

PAPER • OPEN ACCESS

Heart rate dynamics distinguish among atrial fibrillation, normal sinus rhythm and sinus rhythm with frequent ectopy

To cite this article: Marta Carrara *et al* 2015 *Physiol. Meas.* **36** 1873

View the [article online](#) for updates and enhancements.

Related content

- [Local dynamics of heart rate: detection and prognostic implications](#)
Travis J Moss, Douglas E Lake and J Randall Moorman
- [A comparison of entropy approaches for AF discrimination](#)
Chengyu Liu, Julien Oster, Erik Reinertsen *et al.*
- [Study on the P-wave feature time course as early predictors of paroxysmal atrial fibrillation](#)
Arturo Martínez, Raúl Alcaraz and José J Rieta

Recent citations

- [Nonlinear dynamic approaches to identify atrial fibrillation progression based on topological methods](#)
Bahareh Safarbali and Seyed Mohammad Reza Hashemi Golpayegani
- [Ashish Sharma and Shivnarayan Patidar](#)
- [Combining deep neural networks and engineered features for cardiac arrhythmia detection from ECG recordings](#)
Shenda Hong *et al*



IOP Astronomy ebooks

Part of your publishing universe and your first choice for astronomy, astrophysics, solar physics and planetary science ebooks.

iopscience.org/books/aas

Heart rate dynamics distinguish among atrial fibrillation, normal sinus rhythm and sinus rhythm with frequent ectopy

Marta Carrara¹, Luca Carozzi¹, Travis J Moss²,
Marco de Pasquale¹, Sergio Cerutti¹, Manuela Ferrario¹,
Douglas E Lake² and J Randall Moorman²

¹ Department of Electronics, Information and Bioengineering, Politecnico di Milano, Piazza Leonardo da Vinci 32, Milan, Italy

² Department of Medicine, University of Virginia, Charlottesville, VA, Box 800158, USA

E-mail: marta.carrara@polimi.it

Received 1 April 2015, revised 13 May 2015

Accepted for publication 8 June 2015

Published 6 August 2015



CrossMark

Abstract

Atrial fibrillation (AF) is usually detected by inspection of the electrocardiogram waveform, a task made difficult when the signal is distorted by noise. The RR interval time series is more frequently available and accurate, yet linear and nonlinear time series analyses that detect highly varying and irregular AF are vulnerable to the common finding of frequent ectopy. We hypothesized that different nonlinear measures might capture characteristic features of AF, normal sinus rhythm (NSR), and sinus rhythm (SR) with frequent ectopy in ways that linear measures might not. To test this, we studied 2722 patients with 24h ECG recordings in the University of Virginia Holter database. We found dynamical phenotypes for the three rhythm classifications. As expected, AF records had the highest variability and entropy, and NSR the lowest. SR with ectopy could be distinguished from AF, which had higher entropy, and from NSR, which had different fractal scaling, measured as higher detrended fluctuation analysis slope. With these dynamical phenotypes, we developed successful classification strategies, and the nonlinear measures improved on the use of mean and variability alone, even after adjusting for age. Final models using all variables had excellent performance, with positive predictive values for AF, NSR and SR with ectopy as high as 97, 98 and 90%, respectively. Since these classifiers can reliably



Content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

detect rhythm changes utilizing segments as short as 10 min, we envision their application in noisy settings and in personal monitoring devices where only RR interval time series may be available.

Keywords: cardiac rhythm classification, RR series, nonlinear analysis, sample entropy, atrial fibrillation

(Some figures may appear in colour only in the online journal)

1. Introduction

The dynamics of the heart beat have been studied for several decades and there is consensus that advanced mathematical analyses that describe its hallmark features, variability and complexity, yield important clinical information. While the majority of analyses have been of normal sinus rhythm (NSR), dynamical analyses of the RR interbeat intervals time series have also been advanced toward classifying cardiac rhythms. For example, it is reasonable to think that NSR might be distinguished from atrial fibrillation using nonlinear measures, and we previously developed the coefficient of sample entropy for this purpose (Lake and Moorman 2011). Similarly, there have been successful methods developed for atrial fibrillation (AF) detection based on Poincaré plots (Sarkar *et al* 2008) and cumulative distribution functions (Tateno and Glass 2001), as well as linear measures. These approaches might be confused, though, by sinus rhythm (SR) with frequent ectopy, for here the overall variability might rise to the level of atrial fibrillation. Since AF, SR with frequent ectopy, and NSR account for the majority of non-paced stable cardiac rhythms, methods in this area are of widespread importance.

AF commonly eludes diagnosis because of its paroxysmal nature, and, once found, decisions about its therapy are best informed by knowledge of its burden (Prystowsky 2000). For patients with AF, clinicians must make complex decisions about oral anticoagulation, rate control with either drugs or cardiac ablation with pacemaker therapy, and rhythm control with either drugs or ablation procedures. The implications for anticoagulation therapy alone are enormously important, as the rate of stroke goes up sharply with even short episodes of AF (Healey *et al* 2012).

Another clinical diagnosis of importance is SR with frequent ectopic beats. Atrial ectopy can be a harbinger of AF, and ventricular ectopy leads to poorer prognosis even in patients with no obvious heart disease (Lee *et al* 2012), possibly as a result of reversible cardiomyopathy (Yokokawa *et al* 2013).

Usually, AF is clinically diagnosed by the absence of P waves in the electrocardiogram (ECG) and a rapid irregular ventricular rate. However, even the high fidelity monitoring can have a poor performance when the ECG waveform is noisy or distorted. On the other hand, the heart beat patterns in ECG are less prone to artifacts and noise, such as baseline wandering, and they could be used in clinical contexts where ECG is noisy or not available and other pulsatile signals are used, such as arterial blood pressure or plethysmograph (Park *et al* 2009).

The analysis of the RR interbeat intervals time series is frequently applied for rhythm detection in implanted devices, where only ventricular electrographic recordings and RR intervals are available (Lake and Moorman 2011, DeMazumder *et al* 2013). However, non-invasive devices has a reduced confidence in detecting AF based on the heart rate or RR series alone and thus they are not commonly adopted for heart rhythm identification.

We propose a new approach in which entropy measurements are used in conjunction with other time series measurements in the time and nonlinear dynamical domains. The insight is that higher dimensional characterizations of the time series properties successfully separate the classes of rhythm. Here, we test three dynamical measures—entropy, local dynamics and fractal scaling. The coefficient of sample entropy (COSEn) is efficient at detecting AF from NSR, but yields intermediate values for SR with ectopy (Lake and Moorman 2011). The local dynamics score (LDs) detects the combination of reduced heart rate variability (HRV) with ectopy, and is a powerful predictor of four year mortality in ambulatory patients undergoing Holter monitoring (Moss *et al* 2014). Finally, the familiar detrended fluctuation analysis (DFA) (Peng *et al* 1995) informs on long-range correlations and has been studied in aging (Schmitt and Ivanov 2007), heart failure (Huikuri *et al* 2003, Cerutti *et al* 2007) and in predicting AF after coronary bypass surgery (Tarkiainen *et al* 2008).

One approach to rhythm classification using only RR intervals is to examine the dynamics of the time series. For example we expect AF to be irregular, SR with ectopy to have a regular baseline with intermittent departures, and NSR to be more consistently regular than either. Bravi and coworkers have reviewed the lexicon of RR interval time series variabilities and dynamics, categorizing more than 100 into domains of time, frequency, phase, nonlinear dynamical, statistical, geometric, energetic, informational and invariant domains (Bravi *et al* 2011). Many of these methods are at least modestly effective in distinguishing patients at high or low risk, or degrees of illness severity for patients in NSR (that is, with little or no ectopy). Some have been effective in distinguishing AF from other rhythms, principally NSR. Examples include coefficient of sample entropy (Lake and Moorman 2011, DeMazumder *et al* 2013), a histogrammic method (Tateno and Glass 2001), and a Poincaré plot-based method (Sarkar *et al* 2008). There is, however, no unified and effective solution for an algorithm that distinguishes NSR and AF from the condition of SR with atrial and ventricular ectopy.

2. Materials and methods

2.1. Study population

We studied RR interval time series from 24 h Holter recordings collected from 2722 consecutive patients at the University of Virginia (UVa) Heart Station from 12/2004 to 10/2010, and clinical characteristics of these patients are reported in a previous work (Moss *et al* 2014).

The age of the patients varied from 0 to 100 years, with an average value of 47 ± 25 years. AF, premature atrial contraction (PAC) and premature ventricular contraction (PVC) labels were obtained from an automatic classifier (Philips Holter Software). The RR series were subdivided into 377 285 10 min segments, accounting for 96% of the maximum possible recording time. Each 10 min segment was classified as AF if the burden of AF was greater than 5% (i.e. for more than 30 s), as SR with ectopy if the burden of PAC or PVC was more than 10%, as NSR otherwise. After this classification of individual segments, we found that the dataset was composed of 79% NSR, 8% AF and 13% SR with ectopy. The three categories are mutually exclusive and reflect clinical practice.

2.2. Heart rate metrics

We computed means and standard deviations (SD) for 30 s segments and averaged results for each 10 min segment. Thus our measure of standard deviation is an average of standard deviations of up to 20 30 s segments. To investigate the nonlinear dynamics, we computed the

coefficient of sample entropy (COSEn), detrended fluctuation analysis (DFA) and the local dynamics score (LDs).

2.3. Coefficient of sample entropy (COSEn)

COSEn is an entropy measure derived from the sample entropy (SampEn) and designed specifically to detect AF in very short RR interval time series (Lake and Moorman 2011). Generally, entropy estimators measure the degree of regularity of a signal by counting how many template patterns repeat themselves. Repeated patterns imply order and lead to low values of entropy. In particular, sample entropy is the negative natural logarithm of the conditional probability that two sequences of length m that match within tolerance r will also match at the $m + 1$ st point. Defining as A the total number of matches of length $m + 1$ and B the total number of matches of length m , then the conditional probability (cp) is equal to A/B . The sample entropy is computed as:

$$\text{SampEn} = -\ln(cp) = -\ln(A/B) = -\ln(A) + \ln(B) \tag{1}$$

If A and B are equal, which means that the time series is very regular, the entropy measure is zero, whereas if A is smaller than B , this leads to a higher value of entropy.

The choice of the parameters m and r is crucial in order to obtain a reliable estimation of the conditional probability, especially in very short time series. Lake (2006) proposed the conversion of the measured conditional probability to a probability density. The new measure was named quadratic sample entropy (QSE) and it consists of normalizing the sample entropy by the volume of each matching region, $(2r)^m$. Equation (1) becomes

$$\text{QSE} = -\ln\left(\frac{A}{(2r)^{m+1}}\right) - \left(-\ln\left(\frac{B}{(2r)^m}\right)\right) = -\ln(cp) + \ln(2r) = \text{SampEn} + \ln(2r) \tag{2}$$

QSE allows direct comparison of results obtained by using different values of r . Regression analyses showed that heart rate was an important independent predictor of AF (Lake and Moorman 2011). Hence, the COSEn measure requires the subtraction of the natural logarithm of the mean RR interval:

$$\text{COSEn} = \text{SampEn} + \ln(2r) - \ln(\text{mean}(\text{RR})) \tag{3}$$

In this work the choice of r and m reflects the standards and findings already described in (Lake and Moorman 2011), therefore we chose $m = 1$ and $r = 30$ ms. We calculated COSEn over 30 s segments, consistent with the clinical idea that AF must usually last 30 s to be considered (Fuster *et al* 2006).

2.4. Detrended fluctuation analysis

DFA quantifies fractal-like scaling properties of RR interval time series (Peng *et al* 1995). The interbeat intervals time series of total length N is first integrated,

$$y(k) = \sum_{i=1}^k [B(i) - B_{\text{ave}}] \tag{4}$$

where $B(i)$ is the i th interbeat interval and B_{ave} is the average interbeat interval. Next, the integrated time series is divided into non-overlapping boxes of equal length, n . In each box a least-squares line is fitted to represent the trend in that box. Let $y_n(k)$ be the y coordinate of the straight line segments. The integrated time series, $y(k)$, is detrended by subtracting the local

trend $y_n(k)$ in each box. The root-mean-square fluctuation of this integrated and detrended time series is:

$$F(n) = \sqrt{\frac{1}{N} \sum_{k=1}^N [y(k) - y_n(k)]^2} \quad (5)$$

This computation is repeated over all the box sizes. The results from each box length are averaged—fewer points contribute to the plotted value for higher n . Typically, $F(n)$ will increase with the box size n . A linear relationship on a log–log plot of $F(n)$ versus the box size n indicates the presence of power law (fractal) scaling. Under such conditions, the fluctuations can be characterized by a scaling exponent. In fact, if the data are long-range correlated, $F(n)$ increases as a power-law with respect to n , $F(n) = n^\alpha$. Thus, the fluctuation scaling exponent can be determined by a linear fit on the log–log plot. For uncorrelated data, such as white noise, the scaling exponent is $\alpha = 0.5$. A slope larger than 0.5 indicates persistent long-range correlations. In contrast, $0 < \alpha < 0.5$ indicates an anti-persistent type of correlation. The slope equal to 1 is the theoretical value which corresponds to $1/f$ noise, and $\alpha = 1.5$ to Brownian noise. The original calculation of the DFA was proposed for long signals, such as 24 h Holter ECG recordings, but it was used also in shorter interval time series (Peña *et al* 2009). In this work the scaling exponent α was calculated on the 10 min segments only over the box size range of 4 to 12, consistent with other measures computed here on short segments.

2.5. Local dynamics score (LDs)

The LDs is a new index to investigate the local dynamics of short RR series (Moss *et al* 2014). The new idea is to examine how often individual templates in a short series match each other. Given a 12-beat segment, the algorithm consists mainly into counting the number of times each sample matches with the other 11 with a tolerance r of 20 ms. A histogram of the count of templates as a function of the number of matches is constructed. If no points match, a bar of 12 counts appears in the bin 0, and all the other bins are empty, when all 12 points match each other, the histogram will have a bar of 12 counts in bin 11.

The LDs is computed as a linear combination of the values in bin 0, bin 10 and bin 11; the coefficients were normalized so as to sum to 1. For uniform distribution of matches (the counts in all bins are 1), the LDs is 1. Lower scores imply a bell-shaped distribution, and higher scores imply a distribution concentrated on either or both extremes of the histogram. Here, we calculated the score for every 10 min segment from the average of the 12-beat histograms.

This analysis is expected to report mostly on differentiating NSR from SR with PVCs. We note that the value of LDs in AF is not a useful classifier, and we note that extremely frequent ectopy such as a bigeminal pattern yields a low value, as beats are concentrated in the middle bins—in a 12-beat bigeminal series, each template matches its five siblings, and bins 0, 10 and 11 are empty.

2.6. Statistical analysis

One-way Kruskal–Wallis ANOVA was used to compare the index values among the three groups (NSR, AF, SR with ectopy) and post-hoc multiple comparisons were performed by the Wilcoxon rank-sum test using the Bonferroni correction. For this univariate statistical analysis only a single 10 min segment for each patient was randomly selected. To test the overall hypothesis that dynamical measures were useful for rhythm classification, we used several schemes. The strategy was to compare the accuracy of classification using means and

SD alone with the accuracy after addition of the dynamical measures COSEn, LDs and DFA. The first scheme was a system of three multivariate logistic regression models, each used to distinguish one rhythm class from the other two. The final classification of the RR series was obtained by using the highest probability estimate among the three models.

The second approach was a K-nearest neighbors (K-NN) technique (Xiao *et al* 2010), which requires no statistical model. The strategy was to classify a 10 min segment based on the classification of the majority of neighbors. For a new test record x_i , distances from all the points of the training set were sorted in ascending order. Points with the first K smallest distances are chosen and their classifications are used in a popular vote: the new record was classified the same as the most populous classification of the neighbors. We analysed three distance metrics: the Euclidean distance, the standardized Euclidean distance and the Mahalanobis distance and, on the basis of pilot study, selected the standardized Euclidean distance. The accuracy of classification did not increase for $K > 15$, and we used $K = 25$.

The third approach was random forests (Breiman 2001), an ensemble of tree predictors. The principle behind this method is that a group of weak learners can result in a strong learner. Every element is a normal decision tree where, at each node, a number of predictor variables are selected at random from all the predictors. The number of predictors randomly selected at each node was the square root of the number of variables available, and the predictor that provided the best split in terms of maximizing the separation between the observations was chosen. A forest of 100 trees had maximal performance. Each 10 min segment was evaluated by every tree in the forest, and the final classification was based on the popular vote.

All models were validated using a 10-fold cross-validation procedure on the entire dataset.

3. Results

3.1. An example of analysis

Figure 1 shows ECGs of four patients from the UVa Holter database. Four different rhythms are illustrated: NSR (a), AF (b), SR with PVCs (c) and SR with PACs (d).

Figure 2 shows 10 min segments of the RR interval time series from the same ECG recordings displayed in figure 1. In this example we display 10 min records of SR with a high burden of PVCs, equal to 57% (figure 2(c)), and a very high burden of PACs, equal to 73% (figure 2(d)).

The means (SD) of the four series are NSR 1170 (73), AF 749 (161), SR with PVCs 696 (151) SR with PACs 607 (185) ms. It is noteworthy that the SD is as high for the records of SR with ectopy as it is for the AF record, more than twice the variability of NSR. This points to the fundamental limitation of heart rate variability (HRV) as standardly practiced for this task of rhythm classification.

The NSR series led to the lowest value of COSEn, -2.1 , and the AF series led to the highest, -0.5 . SR with PVCs and PACs led to intermediate values: -1.3 and -1.8 , respectively.

Figure 3 shows the DFA results, plotted as the average variance as a function of the box length. As expected, NSR is readily distinguished from AF by its higher slope: 1.1 and 0.5, respectively. Both NSR and AF, though have higher slopes than either of the records with ectopy; 0.2 for the record with PVCs, and 0.3 for the one with PACs. The phenotype of the records with ectopy is a high and unvarying fluctuation, a consequence of the frequent excursions from the baseline. Thus even short boxes contain outliers and lead to large fluctuation values. Since the outlying ectopic beats are uniform in their values and density, the fluctuation values do not change as a function of window lengths.

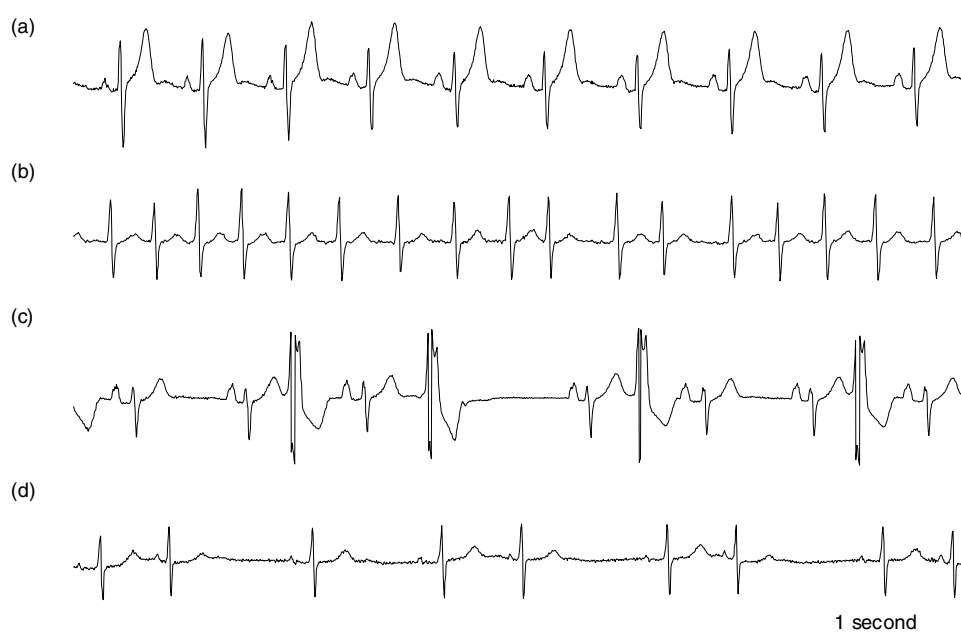


Figure 1. ECGs from four patients of the UVa Holter database in NSR (a), AF (b), SR with PVCs (c) and SR with PACs (d).

LDs was computed on the averaged histogram of template count as a function of template matches calculated every 12 beats. SR with PVCs led to a higher value than NSR, as expected: 1.2 and 0.3, respectively. AF led to variable results (here a high value, 2.0) and LDs is not useful in its classification (Moss *et al* 2014). This record of SR with PACs in a bigeminal pattern led to an intermediate value of 0.5.

3.2. Effect of ectopy on dynamical measures

Figure 4 shows the distribution of each index as a function of the burden of ectopy calculated on all 10 min records, with the 95% confidence interval. Each bin in the figure represents the mean and CI computed every 2% after ranking according to ectopy burden. As expected, ectopic beats decreased the mean RR interval and increased the standard deviation.

Likewise, as expected, ectopic beats increased entropy and the related LDs—after a point, though, entropy fell with increased ectopy. This may be caused by regularized trigeminal and bigeminal rhythms, or by non-sustained atrial and ventricular tachycardia. The DFA slope fell with increasing ectopic burden, as the fluctuation was maximally increased in even the shortest windows.

3.3. Univariate analysis

Table 1 shows the mean values for each group and for each parameter, and the phenotypes of the example series in figures 1 and 2 are recapitulated. NSR had the lowest variability, entropy and LDs, and the highest DFA. AF was found in older patients, and had higher variability and higher entropy. Ectopy was just as variable as AF but had lower DFA.

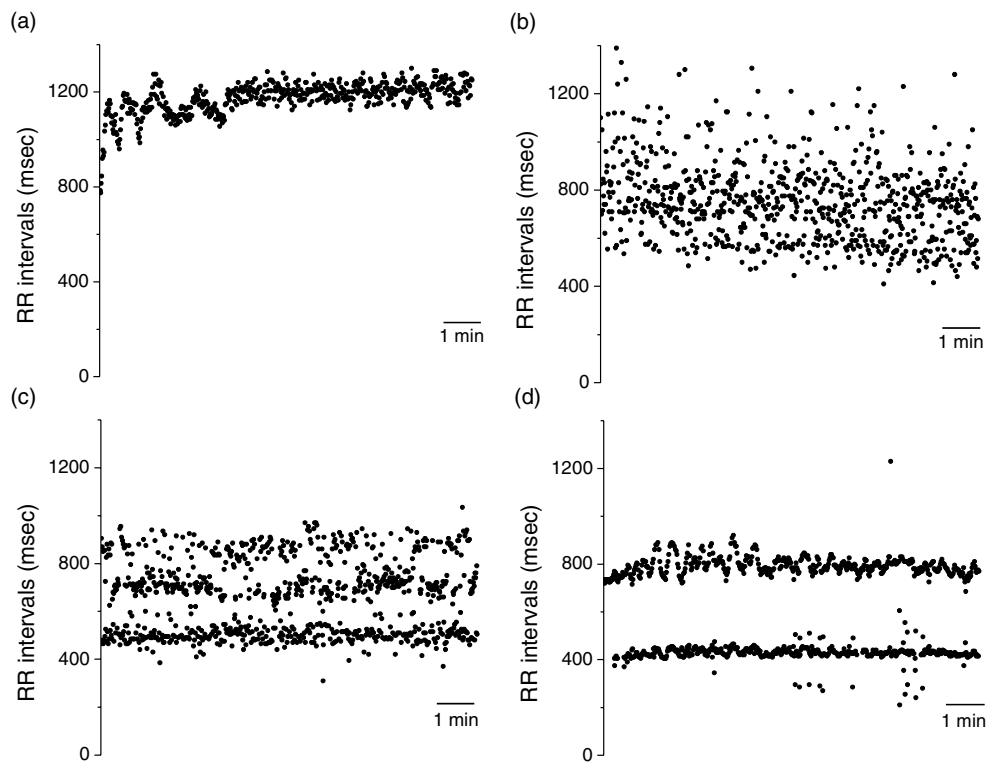


Figure 2. 10 min of RR interval time series from the UVa Holter database for NSR (a), AF (b), and SR with PVCs (c) and SR with PACs (d).

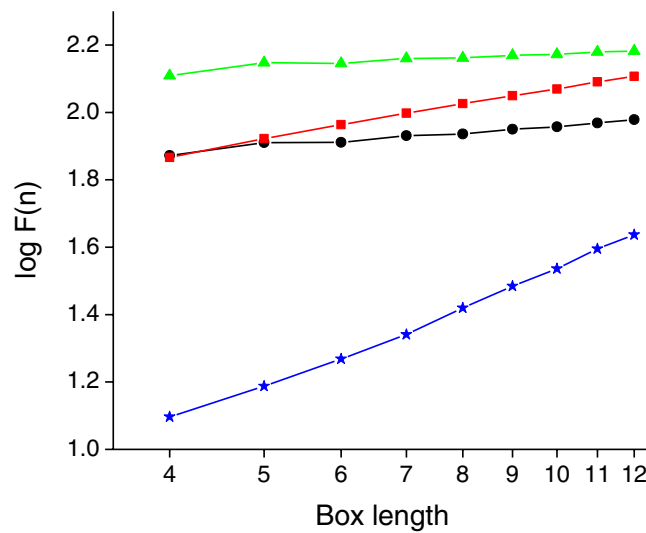


Figure 3. Relationship between the fluctuations function $F(n)$ and the box size n in a double logarithmic plot. Colors and shapes refer to the four different RR interval segments: blue stars corresponds to NSR, red squares to AF, green triangles to SR with PVCs and black dots to SR with PACs.

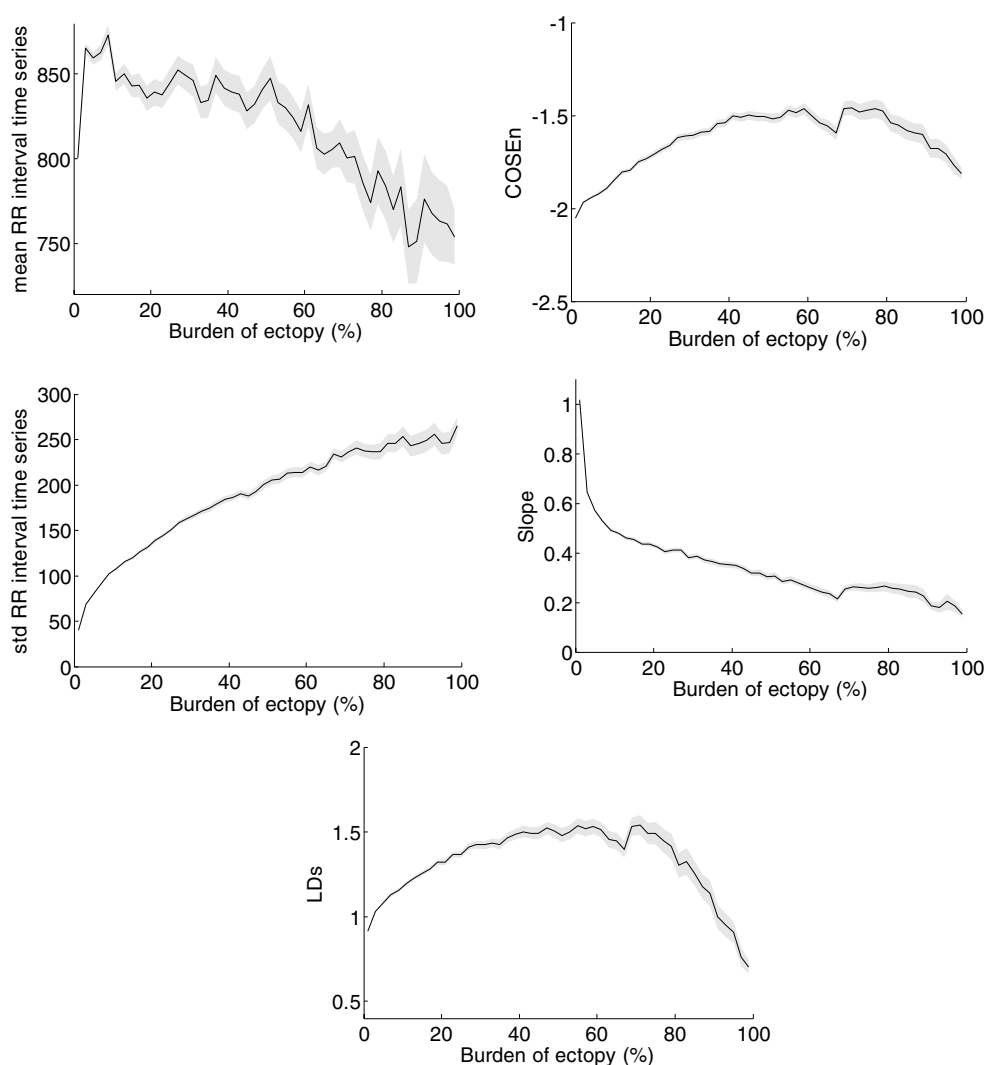


Figure 4. Mean and SD of the RR intervals time series, COSEn, scaling exponent (slope) and LDs as a function of the burden of ectopy. In black is the average trend; the gray region indicates the 95% confidence intervals.

Figure 5 plots the linear (HRV as a function of HR, panels A to C) and nonlinear (COSEn as a function of LDs, panels D to F) measures for all segments. The color represents age (blue is older, A to C) or DFA (blue is higher, D to F).

The goal of the analysis is to distinguish the rhythm categories atrial fibrillation (A and D), normal sinus rhythm (B and E), and sinus rhythm with ectopy (C and F). The major findings are all visually apparent. In the linear and age domain, AF increases with age (A), heart rates in NSR slow with age (B), and SR with ectopy overlaps AF a great deal but occur earlier in life (C). In the nonlinear domain, AF is distinguished by high COSEn (D), and NSR has higher DFA than SR with ectopy (E and F).

Thus no single univariate metric suffices to separate the groups, and we sought to capture these patterns for classification use using multivariable statistical methods.

Table 1. Median and (25th, 75th) percentile values of the parameters for each rhythm group.

Parameters	AF	NSR	SR with ectopy	Kruskal–Wallis test p-value
Mean RR (ms)	763 (660, 884) ^b	795 (676, 928) ^b	820 (707, 942)	< 0.001
Std RR (ms)	149 (116, 186)	38 (24, 58) ^{a,b}	152 (108, 208)	< 0.001
COSEn	-0.5 (-0.8, -0.3) ^b	-2.1 (-2.3, -1.8) ^{a,b}	-1.7 (-2, -1.4)	< 0.001
LDs	1.9 (1.5, 2.4) ^b	0.8 (0.5, 1.3) ^{a,b}	1.3 (1.0, 1.6)	< 0.001
DFA	0.61 (0.56, 0.66) ^b	0.98 (0.66, 1.26) ^{a,b}	0.32 (0.21, 0.5)	< 0.001
Age	72 (65, 79) ^b	47 (23, 66) ^{a,b}	58 (20, 74)	< 0.001

Post-hoc comparisons p -value < 0.05: ^aversus AF, ^bversus SR with ectopy.

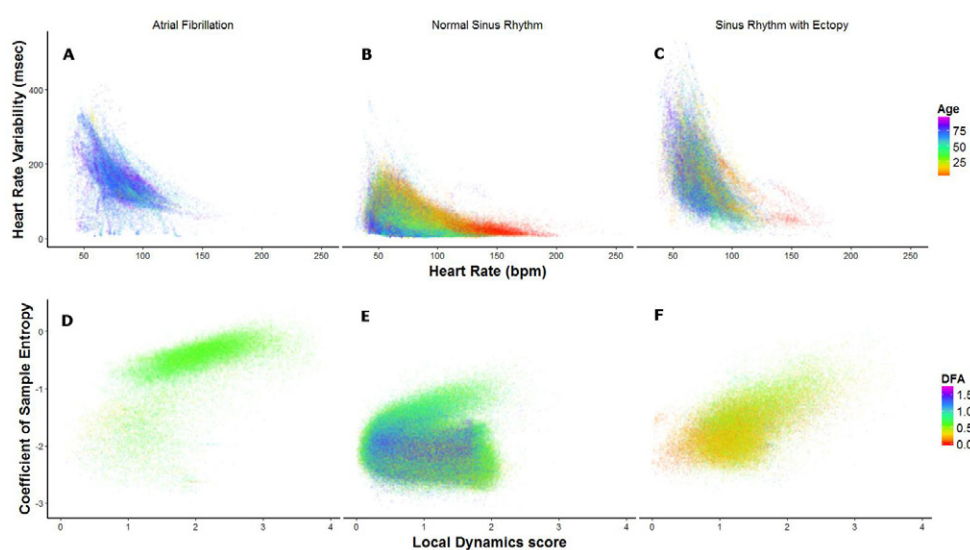


Figure 5. Linear and nonlinear characterization of 377 285 10 min RR interval time series. Panels A to C show linear measures—heart rate variability as a function of heart rate, and age is shown in color. Panels D to F show nonlinear measures—coefficient of sample entropy as a function of local dynamics score, and detrended fluctuation analysis is shown in color.

3.4. Multivariate analyses

The separation of the three rhythm classes in figure 5, especially using nonlinear dynamical measures (panels D to F) suggests that all three multivariable methods should be suitable. The monotonic relationships of the parameters to the classes justifies logistic regression, the distinct separation of neighborhoods justifies k NN analysis, and the simple two-step process of separating AF from non-AF by high entropy, and NSR from SR with ectopy by high DFA justifies decision trees and random forests.

Tables 2 and 3 show the accuracy of labeling 10 min records using the linear measures of HR and HRV along with age (table 2) and the nonlinear dynamical measures of COSEn, LDs and DFA (table 3). Table 4 shows the results for models that combine all features. The columns

Table 2. Contingency matrix of rhythm classification using linear measures—heart rate, heart rate variability, and age.

		AF	NSR	SR with ectopy	PPV
Logistic regression	AF	34%	< 1%	10%	63%
	NSR	13%	99%	31%	92%
	SR with ectopy	53%	1%	60%	62%
k-nearest neighbors	AF	54%	< 1%	20%	60%
	NSR	9%	98%	19%	96%
	SR with ectopy	37%	2%	61%	66%
Random forests	AF	88%	< 1%	6%	89%
	NSR	4%	99%	13%	97%
	SR with ectopy	7%	1%	82%	87%

Percentages are of the 377285 10 min segments analysed. Columns give the classification provided by the Philips Holter software labels and used as a reference. Rows show the classification results from our model. AF, atrial fibrillation; NSR, normal sinus rhythm; SR, sinus rhythm; PPV = positive predictive value.

Table 3. Contingency matrix of rhythm classification using nonlinear dynamical measures—coefficient of sample entropy, local dynamics score, and detrended fluctuation analysis.

		AF	NSR	SR with ectopy	PPV
Logistic regression	AF	80%	1%	6%	83%
	NSR	17%	97%	39%	92%
	SR with ectopy	3%	2%	55%	77%
k-Nearest neighbors	AF	77%	< 1%	2%	93%
	NSR	16%	97%	30%	93%
	SR with ectopy	7%	3%	67%	74%
Random forests	AF	77%	< 1%	2%	93%
	NSR	15%	96%	30%	94%
	SR with ectopy	7%	3%	68%	74%

Percentages are of the 377285 10 min segments analysed. Columns give the classification provided by the Philips Holter software labels and used as a reference. Rows show the classification results from our model. AF, atrial fibrillation; NSR, normal sinus rhythm; SR, sinus rhythm; PPV = positive predictive value.

give the correct classification based on the labels provided by the Holter software. The rows show the classification results from the models we developed. Positive predictive value (PPV) is the proportion of true positive over the total of true and false positive in a classification problem. The averaged values in the contingency tables are approximated to the unity.

The major finding is that the nonlinear dynamical measures led to classification models with higher positive predictive values, and the models using all measures had good accuracy. The random forests analysis led to positive predictive values of 97, 98 and 90% for AF, NSR, and SR with ectopy, respectively.

Clinical practice considers a burden of ectopy to be significant at 10%, and one of our most common misclassification errors was calling a segment NSR when in fact there was slightly

Table 4. Contingency matrix of rhythm classification using all measures.

		AF	NSR	SR with ectopy	PPV
Logistic regression	AF	81%	< 1%	4%	89%
	NSR	10%	99%	19%	96%
	SR with ectopy	9%	< 1%	77%	86%
k-Nearest neighbors	AF	82%	< 1%	4%	90%
	NSR	9%	98%	14%	97%
	SR with ectopy	9%	2%	82%	84%
Random forests	AF	90%	< 1%	1%	97%
	NSR	6%	99%	10%	98%
	SR with ectopy	4%	1%	89%	90%

Percentages are of the 377 285 10 min segments analysed. Columns give the classification provided by the Philips Holter software labels and used as a reference. Rows show the classification results from our model. AF, atrial fibrillation; NSR, normal sinus rhythm; SR, sinus rhythm; PPV = positive predictive value.

more than 10% ectopy. Thus we tested the effect of varying the diagnostic threshold for ectopy from 4 to 20% in steps of 2%—as we raised the threshold, the accuracy for detecting SR with ectopy improved from 55.5% (4% burden) to 89.7% (20% burden), while the accuracy for detecting NSR decreased 4%. The overall accuracy peaked at 94.4% for a threshold of 14% burden.

3.5. Continuous behavior of the classifier

We tested the effectiveness of the classifier in a continuous context, mimicking real-time implementation. Figure 6 shows the performance of the 10 min model by adopting a two minutes update in one of the Holter recordings. Changes in heart rhythm are faithfully reported.

4. Discussion

We studied cardiac rhythm classification using linear and dynamical measures of RR interval time series to discriminate among NSR, AF and SR with ectopy. In 2722 24h ambulatory ECG recordings from patients divided into 10 min RR series, we measured *entropy* using coefficient of sample entropy, *local dynamics* using a new score, and *fractal scaling* using detrended fluctuation analysis. We tested the hypothesis that these measures added to standard measures of HR and HRV using model-based and model-free statistical classifiers. Our major finding is improved rhythm classification using the nonlinear dynamical measures. Of particular importance was accuracy in classifying SR with ectopy, a clinical finding with prognostic impact. The threshold to classify a segment as SR with ectopy, as opposed to NSR, was only a 10% ectopic burden—this reflects clinical practice, and the accuracy increased further if the threshold for diagnosing ectopy was higher.

4.1. Comparison with prior studies

Several groups have addressed the arrhythmia detection problem based on HRV signals using linear and nonlinear measures and different classifiers (table 5). All the papers cited

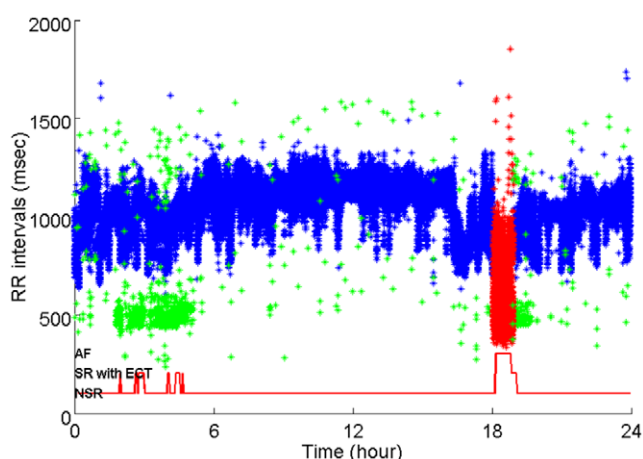


Figure 6. A 24 h Holter recording with demonstration of model output, updated every 2 min. Blue dots represent RR intervals labeled as NSR, green as PVCs, and red as AF. The model output depicted along the lower edge of the plot correctly identifies distinct changes between all three rhythms.

Table 5. Review of literature on rhythm classification.

Reference	Database	Method	Ectopy?
Tateno and Glass 2001	MIT-BIH	Linear: KS two-sample test of ECDF	—
Sarkar <i>et al</i> 2008	MIT-BIH	Nonlinear: Poincare plot	—
Huang <i>et al</i> 2011	MIT-BIH	Linear measures	A, V
Bárdossy <i>et al</i> 2014	MIT-BIH and ICD database	Linear measures, fuzzy logic	A and V combined
Asl <i>et al</i> 2008	MIT-BIH	Linear and nonlinear: GDA and SVM	V
Lake and Moorman 2011	UVa Holter	Nonlinear: entropy	—
DeMazumder <i>et al</i> 2013	ICD database	Nonlinear: entropy	—
Acharya <i>et al</i> 2004	MIT-BIH	Nonlinear measures, fuzzy logic and ANN	V
This paper	UVa Holter	Linear and nonlinear: moments, entropy, DFA, LDs.	A and V separately and combined

KS = kolmogorov–smirnov, ECDF = empirical cumulative distribution function, GDA = generalized discriminate analysis, SVM = support vector machine, DFA = detrended fluctuation analysis, LDs = local dynamics score, ANN = artificial neural network, A = atrial ectopy, V = ventricular ectopy.

include AF in their database, but few of them take into account atrial or ventricular ectopy. Only Huang *et al* (Huang *et al* 2011) and Bardossy *et al* (Bárdossy *et al* 2014) consider PVCs and PACs; the first employed linear measures to identify the transitions between AF and NSR, whereas the second proposed a diagnostic algorithm for implantable cardioverter defibrillators. This reached a very high accuracy, but the real innovation of that work lay in applying a new fuzzy logic-based scheme to standard measures of onset and instability.

4.2. Limitations

Our clinical emphasis constrains our analysis in several ways. First, we assigned the diagnosis of AF when as little as 30 s, or 5% of a 10 min segment was present. This can generate a misclassification of some AF segments into NSR or SR with ectopy, but it is consistent with clinical practice, where episodes lasting this long elicit full consideration in authoritative guidelines (Fuster *et al* 2006), count toward AF burden (Glotzer *et al* 2009), and thus can lead to full AF treatment measures. Second, we consider atrial flutter (AFL) to be the same as AF. This approach, which is based on the similarities in clinical management, is certain to lead to misclassifications. The dynamics of AFL when AV conduction is fixed can never be considered the same as AF, and the regularity with which AFL rhythm can present is responsible for most of the 7 to 10% misclassification error of AF as NSR. Finally, we assign the diagnosis of ectopy when as little as 10% is present, as noted above. This cause an erroneous classification of this portion of low-burden records into NSR, but a burden of 10% can induce clinician to evaluate the use of anti-arrhythmic drugs or ablation procedures, especially if PVC cardiomyopathy is suspected.

The end result of these classification decisions leads us to diagnoses of AF and SR with ectopy even when the rhythm is 90% or more purely normal SR.

Surprisingly, these measures do not distinguish well between atrial and ventricular premature beats, a potentially important clinical distinction. Atrial premature beats may be harbingers of AF, and ventricular premature beats are associated with increased mortality, especially when structural heart disease is present.

4.3. Clinical implications

The clinical importance of accurate cardiac rhythm classification is related to the specific treatments that the presence of the arrhythmia requires. For example, the distinction between AF and SR with ectopy can be difficult without PQRST waveforms, but is nonetheless important for two reasons. First, the diagnosis of AF calls for decisions about anticoagulation, rate control and rhythm control. Second, atrial ectopy may presage AF, and ventricular ectopy may lead to cardiomyopathy or in other ways increase the risk of mortality (Lee *et al* 2012a, Lee *et al* 2012b). An atrial ectopy burden of even less than 1% increases the risk of AF over the next 5 to 15 years (Dewland *et al* 2013). More than 10% ventricular ectopy can be associated with LV dysfunction and clinical heart failure syndromes that are reversed by ablation of the ectopic site (Baman *et al* 2010, Lee *et al* 2012a, Yokokawa *et al* 2013).

4.4. Applications

The utility of this classification are represented by its ability to work on segments of only 10 min, allowing it to be potentially useful in a real time context. Thus these algorithms should be useful for personal monitoring devices. In addition, continuous monitoring in hospitalized patients can be crucially important since changes in the cardiac rhythm are very abrupt and life-threatening. An example of this application is reported in figure 6 where the proposed classifier was proved to be able to track sudden changes in heart rhythm.

5. Conclusion

Multivariable statistical pattern recognition techniques improve on simple measures in correctly distinguishing SR with frequent ectopy, AF and NSR. Moreover, nonlinear measures

add significantly to the mean and the variability. These results should be useful in designing rhythm classifiers from heart rate devices where high fidelity ECG waveforms are not available.

Acknowledgments

Research supported by the American Heart Association and the Italian Ministry of Education, University and Research (MIUR), FIRB2008, Project RBF08VABD. JR Moorman has equity and is Chief Medical Officer of Adult Medical Predictive Devices, Diagnostics and Displays, which has licensed technology for the University of Virginia related to this work.

References

- Acharya R *et al* 2004 Classification of cardiac abnormalities using heart rate signals *Med. Biol. Eng. Comput.* **42** 288–93
- Asl B M, Setarehdan S K and Mohebbi M 2008 Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal *Artif. Intell. Med.* **44** 51–64
- Baman T S *et al* 2010 Relationship between burden of premature ventricular complexes and left ventricular function *Heart Rhythm* **7** 865–9
- Bárdossy A *et al* 2014 Fuzzy logic-based diagnostic algorithm for implantable cardioverter defibrillators *Artif. Intell. Med.* **60** 113–21
- Bravi A, Longtin A and Seely A J E 2011 Review and classification of variability analysis techniques with clinical applications *Biomed. Eng. Online* **10** 90
- Breiman L 2001 Random forests *Mach. Learn.* **45** 5–32
- Cerutti S *et al* 2007 Long-term invariant parameters obtained from 24 h Holter recordings: a comparison between different analysis techniques *Chaos* **17** 015108
- DeMazumder D *et al* 2013 Dynamic analysis of cardiac rhythms for discriminating atrial fibrillation from lethal ventricular arrhythmias *Circ. Arrhythm. Electrophysiol.* **6** 555–61
- Dewland T *et al* 2013 Atrial ectopy as a predictor of incident atrial fibrillation: a cohort study *Ann. Intern. Med.* **159** 721–8
- Fuster V *et al* 2006 ACC/AHA/ESC 2006 guidelines for the management of patients with atrial fibrillation—executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the European Society of Cardiology Committee for Practice Guidelines *J. Am. Coll. Cardiol.* **48** 854–906 (PMID: 16904574)
- Glotzer T V *et al* 2009 The relationship between daily atrial tachyarrhythmia burden from implantable device diagnostics and stroke risk: the TRENDS study *Circ. Arrhythm. Electrophysiol.* **2** 474–80
- Healey J *et al* 2012 Subclinical atrial fibrillation and the risk of stroke *New Engl. J. Med.* **366** 120–9
- Huang C *et al* 2011 A novel method for detection of the transition between atrial fibrillation and sinus rhythm *IEEE Trans. Biomed. Eng.* **58** 1113–9
- Huikuri H V, Mäkikallio T H and Perkiömäki J 2003 Measurement of heart rate variability by methods based on nonlinear dynamics *J. Electrocardiol.* **36** 95–9
- Lake D E 2006 Renyi entropy measures of heart rate Gaussianity *IEEE Trans. Biomed. Eng.* **53** 21–7
- Lake D E and Moorman J R 2011 Accurate estimation of entropy in very short physiological time series: the problem of atrial fibrillation detection in implanted ventricular devices *Am. J. Physiol. Heart Circ. Physiol.* **300** H319–25
- Lee G K *et al* 2012a Premature ventricular contraction-induced cardiomyopathy: a treatable condition *Circ. Arrhythm. Electrophysiol.* **5** 229–36
- Lee V *et al* 2012b The prognostic significance of premature ventricular complexes in adults without clinically apparent heart disease: a meta-analysis and systematic review *Heart* (British Cardiac Society) **98** 1290–8
- Moss T J, Lake D E and Moorman J R 2014 Local dynamics of heart rate: detection and prognostic implications *Physiol. Meas.* **35** 1929–42
- Park J, Lee S and Jeon M 2009 Atrial fibrillation detection by heart rate variability in Poincare plot *Biomed. Eng. Online* **8** 38

- Peña M *et al* 2009 Applying fractal analysis to short sets of heart rate variability data *Med. Biol. Eng. Comput.* **47** 709–17
- Peng C K *et al* 1995 Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series *Chaos* **5** 82–7
- Prystowsky E N 2000 Management of atrial fibrillation: therapeutic options and clinical decisions *Am. J. Cardiol.* **85** 3–11
- Sarkar S, Ritscher D and Mehra R 2008 A detector for a chronic implantable atrial tachyarrhythmia monitor *IEEE Trans. Biomed. Eng.* **55** 1219–24
- Schmitt D T and Ivanov P C 2007 Fractal scale-invariant and nonlinear properties of cardiac dynamics remain stable with advanced age: a new mechanistic picture of cardiac control in healthy elderly *Am. J. Physiol.* **293** R1923–37 (PMID: 17670859)
- Tarkiainen T H *et al* 2008 Preoperative alterations in correlation properties and complexity of R–R interval dynamics predict the risk of atrial fibrillation after coronary artery bypass grafting in patients with preserved left ventricular function *J. Cardiovasc. Electrophysiol.* **19** 907–12
- Tateno K and Glass L 2001 Automatic detection of atrial fibrillation using the coefficient of variation and density histograms of RR and Δ RR intervals *Med. Biol. Eng. Comput.* **39** 664–71
- Xiao Y *et al* 2010 Nearest-neighbor and logistic regression analyses of clinical and heart rate characteristics in the early diagnosis of neonatal sepsis *Med. Decis. Mak.* **30** 258–66
- Yokokawa M *et al* 2013 Recovery from left ventricular dysfunction after ablation of frequent premature ventricular complexes *Heart Rhythm* **10** 172–5