

# Entropy Measures in Heart Rate Variability Data

Niels Wessel<sup>1</sup>, Agnes Schumann<sup>2</sup>, Alexander Schirdewan<sup>3</sup>,  
Andreas Voss<sup>2</sup>, Jürgen Kurths<sup>1</sup>

<sup>1</sup> University of Potsdam, Am Neuen Palais 10, PF 601553,  
D-14415 Potsdam, Germany  
{niels, jkurths}@agnld.uni-potsdam.de

<sup>2</sup> University of Applied Sciences Jena, Carl-Zeiss-Promenade 2, PF 100314  
D-07745 Jena, Germany  
voss@fh-jena.de, schumann@fvk-berlin.de

<sup>3</sup> Franz-Volhard-Hospital, Humboldt-University, Berlin,  
Wiltbergstr. 50, D-13125 Berlin, Germany  
schirdewan@fvk-berlin.de

**Abstract.** Standard parameters of heart rate variability are restricted in measuring linear effects, whereas nonlinear descriptions often suffer from the curse of dimensionality. An approach which might be capable of assessing complex properties is the calculation of entropy measures from normalised periodograms. Two concepts, both based on autoregressive spectral estimations are introduced here. To test the hypothesis that these entropy measures may improve the result of high risk stratification, they were applied to a clinical pilot study and to the data of patients with different cardiac diseases. The study shows that the entropy measures discussed here are useful tools to estimate the individual risk of patients suffering from heart failure. Further, the results demonstrate that the combination of different heart rate variability parameters leads to a better classification of cardiac diseases than single parameters.

## 1 Introduction

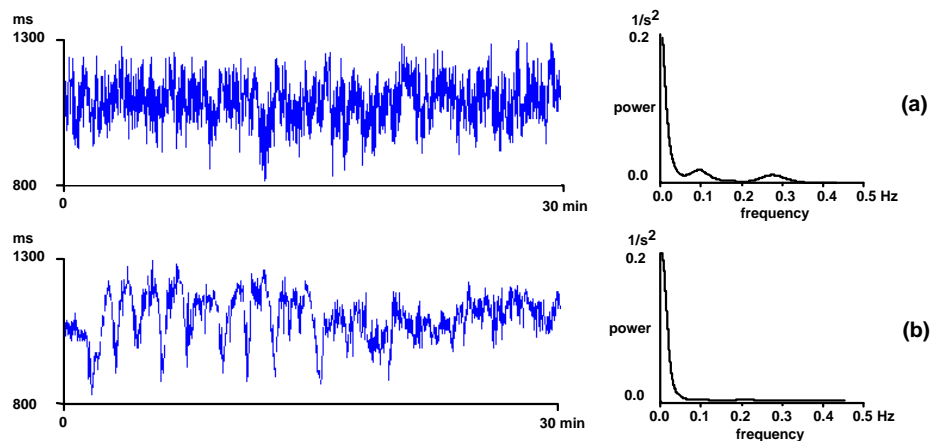
An accurate identification of patients who are at high risk of sudden cardiac death is an important and challenging problem. Heart rate variability (HRV) parameters, calculated from the time series of beat-to-beat-intervals, have been used to predict the mortality risk in patients with structural heart diseases [1,2]. Linear parameters only provide limited information about the underlying complex system, whereas nonlinear descriptions often suffer from the curse of dimensionality. This means that there are not enough points in the time series to reliably estimate these nonlinear measures. Therefore, we favour measures of complexity which are able to characterise quantitatively the dynamics even in rather short time series [3-5]. Recently we could demonstrate that a multivariate approach including these nonlinear as well as linear parameters significantly improves the results of risk stratification [6]. Entropy measures have been used widely in HRV analysis with encouraging results. Most

frequently the ‘approximate entropy’ ApEn is used which was firstly applied to heart rate data in [7,8]. Promising applications of ApEn to HRV data are given for example in [9-13]. Other interesting entropy measures are the ‘tone entropy’ [14], the ‘conditional entropy’ [15], the ‘pattern entropy’ [16], the ‘Kolmogorov entropy’ [17] and the entropy measures based on symbolic dynamics [3,4].

In this contribution we introduce two entropy measures based on periodograms of cardiac beat-to-beat intervals. Both measures - the renormalised and the amplitude adjusted entropy - are calculated from the autoregressive spectral estimation of the time series. The basic idea of these methods is to determine the complexity of cardiac periodograms, however, the renormalised entropy needs and the amplitude adjusted entropy does not need a reference distribution. In this study we investigate the ability of both entropy measures to distinguish between healthy persons and cardiac patients in a clinical pilot study. For the distinction between different kinds of cardiac diseases it is assessed in a multivariate approach whether the amplitude adjusted entropy contributes significantly to other traditional heart rate variability measures.

## 2 Methods

Applications of renormalised entropy to heart rate data based on the Fast Fourier Transform were previously introduced in [3,4]. To overcome the potential lack of reproducibility and time instability of this measure, the autoregressive method  $RE_{AR}$  was developed. Additionally, to avoid the problem of reference selection, the amplitude adjusted entropy  $AE_{AR}$  is introduced here. Figure 1 gives two examples of tachograms, i.e. the time series of the beat-to-beat intervals and the corresponding autoregressive spectral estimations.



**Fig. 1.** Tachograms and autoregressive spectral estimations (a) of a healthy person with normal low and high frequency modulations and (b) of a cardiac patient with a single dominant peak in the very low frequency domain – low and high frequency modulations are absent

The low and high frequency oscillations (0.05-0.4 Hz) are rather low in comparison to the very low frequency peak (<0.05 Hz). An autoregressive spectral estimation is useful to emphasise the different spectral domains, the amplitude adjustment guarantees the comparability of different spectral distributions.

## 2.1 Renormalised Entropy $RE_{AR}$

To compare the degree of complexity of one distribution in relation to a given reference distribution, the latter one is renormalised to a fixed energy. Considering two tachograms with density estimates  $f_0(x)$  and  $f_1(x)$  and using the estimate  $f_0(x)$  as a reference, the renormalised density distribution  $\bar{f}_0(x)$  of  $f_0(x)$  is defined as:

$$\bar{f}_0(x) := \frac{f_0(x)^T}{\int f_0(x)^T dx} \quad (1)$$

where the parameter T is the solution of the integral equation

$$\int \ln f_0(x)^{\bar{f}_0(x)-f_1(x)} dx = 0. \quad (2)$$

The solution of (2) has to be determined numerically. The renormalised entropy  $RE_{AR}$  of the distribution  $f_1(x)$  is defined by the following interchanging algorithm; where  $S(f)$  is the Shannon-entropy of distribution  $f$ , that is

$$S(f) = -\int f(x) \cdot \ln f(x) dx. \quad (3)$$

- I. Calculating of  $\Delta_1 = S(f_1) - S(\bar{f}_0)$  with the distribution  $f_0(x)$  as the reference ( $f_0(x)$  is renormalised). The value of T is denoted  $T_1 = T$ .
- II. Calculating of  $\Delta_2 = S(f_0) - S(\bar{f}_1)$  with the distribution  $f_1(x)$  as the reference ( $f_1(x)$  is renormalised). The resulting T value is denoted  $T_2 = T$ .
- III. If  $T_1 > T_2$ , the distribution  $f_0(x)$  is found to be the more disordered one (in the sense of renormalised entropy - i.s.r.e.) and the renormalised entropy is defined as  $RE_{AR} = \Delta_1$ . Otherwise ( $T_1 < T_2$ )  $f_1(x)$  is the more disordered distribution (i.s.r.e.) and the renormalised entropy is  $RE_{AR} = -\Delta_2$ .

Calculating the renormalised entropy requires estimating the tachogram distributions. Here we use an autoregressive spectral estimation of the filtered and interpolated tachogram. To overcome bias problems a sinusoidal oscillation with a fixed amplitude and frequency was added to the time series. The amplitude of 40 msec was chosen to obtain a dominant peak in the spectral estimation and the frequency was set to 0.4 Hz, which is the upper limit of the high frequency band [18]. A spectral density estimation in the interval [0,0.42] Hz was used to include all

physiological modulations and the calibration peak. Using a reference tachogram from a healthy subject with normal low and high frequency modulations the  $RE_{AR}$  method is designed so that either a decreased HRV or a pathological spectrum leads to positive values of renormalised entropy.

## 2.2 Amplitude adjusted entropy $AE_{AR}$

The technique described in the last section requires determining a reference state. This can be done easily by finding the most disordered spectrum of all data sets from a given control group. But, when analysing new data sets, the problem arise, which distribution should be selected for reference. The most disordered from all control group data sets or should we select for each study an own reference state? The latter choice could lead to an incomparability between different studies. The motivation for designing the amplitude adjusted entropy, therefore, was to find a method which is able to estimate the complexity of a given periodogram independently from a reference state. How can this be done?

One main objective in assessing spectral estimations is to determine phases with a decreased heart rate variability, therefore, the amplitude adjustment described above was adopted. A sinusoidal oscillation with an amplitude of 40 msec and a frequency of 0.4 Hz was superimposed to the original time series. In this way we obtained comparable variability values since they refer to the uniform superimposed variability. A second objective in HRV analysis is the determination of pathological spectra with only singular dominant peaks, since the spectral distributions of healthy persons normally have several peaks due to different cardiovascular modulations. To quantify the intensity of these modulations, the Shannon entropy of the amplitude adjusted spectrum is calculated, i.e. the amplitude adjusted entropy  $AE_{AR}$  is given by

$$AE_{AR} = S(\hat{f}) = -\int \hat{f}(x) \cdot \ln \hat{f}(x) dx \quad (4)$$

where  $\hat{f}(x)$  is the spectral estimation of the time series superimposed by a uniform sinusoidal oscillation. Correspondingly the Shannon entropy of the original (not amplitude adjusted) periodogram is denoted by  $E_{AR}$ .

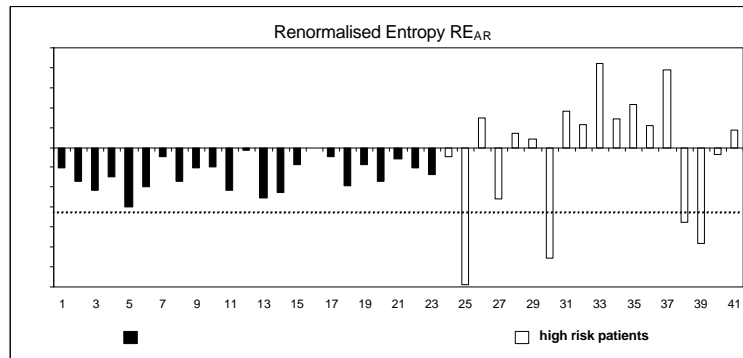
## 3 Results

### 3.1 Clinical pilot study

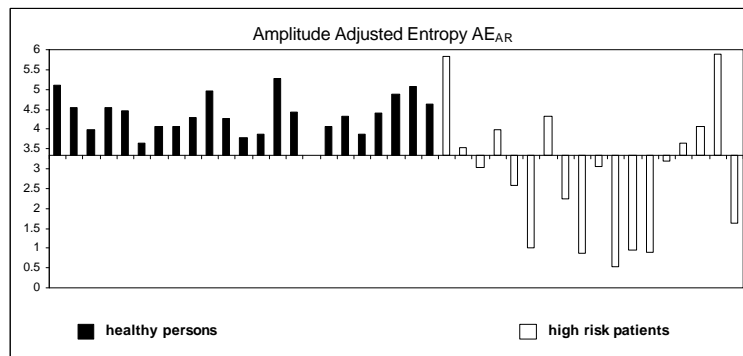
In a clinical pilot study the renormalised entropy  $RE_{AR}$  and the amplitude adjusted entropy  $AE_{AR}$  were applied to data of 18 cardiac patients and 23 healthy subjects. The cardiac patient group consisted of patients after myocardial infarction with documented life threatening ventricular arrhythmias. From the group of healthy subjects, the most disordered tachogram (i.s.r.e.) was determined as the reference for

$RE_{AR}$  calculation ( $RE_{AR}=0$  for healthy person no. 16). The results of this clinical pilot study are shown in Figure 2. The Renormalised entropy  $RE_{AR}$  correctly recognised 15 of 18 high risk patients (with the classification rule: greater zero or under the dotted line at  $-0.33$ ). The Kolmogorov-Smirnov-Z test showed clearly significant differences between both distributions ( $p<0.001$ ). Due to the bimodal distribution of  $RE_{AR}$  in the cardiac patient group, however, no statistical significance could be achieved with the two-tailed t-test for equality of means.

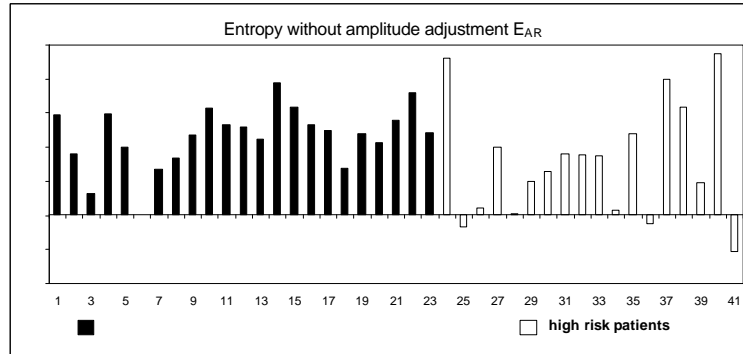
The classification based on the amplitude adjusted entropy reaches a comparable sensitivity at the 100%- specificity level. 11 of 18 cardiac patients were recognised by the classification rule  $AE_{AR}<3.34$  (3.34 is the minimum of  $AE_{AR}$  in the healthy group, compare Fig. 3). The t-test for the equality of means, however, was highly significant ( $4.34\pm 0.49$  for the healthy vs.  $2.85\pm 1.65$  for the high risk group,  $p<0.001$ ) because of the unimodal distribution of  $AE_{AR}$  in both groups.



**Fig. 2.** Results of renormalised entropy  $RE_{AR}$  in a clinical pilot study, black bars represent the control group whereas white bars refer to the cardiac patients



**Fig. 3.** Results of amplitude adjusted entropy  $AE_{AR}$ , black bars represent the control group whereas white bars refer to the cardiac patients



**Fig. 4.** Results of the Shannon entropy without amplitude adjustment  $E_{AR}$ , black bars represent the control group whereas white bars refer to the cardiac patients

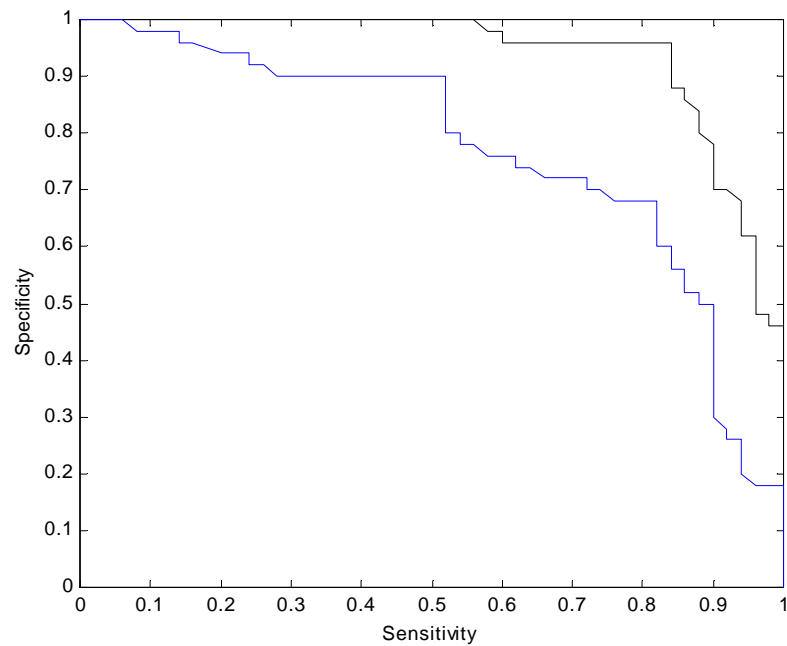
Figure 4 demonstrates the effect of leaving the amplitude adjustment out: The entropy values of the healthy and the high risk group highly overlap. The t-test for equality of means showed no significant differences ( $4.66 \pm 0.45$  for the healthy vs.  $4.29 \pm 0.85$  for the high risk group).

### 3.2 Multiparametric study

The task of separating different cardiac patient groups on the basis of HRV parameters is a demanding problem. If there are differences, it might be possible to find non-invasive marker for specific cardiac diseases. Or, is there even a potential to estimate the severity of cardiac diseases non-invasively? We studied, whether the HRV behaviour of patients suffering from dilated cardiomyopathy (DCM, 41 male, 9 female, age  $52 \pm 10$  years) and patients which have survived an acute myocardial infarction (MI, 42 male, 8 female, age  $58 \pm 9$  years) can be distinguished. So, standard HRV parameters from time and frequency domain were derived from the whole 24h-time series as well as mean values of 5-minute-segments (see table 1). Additionally, the amplitude adjusted entropy values were calculated to assess, whether they provide new information which is not contained in the standard parameter set. The renormalised entropy was not regarded in this study to avoid the problem of reference selection. Since the variability characteristics of both patient groups highly overlap in all individual features a multiparametric approach is pursued. Correlation analysis reveals that the amplitude adjusted entropy calculated from the whole time series is not highly correlated with any standard parameter (pearson correlation coefficient  $< 0.5$ ). This is not true for the short term amplitude adjusted entropy which exhibits correlations above 0.7 with the Shannon and Renyi entropy of the original time series, and with LF, RMSSD and pNN50. Linear discriminant analysis is used to distinguish between the two different patient groups, since this technique is able to construct a linear class boundary which is optimally adjusted to correlations between parameters. To assess the class separability of different parameter

**Table 1.** Standard parameters used in the multiparametric study. For a precise definition see [4,18].

Long term parameters	meanNN	- mean duration of NN
	sdNN	- standard deviation of NN
	sdaNN5	- standard deviation of NN averaged over 5 min
	rmssd	- root mean square of successive NN differences
	pNN50	- the percentage of NN differences > 50 msec
	Shannon	- Shannon entropy of the histogram
	ULF, VLF, LF, HF	- frequency components of the power spectrum
Short term parameters	Shannon	- Shannon entropy of the histogram
	Renyi	- Renyi entropy of the histogram (order 0.25)
	VLF, LF, HF	- frequency components of power spectrum
	LF/HF	- proportion of frequency components



**Fig. 5.** Receiver operator characteristics achieved with the amplitude adjusted entropy (solid line) and without it (dotted line)

sets, crossvalidated recognition rates were calculated providing nearly unbiased estimates of the performance [19-21]. Then all possible parameter sets consisting of 4 parameters were ranked.

It turned out, that the entropy measure  $AE_{AR}$  contributes significantly to the classification performance between the two cardiac diseases. The combination of short and long term  $AE_{AR}$  with meanNN and sdaNN yields a crossvalidated recognition rate of 89 %. Leaving the amplitude adjusted entropy out of consideration, a maximum separation of 75 % could be achieved with a parameter set consisting of all Shannon and Renyi entropies of the histogram combined with the normalised very low frequency component between 0.0033Hz and 0.04Hz. The improvement is demonstrated in Figure 5 where the receiver-operator-curves of both parameter sets are compared. For each sensitivity value a larger value of specificity can be obtained using  $AE_{AR}$ . The Parameters meanNN and sdaNN5 achieve a higher significance value in the t-test than  $AE_{AR}$ , but classifying only on these two standard parameters leads to a significantly reduced performance of only 67%. Thus, the amplitude adjusted entropy provides important additional information for cardiac disease classification in this study.

## 4 Discussion

The renormalised entropy  $RE_{AR}$  is designed in such a way that tachograms with a normal variability and typical periodograms have negative values of  $RE_{AR}$ . Either a decreased HRV or pathological spectra (dominant ULF, VLF or LF peaks) lead to positive values of  $RE_{AR}$ . The results of the clinical pilot study confirmed that in general, healthy persons have negative while most high risk patients have positive renormalised entropy values. The results did not achieve statistical significance in comparing the group mean values, however, in an extensive clinical study  $RE_{AR}$  already demonstrated the usefulness for risk stratification [22]. A high statistical significance between survivors of an acute myocardial infarction who survived a two year follow-up and those who died could be shown. Considering these results and the results of the amplitude adjusted entropy  $AE_{AR}$  in the clinical pilot study, the application of  $AE_{AR}$  to such a large data base seems to be very promising.

For the distinction of DCM and MI patients, which is a more difficult problem than the distinction of cardiac patients and healthy subjects, a multiparametric approach is taken to gain from the information augmentation provided by several parameters. To avoid the problem of selecting a reference spectrum only the amplitude adjusted entropy together with standard parameters from time and frequency domain are considered. It could be shown that the combination of the amplitude adjusted entropy measures with standard HRV parameters leads to better classification results in comparison to the results achieved only with standard parameters. In this study, we got an improvement of more than 10%. This implies that the amplitude adjusted entropy measures contain additional information which is not provided by the standard parameters considered. This suggests further that a multivariate approach



using potent HRV parameters together with clinical parameters may be promising in risk stratification tasks.

Calculating renormalised entropy assumes a fixed reference distribution. In this study, the reference distribution was selected using the interchanging algorithm applied to the data of 23 healthy persons. With the amplitude adjusted entropy, a new method is introduced which does not need a reference distribution. It has to be validated, whether the amplitude adjusted entropy has the same or better properties in the classification of patients after acute myocardial infarction. In summary, both entropy methods  $RE_{AR}$  and  $AE_{AR}$  seem to be potent for risk stratification of patients after myocardial infarction.

## References

1. Kleiger, R.E., Miller, J.P., Bigger, J.T., Moss, A.: Decreased heart rate variability and its association with increased mortality after acute myocardial infarction. *Am J Cardiol* 59 (1987) 256-262
2. Tsuji, H., Larson, M.G., Venditti, F.J. Jr, Manders, E.S., Evans, J.C., Feldman, C.L., Levy, D.: Impact of reduced heart rate variability on risk for cardiac events. The Framingham Heart Study. *Circulation* 94 (1996) 2850-2855
3. Kurths, J., Voss, A., Witt, A., Saparin, P., Kleiner, H.J., Wessel, N.: Quantitative analysis of heart rate variability. *Chaos* 5 (1995) 88-94
4. Voss, A., Kurths, J., Kleiner, H.J., Witt, A., Wessel, N., Saparin, P., Osterziel, K.J., Schurath, R., Dietz, R.: The application of methods of non-linear dynamics for the improved and predictive recognition of patients threatened by sudden cardiac death. *Cardiovasc Res* 31 (1996) 419-433
5. Wessel, N., Ziehmann, Ch., Kurths, J., Meyerfeldt, U., Schirdewan, A., Voss, A.: Short-term Forecasting of Life-threatening Cardiac Arrhythmias based on Symbolic Dynamics and Finite-Time Growth Rates. *Phys Rev E* 61 (2000) 733-739
6. Voss, A., Hnatkova, K., Wessel, N., Kurths, J., Sander, A., Schirdewan, A., Camm, A.J., Malik, M.: Multiparametric Analysis of Heart Rate Variability Used for Risk Stratification Among Survivors of Acute Myocardial Infarction. *Pacing Clin Electrophysiol* 21 (1998) 186-192
7. Pincus, S.M., Gladstone, I.M., Ehrenkranz, R.A.: A regularity statistic for medical data analysis. *J Clin Monit* 7 (1991) 335-345
8. Pincus, S.M., Viscarello, R.R.: Approximate entropy: a regularity measure for fetal heart rate analysis. *Obstet Gynecol* 79 (1992) 249-255
9. Pincus, S.M., Goldberger, A.L.: Physiological time-series analysis: what does regularity quantify? *Am J Physiol* 266 (1994) H1643-1656
10. Makikallio, T.H., Seppanen, T., Niemela, M., Airaksinen, K.E., Tulppo, M., Huikuri, H.V.: Abnormalities in beat to beat complexity of heart rate dynamics in patients with a previous myocardial infarction. *J Am Coll Cardiol* 28 (1996) 1005-1011
11. Palazzolo, J.A., Estafanous, F.G., Murray, P.A.: Entropy measures of heart rate variation in conscious dogs. *Am J Physiol* 274 (1998) H1099-1105
12. Vikman, S., Makikallio, T.H., Yli-Mayry, S., Pikkujamsa, S., Koivisto, A.M., Reinikainen, P., Airaksinen, K.E., Huikuri, H.V.: Altered complexity and correlation properties of R-R

- interval dynamics before the spontaneous onset of paroxysmal atrial fibrillation. *Circulation* 100 (1999) 2079-2084
13. Van Leeuwen, P., Lange, S., Bettermann, H., Gronemeyer, D., Hatzmann, W.: Fetal heart rate variability and complexity in the course of pregnancy. *Early Hum* 54 (1999) 259-269
  14. Oida, E., Moritani, T., Yamori, Y.: Tone-entropy analysis on cardiac recovery after dynamic exercise. *J Appl Physiol* 82 (1997) 1794-1801
  15. Porta, A., Baselli, G., Liberati, D., Montano, N., Cogliati, C., Gnecci-Ruscione, T., Malliani, A., Cerutti, S.: Measuring regularity by means of a corrected conditional entropy in sympathetic outflow. *Biol Cybern* 78 (1998) 71-78
  16. Zebrowski, J.J., Poplawska, W., Baranowski, R.: Entropy, pattern entropy and related methods for the analysis of data on the time intervals between heart beats from 24h electrocardiograms, *Phys Rev E* 50 (1994) 4187-4205
  17. Guzzetti, S., Signorini, M.G., Cogliati, C., Mezzetti, S., Porta, A., Cerutti, S., Malliani, A.: Non-linear dynamics and chaotic indices in heart rate variability of normal subjects and heart-transplanted patients. *Cardiovasc Res* 31 (1996) 441-446
  18. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *Circulation* 93 (1996) 1043-1065
  19. Fukunaga, K., Introduction to statistical pattern recognition. Academic Press, New York-London (1974)
  20. Duda, R.D., Hart, P.E.: Pattern Recognition and Scene Analysis. Wiley, New York (1973)
  21. Fahrmeir, L., Hamerle, A.: Multivariate statistische Verfahren. Walter de Gruyter, Berlin-New York (1984)
  22. Wessel, N.: Komplexe Analyse nichtlinearer Phänomene in kardiologischen Datenreihen. Dissertation, University of Potsdam (1998)