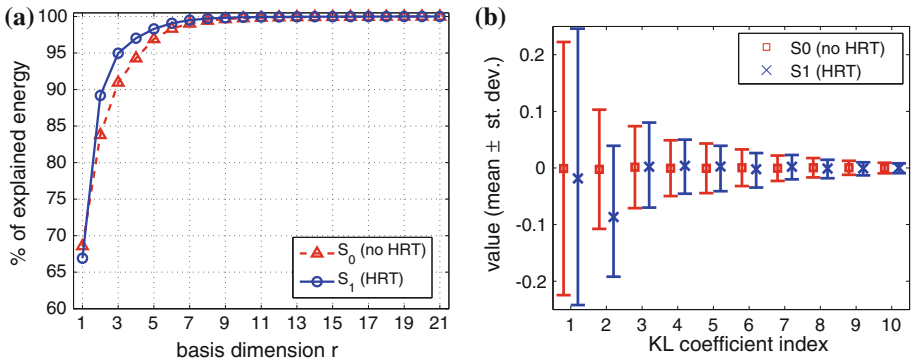

FIGURE 3. (a) Percentage of energy in $\mathcal{S}_0$ and $\mathcal{S}_1$ explained by the turbulence KLT basis as a function of basis dimension *r*. (b) Mean ± 1 standard deviation of the most important KL coefficients for vectors in $\mathcal{S}_0$ and $\mathcal{S}_1$.

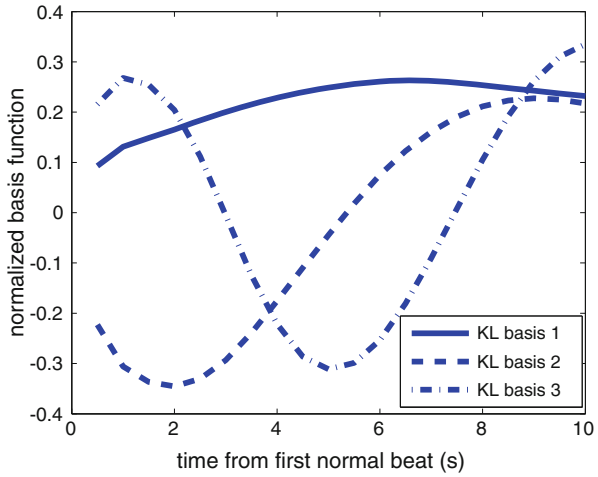

FIGURE 4. First three functions of the KLT basis, i.e., b_1, b_2, and b_3.

The GLRT detector for the hypothesis test in Eq. (2) was derived under the assumption that $\boldsymbol{\theta}_s$ is an unknown deterministic vector and $\mathbf{m}$ is a random vector distributed as white multivariate Gaussian $\mathbf{m} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$, with σ^2 unknown.[18] Since the resulting detector turned out to be invariant to the sign of the response to the VPB, it was subsequently improved by also taking into account *a priori* information on the mean HRT shape.[17] For the improved GLRT detector, the noise model remains the same, but $\boldsymbol{\theta}_s$ is assumed to be a known constant vector $\boldsymbol{\theta}_s = \boldsymbol{\mu}$, with $\boldsymbol{\mu}$ being a mean HRT shape vector estimated from a training dataset. The test statistic is then[17]:

$$T_{\mu}(\mathbf{x}) = \frac{\mathbf{x}^{\mathrm{T}}\mathbf{x}}{(\mathbf{x} - \mathbf{B}\boldsymbol{\mu})^{\mathrm{T}}(\mathbf{x} - \mathbf{B}\boldsymbol{\mu})}. \tag{3}$$

HRT Detection Using Covariance Information

In this paper, we reformulate the model in Eq. (2) to become one which instead treats the coefficients $\boldsymbol{\theta}_x = \mathbf{B}^{\mathrm{T}}\mathbf{x}$ as observations, being related to the reduced subspace defined by the first three KLT basis functions. This model offers detailed characterization of the observations as not only the mean $\boldsymbol{\mu}$ is included, as

in Smith *et al.*,[17] but also the covariance of the three coefficients in $\boldsymbol{\theta}_{\mathbf{x}}$. Hence, the detection problem can be formulated in the coefficient space as

$$\begin{aligned}\mathcal{H}_0 &: \boldsymbol{\theta}_{\mathbf{x}} = \mathbf{B}^{\mathrm{T}}\mathbf{m} = \boldsymbol{\theta}_{\mathbf{m}} \\ \mathcal{H}_1 &: \boldsymbol{\theta}_{\mathbf{x}} = \boldsymbol{\theta}_{\mathbf{s}} + \mathbf{B}^{\mathrm{T}}\mathbf{m} = \boldsymbol{\theta}_{\mathbf{s}} + \boldsymbol{\theta}_{\mathbf{m}}.\end{aligned} \tag{4}$$

Modeling both the HRT and background HRV components as correlated multivariate Gaussian random vectors with means $\boldsymbol{\mu}$ and $\mathbf{0}$, respectively, we have a detection problem of the form

$$\begin{aligned}\mathcal{H}_0 &: \boldsymbol{\theta}_{\mathbf{x}} \sim \mathcal{N}(\mathbf{0}, \Sigma_0) \\ \mathcal{H}_1 &: \boldsymbol{\theta}_{\mathbf{x}} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma_1),\end{aligned} \tag{5}$$

where $\boldsymbol{\mu}$, and the covariance matrices Σ_0 and Σ_1 can be estimated from a labeled training set.

Applying the Neyman–Pearson (NP) criterion to the hypothesis problem, defined by Eqs. (4) and (5), it can be shown that the detection statistic maximizing the probability of detection P_{D} for a given probability of false alarm P_{FA} is given by the likelihood ratio[9]:

$$\mathcal{L}(\mathbf{x}) = \frac{p(\mathbf{x};\mathcal{H}_1)}{p(\mathbf{x};\mathcal{H}_0)} = \frac{|\Sigma_0|^{\frac{1}{2}}\exp\left(-\frac{1}{2}(\boldsymbol{\theta}_{\mathbf{x}}-\boldsymbol{\mu})^{\mathrm{T}}\Sigma_1^{-1}(\boldsymbol{\theta}_{\mathbf{x}}-\boldsymbol{\mu})\right)}{|\Sigma_1|^{\frac{1}{2}}\exp\left(-\frac{1}{2}\boldsymbol{\theta}_{\mathbf{x}}^{\mathrm{T}}\Sigma_0^{-1}\boldsymbol{\theta}_{\mathbf{x}}\right)}. \tag{6}$$

Taking the logarithm of Eq. (6) and discarding constant terms, we obtain the equivalent detection statistic,

$$T_{\Sigma}(\mathbf{x}) = \boldsymbol{\theta}_{\mathbf{x}}^{\mathrm{T}}\Sigma_0^{-1}\boldsymbol{\theta}_{\mathbf{x}} - (\boldsymbol{\theta}_{\mathbf{x}}-\boldsymbol{\mu})^{\mathrm{T}}\Sigma_1^{-1}(\boldsymbol{\theta}_{\mathbf{x}}-\boldsymbol{\mu}). \tag{7}$$

It must be noted that Σ_0 and Σ_1 are not explicitly related to each other in Eqs. (5) and (7), and therefore the model, and, consequently, the detection statistic Eq. (7) is general and does not necessarily assume additivity of HRT and HRV.

Assessment of Detection Performance

We evaluated the ability of the proposed NP detector $T_{\Sigma}(\mathbf{x})$ (Eq. 7), the GLRT detector $T_{\boldsymbol{\mu}}(\mathbf{x})$ (Eq. 3), and the standard indices TS and TO to discriminate between the presence or absence of HRT.

Each of the two datasets $\mathcal{S}_0$ and $\mathcal{S}_1$ was divided into two halves: the first half ($\mathcal{S}_{0\mathrm{ts}}$ and $\mathcal{S}_{1\mathrm{ts}}$) was used as test set, whereas the second half ($\mathcal{S}_{0\mathrm{tr}}$ and $\mathcal{S}_{1\mathrm{tr}}$) was used as training set. The model parameters $\boldsymbol{\mu}$, needed in both $T_{\Sigma}(\mathbf{x})$ and $T_{\boldsymbol{\mu}}(\mathbf{x})$, and Σ_0, Σ_1, needed for $T_{\Sigma}(\mathbf{x})$, were estimated as the sample mean vector and sample covariance matrices of vectors $\boldsymbol{\theta}_{\mathbf{x}} = \mathbf{B}^{\mathrm{T}}\mathbf{x}$ in datasets $\mathcal{S}_{0\mathrm{tr}}$ and $\mathcal{S}_{1\mathrm{tr}}$. Their respective values are given in the Appendix.

Usually the tachograms following different VPBs are averaged in order to improve the poor SNR of a single response.[14] To assess the detection performance of the HRT indices computed from averaged responses, we built average-response datasets by averaging the responses of M consecutive VPBs belonging to the same patient. In this way, we built datasets for $M = \{10, 50, 100\}$ averaged responses in addition to the original single-VPB dataset ($M = 1$).

Detection performance is given in terms of ROC curves for different values of M. The performance was quantified by computing the area under the ROC curve (AUC) and the probability of detection ($P_{\mathrm{D}}^{0.1}$) for a given probability of false alarm $P_{\mathrm{FA}} = 0.1$.

Statistical Methods for Evaluation of Risk Stratification Ability

The different HRT indices were computed from the averaged responses using all VPBs suitable according to the criteria stated in the section"Dataset for assessment of detection performance." The indices were also computed using only the suitable VPBs of the first 4 h of the Holter recording. The cut-off points for dichotomization were the commonly used values of 0% and 2.5 ms/beat for TO and TS, respectively, and the lower tertile of the population for the model-based indices.

The KLT basis $\mathbf{B}$ and the statistical parameters $\boldsymbol{\mu}, \Sigma_0$, and Σ_1, used to compute the statistics $T_{\Sigma}(\mathbf{x})$ and $T_{\boldsymbol{\mu}}(\mathbf{x})$ for risk stratification, were identical to those used for assessing detection performance, and were thus estimated independently from the MUSIC dataset.

The two-tailed non-parametric Mann–Whitney U-test was used to test differences in parameters between patients in the different outcome groups. Kaplan–Meier mortality curves were calculated to test the association of the HRT indices with the endpoints of the study, and compared using the log-rank test. Cox's univariate regression models were used for the mortality analysis, expressing results as hazard ratios. Statistical significance was defined as $p < 0.05$.

RESULTS

HRT Detection Performance

Figure 5a shows the scatter plot of a selection of vectors $\mathbf{x} \in \mathcal{S}_0, \mathcal{S}_1$ represented in the subspace defined by the first three KL basis functions. Each point represents the shape of the turbulence after a single VPB. A substantial overlap can be observed between the two datasets. It should be recalled that $T_{\Sigma}(\mathbf{x})$ in Eq. (7) provides a quadratic separation surface between the two groups. Figure 6a shows the ROC curves in the test set for the four different indices.

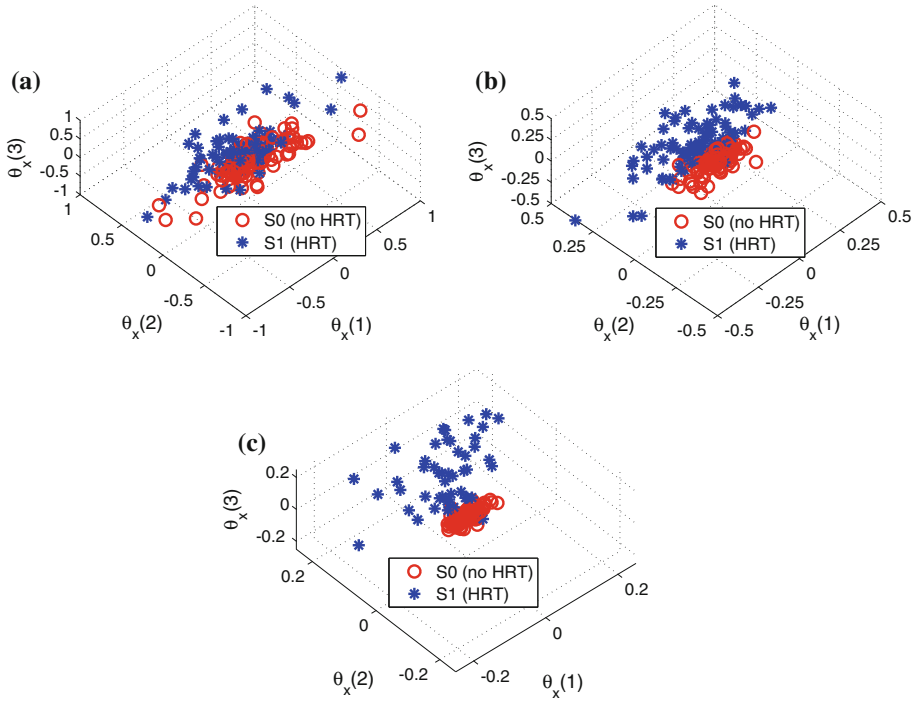

FIGURE 5. Scatter plot of θ_x for (a) single post-VPB responses, and averages of (b) 10 and (c) 100 consecutive post-VPB responses.

Figures 5b and 5c show that the groups become more separable when the responses to 10 or 100 consecutive VPBs are averaged. The ROC curves for different numbers of VPBs used for averaging are shown in Figs. 6b–6d. Note that the values of $\boldsymbol{\mu}, \Sigma_0$, and Σ_1 have not been re-estimated for averaged VPBs, but were first computed for individual VPBs in the training set and then used in the test set regardless of averaging.

The detection performance is summarized in Table 1 in terms of AUC and $P_D^{0.1}$.

Risk Stratification

During the follow-up, 29 patients (32%) died, of which 27 were from cardiovascular causes (13 of sudden cardiac death and 14 of heart failure progression). Table 2 summarizes the distribution of age, LVEF, and HRT indices in the studied population. Using the cut-off points of 0% and 2.5 ms/beat for TO and TS, 21 patients (23%) had abnormal TO and 42 patients (47%) had abnormal TS. With the above-mentioned lower tertile as cut-off point, 30 patients had abnormal $T_\Sigma(\mathbf{x})$ and $T_{\boldsymbol{\mu}}(\mathbf{x})$.

Table 3 shows the differences in HRT indices and other clinical variables between patients with cardiac death and survivors. Patients who died had lower LVEF (median 33% vs. 39%, $p = 0.004$) and more frequently exhibited higher NYHA functional class (NYHA III 44% vs. 10%, $p < 0.001$).

As for indices characterizing HRT, cardiac death victims had significantly lower values of TS and $T_\Sigma(\mathbf{x})$, whereas differences in the values of TO and $T_{\boldsymbol{\mu}}(\mathbf{x})$ did not attain significance. Once categorized according to the cut-off points, only the number of patients with abnormal $T_\Sigma(\mathbf{x})$ was significantly different between the two groups (44 vs. 21%, $p = 0.030$). The rest of dichotomized indices did not achieve significance in the study group. When combining TO and TS, the share of patients with abnormal combined TO/TS, defined as presenting abnormal values in at least one of the parameters TO and TS (HRT category 1 or 2 according to the HRT literature), differed significantly between the two groups (74 vs. 48%, $p = 0.021$).

Kaplan–Meier mortality curves, indicating probability of cardiac death during follow-up using the respective dichotomized values, are shown in Figs. 7 and 8. Patients with abnormal $T_\Sigma(\mathbf{x})$ were characterized by a more than 2.5-fold higher cardiac mortality estimated at third year of follow-up (45 vs. 18%, $p = 0.005$). The mortality was nearly twofold higher for patients stratified by TS and TO; however, statistical significance was only observed for TS ($p = 0.043$). It is noteworthy that the Kaplan–Meier curves diverged already at the first year of follow-up for $T_\Sigma(\mathbf{x})$, (32 vs. 5%), whereas abnormal TS was related with gradual increase in mortality risk during long-term observation. Sudden cardiac death risk was also higher in

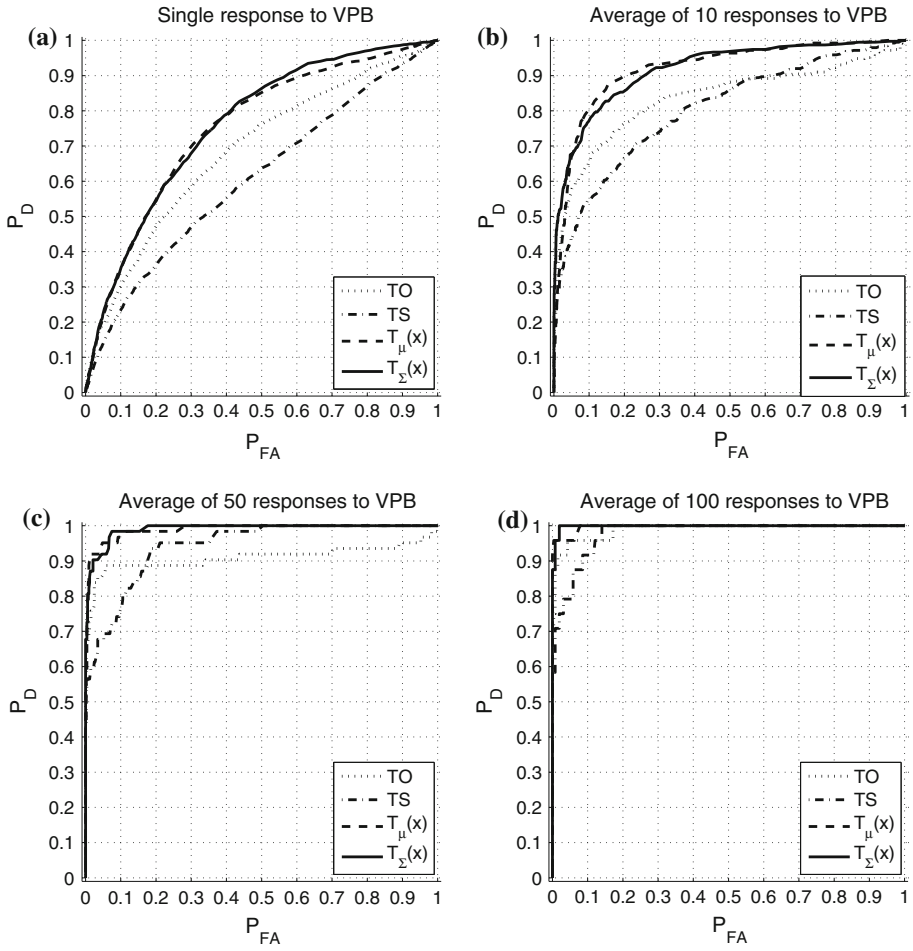

FIGURE 6. ROC curves of detection performance for the test set $\mathcal{S}_1$ using (a) single post-VPB responses, and averages of (b) 10, (c) 50, and (d) 100 responses.

TABLE 1. Detection performance of TS, TO, $T_\Sigma(\mathbf{x})$, and $T_\mu(\mathbf{x})$ according to the number of averaged responses *N*.

Index	$N = 1$	$N = 10$	$N = 50$	$N = 100$
TS				
AUC	0.61	0.81	0.94	0.978
$P_D^{0.1}$	0.24	0.55	0.76	0.92
TO				
AUC	0.69	0.84	0.91	0.989
$P_D^{0.1}$	0.30	0.66	0.89	0.96
$T_\Sigma(\mathbf{x})$				
AUC	0.76	0.92	0.99	0.999
$P_D^{0.1}$	0.35	0.77	0.98	1.0
$T_\mu(\mathbf{x})$				
AUC	0.75	0.92	0.99	0.997
$P_D^{0.1}$	0.35	0.80	0.97	1.0

TABLE 2. Clinical and HRT characteristics in the study population.

	Mean ± SD	Median	Lower tertile	Higher tertile
Age	64 ± 9	64.5	59	68
LVEF (%)	36 ± 10	35.5	33.3	39.0
TS (ms/beat)	4.26 ± 4.3	2.82	1.93	5.49
TO (%)	−1.02 ± 1.56	−0.81	−1.65	−0.30
$T_\mu(\mathbf{x})$	0.72 ± 1.03	0.43	0.22	0.70
$T_\Sigma(\mathbf{x})$	2.87 ± 8.71	0.37	−0.75	2.24

patients with abnormal $T_\mu(\mathbf{x})$ (20 vs. 12%), but did not reach statistical significance.

The results of Cox's univariate regression analysis are summarized in Table 4. Abnormal $T_\Sigma(\mathbf{x})$ exhibited the strongest association with cardiac death of all the HRT indices (hazard ratio = 2.8, CI = 1.32–5.97, p = 0.008), the other index with significant but weaker association being TS ≤ 2.5 (hazard ratio = 2.2, CI = 1.01–4.80, p = 0.048). Interestingly, abnormal $T_\Sigma(\mathbf{x})$ also exhibited stronger hazard ratio than abnormal combined TO/TS, indicating impaired HRT reaction according to classical approach.

When computing the HRT indices using only the first 4 h of the Holter recording, only 79 patients

TABLE 3. Differences in clinical variables and HRT indices between risk groups.

Variable	Survivors ($N = 61$)	Cardiac death ($N = 27$)	p-Value
Age (years)	63(63) ± 8	65(68) ± 12	0.154
Age >65 years	24 (39%)	16 (59%)	0.084
Gender (male)	52 (85%)	26 (96%)	0.132
LVEF (%)	**38 (39) ± 10**	**32 (33) ± 10**	**0.004**
LVEF ≤ 35%	**24 (39%)**	**20 (74%)**	**0.003**
NYHA class III	**6 (10%)**	**12 (44%)**	**<0.001**
Diabetes	31 (51%)	19 (70%)	0.088
TO (%)	−1.23 (−1.01) ± 1.48	−0.62 (−0.43) ± 1.68	0.060
TO ≥ 0%	11 (18%)	9 (33%)	0.114
TS (ms/beat)	**4.61 (3.18) ± 4.35**	**3.44 (1.80) ± 4.30**	**0.025**
TS ≤ 2.5	25 (41%)	17 (63%)	0.057
Combined TO/TS[a]	**29 (48%)**	**20 (74%)**	**0.021**
$T_\Sigma(\mathbf{x})$	**2.77 (0.93) ± 5.80**	**3.26 (−0.78) ± 13.47**	**0.048**
$T_\Sigma(\mathbf{x}) \leq -0.75$	**13 (21%)**	**12 (44%)**	**0.030**
$T_\mu(\mathbf{x})$	0.87 (0.57) ± 1.20	0.42 (0.31) ± 0.39	0.058
$T_\mu(\mathbf{x}) \leq 0.22$	18 (30%)	11 (41%)	0.301

Data are given as mean (median) ± SD, or # (%). Boldface indicates significant differences.
[a]Abnormal combined TO/TS refers to patients with at least one abnormal parameter (TO ≥ 0% and/or TS ≤ 2.5 ms/beat).

remained with at least 1 suitable VPB. In this subpopulation only abnormal $T_\mu(\mathbf{x})$ and $T_\Sigma(\mathbf{x})$ indicated patients with significantly higher risk of cardiac death (3-year cardiac mortality 49 vs. 19%, $p = 0.004$, and 44 vs. 20%, $p = 0.013$ for $T_\mu(\mathbf{x})$ and $T_\Sigma(\mathbf{x})$, respectively). In these shorter recordings $T_\mu(\mathbf{x})$ had the strongest predictive value when assessed by Cox's analysis (hazard ratio = 2.96, CI = 1.37–6.39, $p = 0.006$), followed by $T_\Sigma(\mathbf{x})$ (hazard ratio = 2.56, CI = 1.18–5.54, $p = 0.017$). Abnormal TS and TO were non-significant.

DISCUSSION

Models and Methods

To characterize an HR signal when HRT is present or absent, we have created a labeled dataset of RR interval series from Holter ECG signals as described in the section "Datasets," thus circumventing the need for simulation. The training part of this dataset has also been used to estimate $\boldsymbol{\mu}$, Σ_0, and Σ_1 of the model-based indices. It is particularly interesting to note that these estimates produced excellent results on risk stratification when $T_\Sigma(\mathbf{x})$ was employed, despite the fact that the estimates were based on a population completely unrelated to the patients with ischemic cardiomyopathy and mild-to-moderate CHF.

Figure 3 shows that the background HRV, which can be understood as the noise term in the model of Eq. (2), is also quite well represented in the signal subspace. Due to this overlap, a subspace energy detector will perform poorly. Therefore, we have focused on the signal subspace (as a good representation for both groups of signals) and have designed an NP detector by modeling both datasets as multivariate Gaussian with different covariance matrices (Eq. 4). Note that the signal model in Eq. (4), put forward in this study, is more general than the one in Eq. (2) as the interaction between background HRV and HRT does not have to be additive.

While TO and TS quantify certain parts of the turbulence response, i.e., the initial acceleration and the velocity of the posterior recovery, respectively, $T_\Sigma(\mathbf{x})$ and $T_\mu(\mathbf{x})$ characterize the whole shape of HRT and are therefore referred to as HRT shape indices.

As for the conceptual differences between $T_\Sigma(\mathbf{x})$ and $T_\mu(\mathbf{x})$, the NP statistic (7) disregards any part of the response orthogonal to the three-dimensional KLT subspace. This can be interpreted as linear, time-variant filtering of the data prior to detection. From Eqs. (3) and (7) it is obvious that both statistics can be interpreted in terms of the distances from the observed data to the origin and to the average turbulence shape $\boldsymbol{\mu}$, the main difference being that $T_\Sigma(\mathbf{x})$ normalizes the distances with the covariances of the two types of responses (with and without turbulence), while $T_\mu(\mathbf{x})$ tacitly assumes that all three dimensions are uncorrelated and have equal variance.

HRT Detection Performance

The two shape indices achieved a given detection performance with less averaging than did TO and TS as shown in Fig. 6 and Table 1; this result was already noted for the GLRT.[17] This improved performance can be interpreted as a gain in signal-to-noise ratio. On the other hand, the detection performance in our

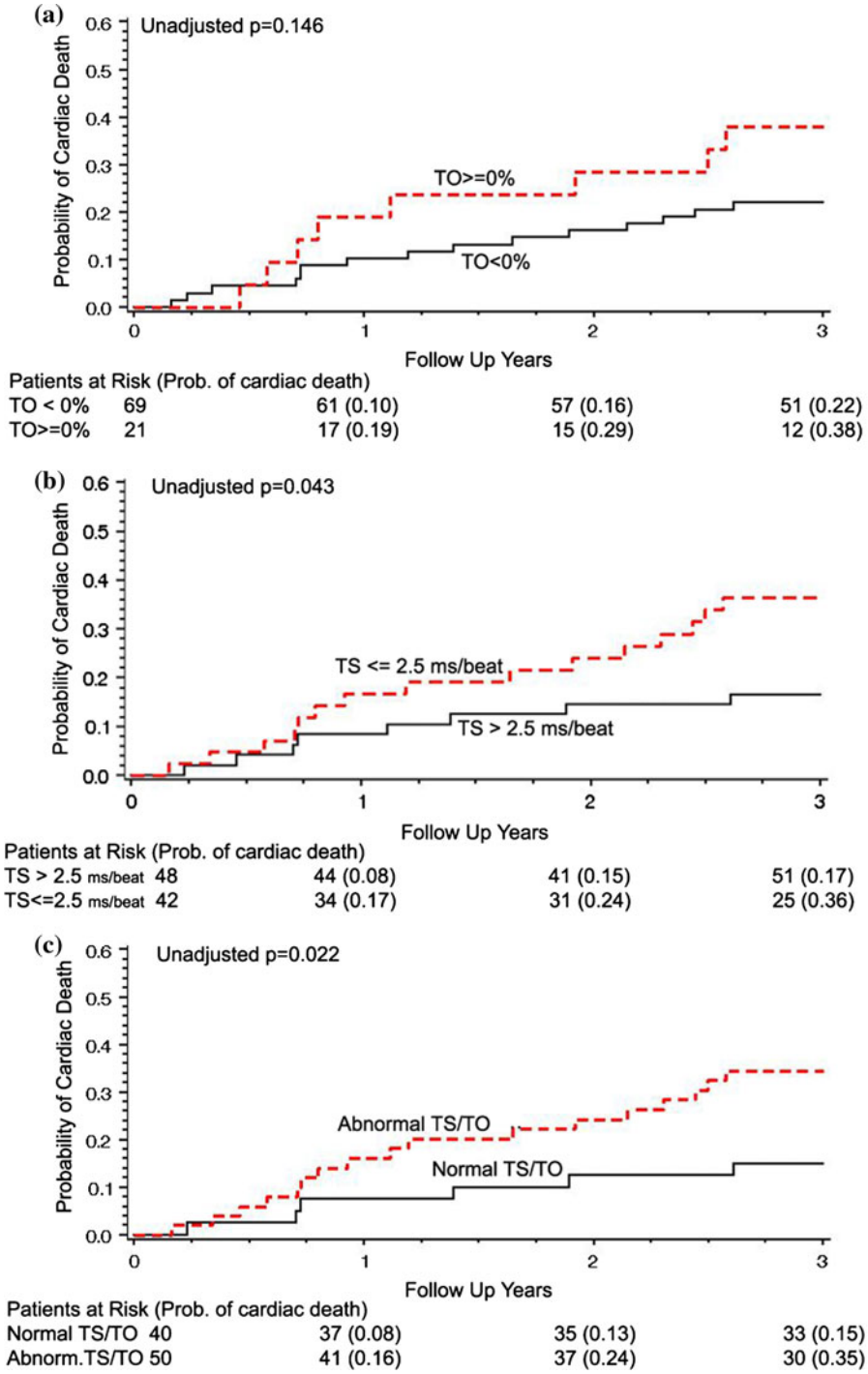

FIGURE 7. Kaplan–Meier mortality curves for (a) TO, (b) TS, (c) combined TO/TS. *p*-Values are given for long-rank tests. Numbers in parenthesis below the graph represent the probability of cardiac death in the first, second, and third year of the follow-up.

labeled dataset is virtually identical for both $T_{\Sigma}(\mathbf{x})$ and $T_{\mu}(\mathbf{x})$ in all cases.

When dealing with single VPBs, the detection performance is poor for all methods, although considerably better for $T_{\Sigma}(\mathbf{x})$ and $T_{\mu}(\mathbf{x})$. Thus, fixing $P_{\mathrm{FA}} = 0.1$ (10% of series without turbulence, i.e., with abnormal response, would be considered to have a normal response), P_{D} is 0.35 for $T_{\Sigma}(\mathbf{x})$ and 0.35 for $T_{\mu}(\mathbf{x})$, but 0.25 for TO and 0.15 for TS. When applied to averages of ten responses, detection performance improved, with $P_{\mathrm{D}}^{0.1}$ values of 0.76 and 0.80 for $T_{\Sigma}(\mathbf{x})$ and $T_{\mu}(\mathbf{x})$, respectively, and 0.66 and 0.40 for TO and

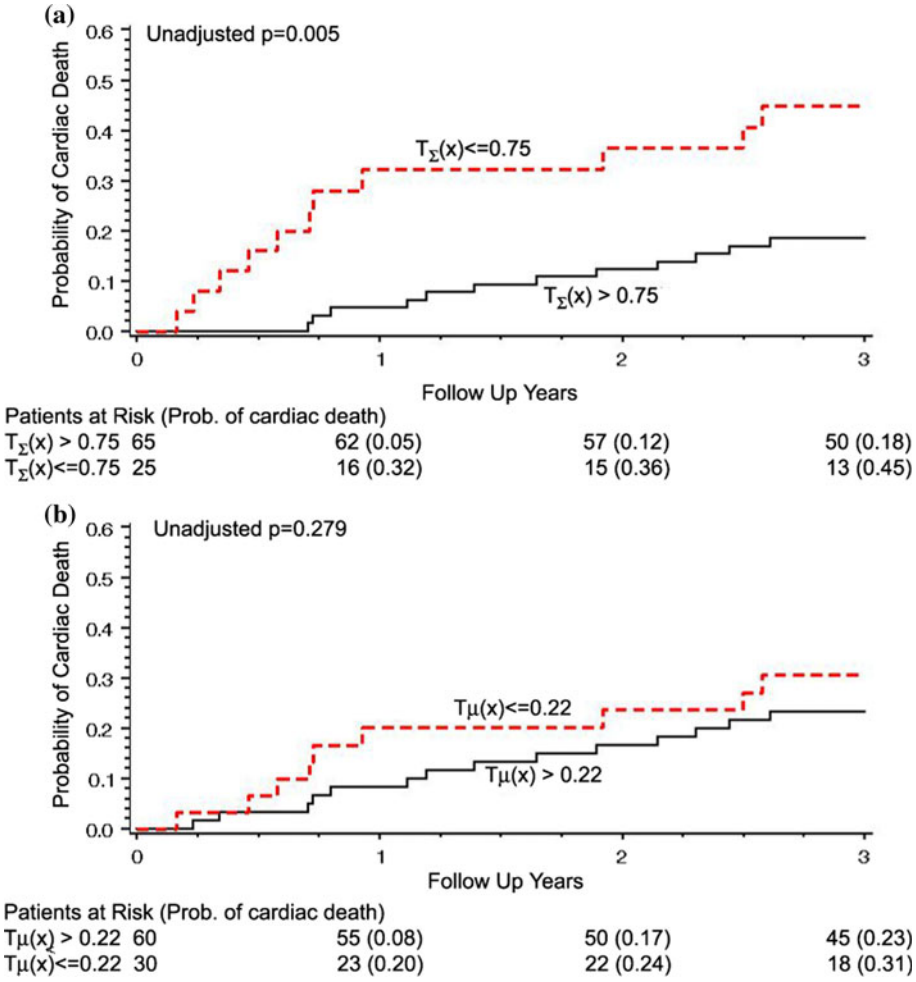

FIGURE 8. Kaplan–Meier mortality curves for (a) $T_\Sigma(x)$, (b) $T_\mu(x)$. *p*-Values are given for long-rank tests. Numbers in parenthesis below the graph represent the probability of cardiac death in the first, second, and third year of the follow-up.

TABLE 4. Univariate association of HRT risk variables with total mortality.

Variable	Hazards ratio	95% CI	*p*-Value
TO ≥ 0%	1.79	0.81–3.99	0.152
TS ≤ 2.5	**2.20**	**1.01–4.80**	**0.048**
Combined TO/TS[a]	**2.63**	**1.11–6.22**	**0.028**
$\boldsymbol{T_\Sigma(x) \leq -0.75]}$	**2.80**	**1.32–5.97**	**0.008**
$T_\mu(x) \leq 0.22$	1.52	0.71–3.28	0.289

Boldface indicates significant differences.
[a]Abnormal combined TO/TS refers to patients with at least one abnormal parameter (TO ≥ 0% and/or TS ≤ 2.5 ms/beat).

TS, respectively. Almost perfect discrimination was obtained when computing the parameters on averages of 100 responses. Similar conclusions are obtained from the comparison of the AUC results.

It is indeed surprising that TO detects HRT better than TS, while most clinical studies have found TS to be more powerful than TO for risk stratification.[4,6,12,14] It should be emphasized that the evaluation of the detection performance does not necessarily foretell the power of an index for risk stratification. Under the hypothesis that there is a connection between HRT and risk, a useful clinical index must have good detection performance; otherwise, it would be impossible to determine when the clinically useful phenomenon appears. However, it is not clear which specific HRT features are related to risk.

Risk Stratification

The concept of abnormal HRT as a risk marker of mortality was introduced in 1999.[14] Over the last decade several smaller and large-scale clinical trials have documented that impaired response of the sinus node to VPB predicts increased risk of all-cause mortality, as well as sudden death.[1,3–6,12–14] Prognostic value of HRT has been evaluated mainly in large

cohorts of postinfarction patients.[1,5,6,12,14] All these studies documented that HRT category 2 (both TO and TS abnormal) is the strongest risk predictor of mortality, followed by HRT category 1 (one of the parameters abnormal). Although TO and TS exhibit additive value in risk stratification, TS is usually found to provide more powerful information on risk.

In our study, risk stratification was assessed on a dataset of patients with ischemic cardiomyopathy and mild-to-moderate heart failure. Impaired HRT quantified by the newly proposed statistic $T_\Sigma(\mathbf{x})$ identified patients at high risk of cardiac death during a median of 44 months of follow-up. Patients with $T_\Sigma(\mathbf{x}) < 0.75$ were characterized by nearly a threefold higher risk of unfavorable outcome. This is the first time when model-based HRT indices were evaluated on a clinical setting for predicting cardiac death, and were proven to provide significant prognostic information. Our data showed that $T_\Sigma(\mathbf{x})$ is a stronger risk predictor than TS and TO. It is worth emphasizing that abnormal $T_\Sigma(\mathbf{x})$ was able to define high-risk patients very early in the follow-up, identifying patients with over sixfold higher mortality rate (32 vs. 5%) at the first year of follow-up. This result is of special significance since early initiation of more aggressive pharmacological and device treatment is important for high-risk patients.

While the exact reason is unclear to why $T_\Sigma(\mathbf{x})$ is better for risk stratification than $T_{\boldsymbol{\mu}}(\mathbf{x})$, it is obvious that the covariance information plays an important role to achieve these results. For example, it may be noted from $\boldsymbol{\mu}$, given in the Appendix, that the KL coefficients associated with the first and second basis functions have opposite signs in the average turbulence, whereas Σ_0 shows that the noise (HRV) KL coefficients have positive correlation: this fact is not taken into account in $T_{\boldsymbol{\mu}}(\mathbf{x})$. Also, while $T_{\boldsymbol{\mu}}(\mathbf{x})$ treats all the three KL coefficients identically regardless of their different variances, $T_\Sigma(\mathbf{x})$ is related to the Mahalanobis distance in the HRT subspace.

Another interesting finding is related to the fact that $T_\Sigma(\mathbf{x})$ and $T_{\boldsymbol{\mu}}(\mathbf{x})$, contrary to TO and TS, were found to be significant risk predictors when assessed from short-term recordings. Due to methodological issues related to the averaging process and the number of VPBs, current guidelines recommended HRT-based risk stratification to be limited to 24-h recordings with sinus rhythm and at least five VPBs eligible for analysis.[3] In 24-h Holters HRT can be calculated in 60–90% of patients if recordings with at least one VPB are considered as suitable for the analysis. Requiring instead at least five VPBs, a significant number of patients are no longer eligible for risk stratification. Furthermore, long-term Holter monitoring is frequently considered as costly, and therefore it is discarded as additional procedure. Use of the shape indices based on short recordings with few VPBs would overcome these limitations.

CONCLUSIONS

In this article, we have derived and evaluated a model-based HRT detector on the IPFM modulating signal using a second-order statistics approach within the signal subspace defined by the most relevant KLT functions. This detector, which involves an index that characterizes HRT shape, offers better performance than the commonly used TO and TS, and it attains similar performance with less averaging.

The proposed HRT index $T_\Sigma(\mathbf{x})$ was predictive for cardiac death in ischemic heart failure patients, with higher predictive value than TO and TS, either analyzed separately or together. The index proved to be useful for early risk stratification, as it identified patients with sixfold higher mortality at first year of follow-up, in contrast to a much more gradual increase in risk as identified by TO and TS. As documented by analysis of shorter recordings, $T_\Sigma(\mathbf{x})$ may be of special interest in patients with few VPBs. These promising results encourage further investigation in larger cohorts of patients in order to ascertain the predictive value of $T_\Sigma(\mathbf{x})$ and its relation to other clinical variables.

APPENDIX

The model parameters $\boldsymbol{\mu}, \Sigma_0$, and Σ_1, estimated from datasets $\mathcal{S}_{0\text{tr}}$ and $\mathcal{S}_{1\text{tr}}$, and used for computing $T_{\boldsymbol{\mu}}(\mathbf{x})$ and $T_\Sigma(\mathbf{x})$ have the following values:

$$\boldsymbol{\mu} = 10^{-3}[-34.699 \quad 107.246 \quad 0.112]^{\mathrm{T}} \tag{8}$$

$$\Sigma_0 = 10^{-3}\begin{bmatrix} 8.061 & 0.932 & -0.368 \\ 0.932 & 1.442 & 0.330 \\ -0.368 & 0.330 & 1.103 \end{bmatrix} \tag{9}$$

$$\Sigma_1 = 10^{-3}\begin{bmatrix} 32.262 & 3.726 & 0.004 \\ 3.726 & 6.603 & -0.012 \\ 0.004 & -0.012 & 5.454 \end{bmatrix} \tag{10}$$

ACKNOWLEDGMENTS

This study was supported by Projects TEC-2007-68076-C02-02 from CICYT, GTC T-30 from DGA (Spain), CIBER de Bioingeniería, Biomateriales y Nanomedicina, an initiative of the Instituto de Salud Carlos III, and the Swedish Research Council.

REFERENCES

[1]Barthel, P., R. Schneider, A. Bauer, K. Ulm, C. Schmitt, A. Schömig, and G. Schmidt. Risk stratification after acute myocardial infarction by heart rate turbulence. *Circulation* 108(10):1221–1226, 2003. doi:10.1161/01.CIR.0000088783.34082.89.

[2]Bauer, A., M. Malik, P. Barthel, R. Schneider, M. A. Watanabe, A. J. Camm, A. Schömig, and G. Schmidt. Turbulence dynamics: an independent predictor of late mortality after acute myocardial infarction. *Int. J. Cardiol.* 107(1):42–47, 2006. doi:10.1016/j.ijcard.2005.02.037.

[3]Bauer, A., M. Malik, G. Schmidt, P. Barthel, H. Bonnemeier, I. Cygankiewicz, P. Guzik, F. Lombardi, A. Mnller, A. Oto, R. Schneider, M. Watanabe, D. Wichterle, and W. Zareba. Heart rate turbulence: standards of measurement, physiological interpretation, and clinical use: International Society for Holter and Noninvasive Electrophysiology consensus. *J. Am. Coll. Cardiol.* 52(17):1353–1365, 2008. doi:10.1016/j.jacc.2008.07.041.

[4]Cygankiewicz, I., W. Zareba, R. Vazquez, M. Vallverdú, J. R. Gonzalez-Juanatey, M. Valdes, J. Almendral, J. Cinca, P. Caminal, A. Bayés de Luna, and M. S. Insuficiencia Cardiaca Investigators. Heart rate turbulence predicts all-cause mortality and sudden death in congestive heart failure patients. *Heart Rhythm* 5(8):1095–1102, 2008. doi:10.1016/j.hrthm.2008.04.017.

[5]Exner, D. V., K. M. Kavanagh, M. P. Slawnych, L. B. Mitchell, D. Ramadan, S. G. Aggarwal, C. Noullett, A. V. Schaik, R. T. Mitchell, M. A. Shibata, S. Gulamhussein, J. McMeekin, W. Tymchak, G. Schnell, A. M. Gillis, R. S. Sheldon, G. H. Fick, H. J. Duff, and Investigators, R.E.F.I.N.E. Noninvasive risk assessment early after a myocardial infarction the REFINE study. *J. Am. Coll. Cardiol.* 50(24):2275–2284, 2007.

[6]Ghuran, A., F. Reid, M. T. L. Rovere, G. Schmidt, J. T. Bigger, A. J. Camm, P. J. Schwartz, M. Malik, and Investigators, A.T.R.A.M.I. Heart rate turbulence-based predictors of fatal and nonfatal cardiac arrest (the autonomic tone and reflexes after myocardial infarction substudy). *Am. J. Cardiol.* 89(2):184–190, 2002.

[7]Hallstrom, A. P., P. K. Stein, R. Schneider, M. Hodges, G. Schmidt, and K. Ulm. Structural relationships between measures based on heart beat intervals: potential for improved risk assessment. *IEEE Trans. Biomed. Eng.* 51(8):1414–1420, 2004. doi:10.1109/TBME.2004.828049.

[8]Jager, F., A. Taddei, G. B. Moody, M. Emdin, G. Antolic, R. Dorn, A. Smrdel, C. Marchesi, and R. G. Mark. Long-term ST database: a reference for the development and evaluation of automated ischaemia detectors and for the study of the dynamics of myocardial ischaemia. *Med. Biol. Eng. Comput.* 41(2):172–183, 2003.

[9]Kay, S. M.: Fundamentals of Statistical Signal Processing. Detection theory. Upper Saddle River: Prentice Hall, 1993.

[10]Lin, L. Y., L. P. Lai, J. L. Lin, C. C. Du, W. Y. Shau, H. L. Chan, Y. Z. Tseng, and S. K. S. Huang. Tight mechanism correlation between heart rate turbulence and baroreflex sensitivity: sequential autonomic blockade analysis. *J. Cardiovasc. Electrophysiol.* 13(5):427–431, 2002.

[11]Marine, J. E., M. A. Watanabe, T. W. Smith, and K. M. Monahan. Effect of atropine on heart rate turbulence. *Am. J. Cardiol.* 89(6):767–769, 2002.

[12]Mäkikallio, T. H., P. Barthel, R. Schneider, A. Bauer, J. M. Tapanainen, M. P. Tulppo, G. Schmidt, and H. V. Huikuri. Prediction of sudden cardiac death after acute myocardial infarction: role of Holter monitoring in the modern treatment era. *Eur. Heart J.* 26(8):762–769, 2005. doi:10.1093/eurheartj/ehi188.

[13]Moore, R. K. G., D. G. Groves, P. E. Barlow, K. A. A. Fox, A. Shah, J. Nolan, and M. T. Kearney. Heart rate turbulence and death due to cardiac decompensation in patients with chronic heart failure. *Eur. J. Heart Fail.* 8(6):585–590, 2006. doi:10.1016/j.ejheart.2005.11.012.

[14]Schmidt, G., M. Malik, P. Barthel, R. Schneider, K. Ulm, L. Rolnitzky, A. J. Camm, J. T. Bigger, and A. Schömig. Heart-rate turbulence after ventricular premature beats as a predictor of mortality after acute myocardial infarction. *Lancet* 353:1390–1396, 1999.

[15]Schneider, R., P. Barthel, and M. Watanabe. Heart rate turbulence on Holter. In: Dynamic Electrocardiography, edited by M. Malik, and A. J. Camm. New York: Blackwell Futura, 2004, pp. 190–193.

[16]Segerson, N. M., S. L. Wasmund, M. Abedin, R. K. Pai, M. Daccarett, N. Akoum, T. S. Wall, R. C. Klein, R. A. Freedman, and M. H. Hamdan. Heart rate turbulence parameters correlate with post-premature ventricular contraction changes in muscle sympathetic activity. *Heart Rhythm* 4(3):284–289, 2007. doi:10.1016/j.hrthm.2006.10.020.

[17]Smith, D., K. Solem, P. Laguna, J. P. Martínez, and L. Sörnmo. Model-based detection of heart rate turbulence using mean shape information. *IEEE Trans. Biomed. Eng.* 57(2):334–342, 2010.

[18]Solem, K., P. Laguna, J. P. Martínez, and L. Sörnmo. Model-based detection of heart rate turbulence. *IEEE Trans. Biomed. Eng.* 55(12):2711–2722, 2008.

[19]Watanabe, M. A. Heart rate turbulence: a review. *Indian Pacing Electrophysiol. J.* 3(1):10–22, 2003.

[20]Wichterle, D., V. Melenovsky, J. Simek, J. Malik, and M. Malik. Hemodynamics and autonomic control of heart rate turbulence. *J. Cardiovasc. Electrophysiol.* 17(3):286–291, 2006.