

RBF network classification of ECGs as a potential marker for sudden cardiac death

Hans A. Kestler^{1,2} Friedhelm Schwenker¹

Günther Palm¹

¹Department of Neural Information Processing and

²Department of Medicine II – Cardiology

University of Ulm, D-89069 Ulm

Germany

Non-invasive risk assessment after myocardial infarction is a major but still unresolved goal in clinical cardiology. Various parameters such as ventricular late potentials, T-wave alternans, and repetitive ventricular extrasystoles have been shown to indicate an increased risk of sudden cardiac death. However, the practical use of these arrhythmic markers into clinical decision making remains difficult. In this chapter we will describe two approaches of risk stratification with RBF networks using high-fidelity ECG recordings. Based on these high-fidelity recordings different aspects of conduction defects are exemplarily investigated. The first utilizes established features derived from signal averaged QRS complexes (heartbeats) and the second investigation centers on capturing morphology changes within the QRS complex.

1 Introduction

The non-invasive risk stratification of patients prone to sudden cardiac death is an important problem in modern cardiology. The incidence of sudden cardiac death (SCD) in the area of Germany is about 80,000 to 160,000 cases per year. Apart from cases of ventricular fibrillation (VF)

related to myocardial ischaemia, the main reason for SCD is the occurrence of ventricular tachyarrhythmias as a cause of a chronic arrhythmogenic substrate. Sudden cardiac death most often occurs in the presence of coronary artery disease (CAD) (90% of SCD patients). In a significant number of patients (13% – 20%) it is the initial symptom of CAD. Ventricular fibrillation emerges in most cases (70%) secondarily from a ventricular tachycardia.

Various parameters such as ventricular late potentials, T-wave alternans, and repetitive ventricular extrasystoles have been shown to indicate an increased risk of sudden cardiac death. However, the practical use of these arrhythmic markers into clinical decision making remains difficult. The positive predictive value of all non-invasive parameters is limited especially when not combined with a reduced left ventricular function. On the other hand the available therapeutic options, the implantable cardioverter defibrillator or long term amiodarone drug therapy, have side effects, strain the patient and are cost intensive, thus requiring a highly selective usage. Recently the MUSTT [1] and the MADIT [2, 3] study proved the effect of defibrillator therapy in post infarction patients pre-selected by a reduced left ventricular function and spontaneous non-sustained ventricular tachycardia, with inducible sustained ventricular tachycardia during electrophysiologic study. In this chapter we summarize the clinical results of the currently available methods for non-invasive risk assessment and describe two approaches of risk stratification with RBF networks based on high-fidelity ECG recordings. In post-infarction studies with animals it has been shown that the substrate for singular or repeated ventricular arrhythmias is a localized damaged myocardium with abnormal conduction characteristics [4–6]. This causes slow or irregular propagation of activation. It is possible to detect these delayed signals with the high-resolution electrocardiogram. Based on these high-fidelity recordings different aspects of conduction defects are exemplarily investigated. The first is based on established features derived from signal averaged QRS complexes (heartbeats) and mainly describes a prolongation of cardiac excitation that extends beyond the normal heartbeat (**late potentials**). The second investigation centers on capturing morphology changes of the QRS complex (segment of the ECG signal associated with depolarization) in beat-to-beat recordings.

Within this problem context of non-invasive risk stratification the topics of data acquisition, group description and evaluation are treated together

with the initialization and training algorithms of the used RBF networks.

2 Medical Background: Review of Non-Invasive Risk Stratification in Patients after Myocardial Infarction

Even in the thrombolytic era with its generally accepted reduction of hospital mortality, patients with acute myocardial infarction remain at an increased risk of sudden death at least in the first year after the acute event [7]. Based on large multi center studies post myocardial infarction there is still a remarkable rate of total death (7% to 23%) and sudden cardiac death (4% to 8%) within the first one or two years after infarction, particularly in patients with reduced left ventricular function, as seen from the placebo arms of several studies [8–12] (Table 1).

Various non-invasive tests have been developed and clinically evaluated for risk assessment after myocardial infarction to detect subgroups of high risk patients for preventive treatment. From the current concept arrhythmogenesis is an integrative process dependent on the existence of an arrhythmogenic substrate (the infarction zone with abnormal automatism and slowed conduction), the autonomic tone, spontaneous trigger events (ventricular extrasystoles) [13], and other modulating environmental factors (electrolytes, drugs). The available non-invasive tests cover different aspects of this scenario (Table 2). Late potential analysis concentrates on the detection of myocardium with slow conduction. Holter monitoring allows to measure the incidence and complexity of spontaneous trigger events. Baroreflex sensitivity and heart rate variability reflect the autonomic tone. Most recently T-wave alternans has been introduced into clinical application as a marker of repolarization abnormalities.

Ambulatory Electrocardiography. With the ambulatory ECG the different types of ectopic beats can be documented that may trigger the initiation of ventricular tachycardias and/or fibrillation. Among those are frequent and repetitive ventricular premature beats (VPBs) up to non-sustained or sustained ventricular tachycardias. In addition to the quantification of spontaneous arrhythmias, modern Holter ECG equipment allows ST-segment analysis for ischaemia detection, heart rate variabil-

Table 1. Mortality rates after myocardial infarction (total deaths and sudden cardiac deaths) in the era of thrombolysis, as taken from the placebo arms of recent large randomized trials. Abbreviations: SCD: sudden cardiac death, FU: follow-up, yrs: years, AMI: acute myocardial infarction, LVEF: left ventricular ejection fraction, f/r VPBs: frequent (>10/h) or repetitive (≥ 3 beats) ventricular premature beats, HF: heart failure, CHF: congestive heart failure, AIRE: Acute Infarction Ramipril Efficacy study, SAVE: Survival And Ventricular Enlargement study, EMIAT: European Myocardial Infarct Amiodarone Trial, CAMIAT: Canadian Amiodarone Myocardial Infarction Arrhythmia Trial, AMIO: Amiodarone.

Author	Year	Study	Target	Number	Total	SCD	FU (yrs)
AIRE Investigators [11]	1993	AIRE	AMI + HF	992	23.0%	–	1.25
Lamas et al. [8]	1995	SAVE	AMI	946	6.6%	–	1 (3.5)
Julian et al. [9]	1997	EMIAT	AMI, LVEF \leq 40%	743	12.0%	8.2%	1.75
Cairns et al. [10]	1997	CAMIAT	AMI, f/r VPBs	596	8.3%	5.2%	1.79
Amiodarone Trials Meta-Analysis Investigators [12]	1997	AMIO META ANALYSIS (13 Trials)	AMI + CHF	6553 (5101+1452)	12.3%	5.7%	1.4

Table 2. Association of certain electrocardiographic methods and parameters to cellular functions of the myocardium and to electrophysiological mechanisms of tachyarrhythmias. Abbreviations: ECG: electrocardiogram, SAECG: signal averaged surface ECG, VLP: ventricular late potentials, HR: heart rate, HRECG: high resolution ECG.

ECG - Parameter	Cellular Function	VT/VF – Mechanism
SAECG: VLP	Conduction delay (static)	Reentry Bystander area?
Holter-ECG: spontaneous arrhythmias	Trigger events	Abnormal automatism Reentry-induction
Real time-HRECG: QRS variability (QRV)	Dynamic conduction delay	Reentry
Real time-HRECG: variant VLP	Wenckebach conduction delay	Reentry Bystander area?
Surface ECG: QT-Dispersion	Dispersion of local repolarization	Reentry
Holter ECG: QT-Variability	Dynamic behaviour of action potential duration	Reentry Autonomic tone Triggered Activity
Surface ECG: T-wave alternans	Dynamic behaviour of repolarization	Reentry Autonomic tone
Holter ECG: HR-variability	Sympathetic- parasympathetic balance	Reentry? Autonomic tone

ity analysis, the detection of ventricular late potentials by signal averaging, and QT variability analysis. Late potentials may be considered as a marker of delayed depolarization from areas of damaged ventricular myocardium, thus forming one pre-requisite of a reentrant circuit. However, it has been claimed from both experimental and clinical studies that in a significant proportion late potentials may represent only bystander areas, which are not an essential anatomical part of the reentrant circuit. Because dynamic events of ventricular depolarization are lost by the averaging process, beat-to-beat real time high resolution ECG equipment have been developed and clinically tested [14, 15], and indeed increased beat-to-beat variations of the QRS-complex (QRV) and the T-wave have been demonstrated in patients at high risk for malignant arrhythmias [15–17]. Two new ECG methods have been extensively investigated in recent years, aimed at detecting repolarization inhomogeneities, namely QT-dispersion from the conventional 12-lead surface ECG [18] and QT-variability mainly from 24-hour ambulatory ECG recordings [19]. Most recently a special technique and algorithm has been described to detect beat by beat T-wave alternans at the microvolt level [20]. T-wave macro-alternans is a well known phenomenon in severe myocardial ischaemia, experimental myocardial infarction, vasospastic angina and other pathophysiological conditions (neuro-hormonal imbalance). Increased beat-by-beat micro-fluctuations of the amplitude and shape of the T-wave (repolarization process) are thought to reveal an increased local inhomogeneity of repolarization within the entire ventricular myocardium [21].

Assessment of imbalances of the autonomic nervous system. To detect influences of the autonomic nervous system, heart rate variability (HRV) has been tested in short term (5 minutes to hours) and long term (24 hours) approaches, using either surface ECG monitoring or 24 hour Holter recorders, both in the time and frequency domain [22–24]. A large number of time domain indices have been developed and clinically tested such as SDNN, SDANN, RMSSD, NN50 and pNN50 (all statistical methods), and HRV-triangular index, TINN, Differential index and Logarithmic index (all geometrical methods). According to the Task Force Committee on Heart Rate Variability both SDNN and HRV triangular index should be used for the assessment of overall HRV, whereas short term components of HRV may be estimated by SDANN and RMSSD [25]. Power spectral analysis using fast Fourier transform

allows the separation of the total power spectrum into high frequency (HF), low frequency (LF), very low frequency (VLF) and the ultra-low frequency range (ULF). The physiological correlate to HF components is attributed to ventilation, of LF to baroreflexes, of VLF to sympathetic activity (?), and of ULF possibly to the activity of the Renin-Angiotensin system. The LF/HR ratio may reflect sympatho-vagal balance, sympathetic modulations or baroreflex activity [24]. It should be stressed that for physiological and mathematical reasons there are strong correlations between certain time and frequency parameters of heart rate variability: SDNN, HRV triangular index, and TINN correlate to the total frequency power spectrum; SDNN index correlates to mean 5-min total power; RMSSD, SDSD, NN50 count, pNN50, differential index and logarithmic index correlate to the HF spectrum [25]. Changes in HRV have been found in patients after myocardial infarction, in diabetic neuropathy, in transplanted hearts, in cardiac failure and in patients with tetraplegia [25]. A number of drugs like beta-adrenergic blockers, antiarrhythmic drugs and muscarinic receptor blockers have been found to influence HRV. Beta blockers enhance HRV, antiarrhythmics like flecainide and propafenone decrease HRV, and scopolamine paradoxically also decreases HRV. For risk stratification in post-MI patients both SDNN and HRV triangular index proved to be a powerful tool for predicting sudden cardiac death. In the study of Kleiger et al. [22] a SDNN cut-off value of 50-100 ms differentiated well between patients at low risk (50-100 ms) and those at high risk (<50 ms), as did a cut-off value of the HRV triangular index of 20 in the study of Malik et al [23].

Electrophysiological testing (EPS). Between 1982 and 1992 a considerable number of studies have been published on the value of electrophysiological testing (EPS) for predicting the risk of sustained VT or SCD in post-MI patients [26–39]. The parameter used initially, repetitive ventricular response, namely (>3 VPBs following one or two extrastimuli) proved to be too unspecific and over sensitive as a marker of ventricular electrical instability, and today only the induction of sustained VT is considered a meaningful and more specific risk marker for ventricular vulnerability [40]. Among 14 selected prognostic studies of EPS following acute myocardial infarction only 8 proved to be prognostic for SCD (Table 3). The positive predictive value of EPS following MI, ranges between 13% to 42%, whereas the negative predictive value is as high as

95% to 100%.

In the clinical setting the value of all non-invasive methods and risk markers on its own, including of programmed ventricular stimulation, is limited since the positive predictive values of all of these methods range from 4% to 42% (Table 4), i.e. there is a significant number of false positive candidates with a pathological finding in either of these tests, who never will experience a sudden cardiac death in the first one or two years after myocardial infarction. A combination of two or more non-invasive risk markers, e.g. left ventricular ejection fraction below 40%, ventricular late potentials, frequent and/or repetitive extrasystoles, may enhance the positive predictive value to 50% or more, but at the expense of sensitivity, which may drop below 50% [42–45].

Risk stratification in patients with cardiomyopathies and chronic congestive heart failure. Risk stratification in patients with hypertrophic obstructive cardiomyopathy (HOCM) or dilative cardiomyopathy (DCM) seems to be less substantiated, since at least PVS proved to be too insensitive and unspecific for differentiating patients at high risk of SCD. Therefore other risk markers for HOCM such as sudden unexpected death in family relatives, sudden unexplained syncope in the individual patient or syncope on exertion, the presence of sustained VT and a high pressure gradient with exaggerated septal hypertrophy seem to be more predictive [47]. For DCM the degree of left ventricular dysfunction, the clinical status of cardiac failure, and the presence of non-sustained VT may be predictive of increased mortality. The degree of left ventricular dysfunction may also be useful for predicting reduced survival rates in patients with other types of chronic severe heart failure (CHF), which may be tested by the degree of exercise tolerance. This has been shown by Mancini et al. [48], who demonstrated that CHF patients on a waiting list for heart transplantation being able to perform an exercise work load with a peak oxygen uptake of >14 ml/kg/min had an equal survival rate compared to those who already had a heart transplant. In contrast, those patients with a peak exercise capacity of < 14 ml/kg/min had a one year survival rate of only 47%, and a two year survival rate of 32%. The role of left ventricular dysfunction and particularly of repetitive ventricular arrhythmias including non-sustained VT in predicting the risk of sudden arrhythmic death remains unclear in patients with severe CHF. At least in those with sustained VT or arrhythmic syncope an implantable

Table 3. Prognostic significance of programmed ventricular stimulation (PVS) for the incidence of ventricular tachycardia (VT) or fibrillation (VF) and for sudden cardiac death (SCD) in the first to second year after acute myocardial infarction. Abbreviations: n: patient number studied, FU: follow-up, Ind +: inducible by PVS, Ind -: non-inducible by PVS, Rec: recurrences.

Author	Year	n	FU month	Ind+ (%)	Rec (%)	Ind- (%)	Rec (%)	Prognostic for SCD
Breithardt [26]	1982	132	15	46	16	54	4	+
Hamer [28]	1982	70	12	17	33	83	9	+
Richards [29]	1983	165	12	23	21	77	2	+
Marchlinski [30]	1983	46	18	22	6	78	14	-
Gonzalez [27]	1984	84	20	23	0	77	6	-
Waspe [31]	1985	50	23	34	41	61	0	+
Roy [32]	1985	150	10	23	6	72	2	-
Santarelli [33]	1985	50	11	46	0	54	0	-
Bhandari [34]	1985	45	12	44	10	56	3	-
Kowey [35]	1990	187	18	64	21	36	29	-
Iesaka [36]	1990	133	21	19	47	81	3	+
Bourke [37]	1991	1209	12	6	19	94	3	+
Bhandari [41]	1992	86	18	22	32	78	7	+
Steinbeck [39]	1992	152	11	16	13	84	2	+

Table 4. Predictive value of different non-invasive and invasive markers for sudden cardiac death given with their corresponding cut-off values. Abbreviations: VTA: ventricular tachyarrhythmias, SCD: sudden cardiac death, ECG: electrocardiogram, VT: ventricular tachycardia, VPB: ventricular premature beat, QTD: QT-dispersion, TD: time domain, AE: serious arrhythmic event.

Method	Marker	Positive Predictive Value (%)	Negative Predictive Value (%)
Holter ECG	non-sustained VT Pairs \geq 10/24h Salvos \geq 2/24h frequent VPBs $>$ 10/h	10 – 20	70 – 90
Signal Averaged ECG (at rest, Holter)	Ventricular late potentials (2/3 Simson criteria positive)	4 – 30	95 – 99
12 lead ECG	QTD \geq 80ms	inconsistent, 0 – 100 (see Zabel et al. [46])	inconsistent, 0 – 100 (see Zabel et al. [46])
Surface ECG	T-wave alternans ratio \geq 3.0	54 (AE)	90 (AE)
Holter ECG	Heart Rate Variability SDNN \geq 50ms (TD)	17	77
Electrophysiological Study	Repetitive ventricular response, sustained ventricular tachycardia	13 – 42	96 - 100

cardioverter defibrillator (ICD) should be implanted prophylactically in order to bridge the patient alive to the time of heart transplantation.

Sequential non-invasive and invasive risk stratification protocols.

Instead of an alternative application of non-invasive or invasive strategies, a two-step risk stratification procedure using non-invasive screening followed by an invasive electrophysiological study seems to be more appropriate, as has been proposed and investigated by Pedretti et al. [49]. In a total of 303 post-myocardial infarction patients a number of clinical risk markers such as age, previous myocardial infarction, non-Q-wave infarction, left ventricular dyskinesia, ejection fraction < 40%, filtered QRS duration of ≥ 115 ms (VLP), VPBs > 6/hour, presence of non-sustained VT, HRV index of < 29 and mean RR interval < 750 ms were tested by stepwise logistic regression analysis. The following markers proved to be most significant: EF < 40%, VLP present, and detection of repetitive VPBs. Patients with none or one non-invasive risk marker were classified as low risk patients and followed for a mean of 15 months post infarction. Those with two or three non-invasive risk markers were classified as intermediate to high risk and subjected to invasive electrophysiologic testing. Patients with inducible sustained VT were classified as high risk patients, and those without inducibility of sustained VT to the low risk group. The mortality rate for SCD was 13/20 (65%) in the group with inducible sustained VT, whereas mortality in the low risk group was 3/263 patients (1.1%). The sensitivity for predicting SCD by the two-step program was 81%, the specificity 89%, the positive predictive value was 65% and the negative predictive value was 99%, i.e. 65% of patients were correctly classified as SCD candidates. The study of Pedretti et al. [49] has been confirmed in essence by the publication of Zoni-Berisso et al. [50].

Current status. In daily practice the diagnosis of an increased arrhythmic risk has to be followed by the clinical decision whether the risk of the individual patient is high enough to require aggressive and cost intensive treatment strategies such as long term amiodarone or cardioverter defibrillator implantation. Despite the large number of studies dealing with risk assessment from different parameters, data concerning proven therapeutic consequences of risk assessment are still limited to very high

risk patient groups. The MUSTT [1] and MADIT [2, 3] trials ascertain the use of a stepwise approach based on non-sustained ventricular tachycardia (nsVT) and a reduced left ventricular ejection fraction (EF) for the selection of patients for electrophysiologic testing (EPS). In patients with inducible sustained VT implantation of an cardioverter defibrillator was proven to be an effective therapy. Despite the fact that these studies could demonstrate the successful transposition of risk assessment into therapy, there are still many open questions. The level of risk as well as the optimum of equivalent combinations of risk factors have to be defined. Currently all risk assessment strategies are based on a significantly depressed left ventricular function. However it remains to be defined whether LV dysfunction needs to be always heavily reduced or whether certain combinations of risk parameters require treatment even with moderate or low LV-impairment. Only recently Ikeda et al. [51] showed that the combined assessment of T-wave alternans and late potentials is associated with a high positive predictive value for an arrhythmic event after acute MI. Furthermore current risk assessment strategies mostly serve as pre-screening of patients for subsequent programmed ventricular stimulation still regarded as the gold standard of risk evaluation. Although this might be acceptable in post MI patients, EPS does not capture all relevant mechanisms of malignant ventricular tachyarrhythmias [52]. Especially in patients with cardiomyopathy the inducibility of sustained VT during electrophysiological stimulation is less specific. In clinical practice a two step stratification program in post infarction patients seems to be an adequate and cost effective approach. The positive predictive value of all individual non-invasive parameters is limited. However the negative predictive value is relatively high ($> 90\%$) and can be further increased by combining them using the OR relation. Thus pre-selection of high risk patients with a set of non-invasive markers will markedly enhance the specificity and predictive accuracy of the subsequent invasive electrophysiological test. A different test characteristic can be reached by combining several screening parameters in an AND relation resulting in an improved positive predictive value of the subsequent EPS. Furthermore the use of multiple pre-screening non-invasive tests allows to cover different pathophysiological aspects of arrhythmogenesis (see Table 2). A potential risk stratification strategy including both established parameters ($EF \leq 40\%$, non-sustained VT, see MUSTT, MADIT) and new non-invasive tests is shown in Figure 1. After myocardial infarction patients

will first be studied for post-MI ischaemia (angina, ST-segment depression) and will be revascularized if needed (PTCA or bypass grafting). Two to three weeks post MI, left ventricular function will be measured (echo, angiogram, scintigraphy), and a Holter ECG will be recorded. In addition a set of new risk markers including ventricular late potentials, heart rate variability, T-wave alternans, and the recently described heart rate turbulence [53] can be recorded. In patients with $EF \leq 40\%$ and non-sustained VT (which corresponds to MUSTT and MADIT) or a reduced LV function in combination with at least three other risk markers an electrophysiologic study will be performed. Patients not undergoing EP study as well as patients with a negative EP test result will receive beta blocker therapy. Patients with inducible VT (due to the protocol always on top of the above mentioned pre-screening risk factors) represent a true high risk population requiring ICD therapy. Further studies are necessary to clarify the role of long term amiodarone drug therapy plus beta blocker as an alternative to the ICD strategy in certain patient subgroups. New technologies of signal analysis have broadened our apparatus for risk assessment allowing to measure very diverse aspects of arrhythmogenesis. The above proposed strategy is a careful approach to include these new technologies into clinical practice although other combinations of risk markers might replace the reduced left ventricular ejection fraction [51]. Several studies with long term follow-up are necessary to clarify their final role in different patient subgroups and pathophysiological entities.

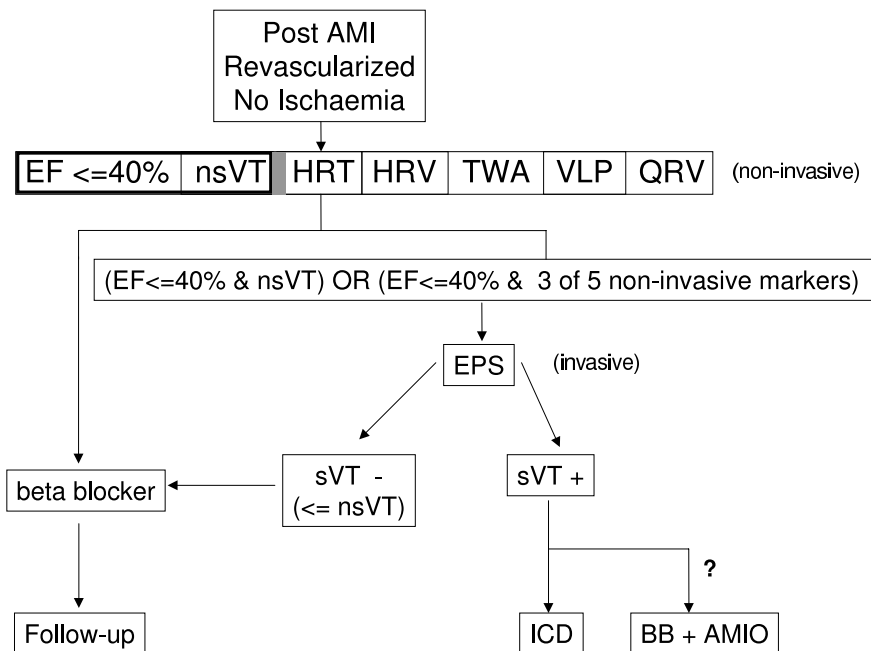


Figure 1. Flow chart showing the proposed two-step risk stratification procedure in patients after acute myocardial infarction (see text). Abbreviations: EF: ejection fraction, VEA: ventricular ectopic activity, HRT: heart rate turbulence, HRV: heart rate variability, TWA: T-wave alternans, VLP: ventricular late potentials, QRV: QRS microvariability, nsVT: non-sustained ventricular tachycardia, EPS: electrophysiological testing (by programmed ventricular stimulation), sVT: sustained VT, ICD: implantable cardioverter defibrillator, BB: beta blocking drug, AMIO: amiodarone.

3 Selected methods for training RBF classifiers

In this section we present the different RBF initialization and training strategies used in this investigation.

In the classification scenario a neural network performs a mapping from a continuous input space $X (= \mathbb{R}^d)$ into a finite set of classes $Y = \{\omega_1, \dots, \omega_l\}$. In the training phase the parameters of the network are determined from a finite training set: $S = \{(\mathbf{x}^\mu, \omega^\mu) | \mu = 1, \dots, N\}$, each feature vector $\mathbf{x}^\mu \in \mathbb{R}^d$ is labeled with its class membership $\omega^\mu \in Y$. In the recall phase further unlabeled observations $\mathbf{x} \in \mathbb{R}^d$ are presented to the network which estimates their class membership ω .

Here, we restrict ourselves to Gaussian basis functions [54] of the type

$$\phi_j(\mathbf{x}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{c}_j\|^2}{2\sigma_j^2}\right)$$

where \mathbf{x} is the d -dimensional input vector with elements $x_i \in \mathbb{R}$, and $\mathbf{c}_j \in \mathbb{R}^d$ is the vector determining the center of the basis function ϕ_j and has elements $c_{ji} \in \mathbb{R}$, $\|\cdot\|$ denotes the Euclidean norm. The radial basis function neural network mapping with M basis functions is then

$$y_k(\mathbf{x}) = \sum_{j=1}^M w_{kj} \phi_j(\mathbf{x}) + w_{k0}$$

where the w_{k0} denote the biases, which may be absorbed into the summation by including an extra basis function ϕ_0 whose activation is set equal to 1. This mapping can be represented as the network diagram of Figure 2 with the radial basis functions in the hidden layer and linear summation on the output layer. In our classification scenario the number of output units corresponds to the number of classes (1 of l coding). Categorization is performed by assigning the input vector \mathbf{x} the class of the output unit with maximum activation:

$$\text{class}(\mathbf{x}) = \omega_{k^*} \quad \text{where} \quad k^* = \underset{k=1, \dots, l}{\text{argmax}} y_k(\mathbf{x}).$$

To simplify notation and without any loss of generality we identify

$$\omega_k = k \quad \text{and thus} \quad Y = \{1, \dots, l\}.$$

Typically, training an RBF network is separated into two phases:

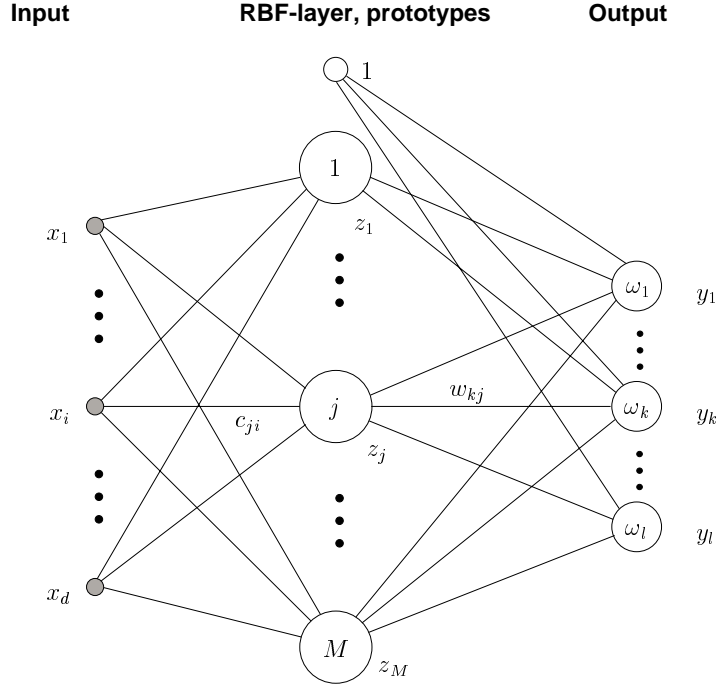


Figure 2. Architecture of the radial basis function neural network with Gaussian basis functions $z_j = \phi_j(\mathbf{x}) = \exp\left(-\frac{\|\mathbf{x}-\mathbf{c}_j\|^2}{2\sigma_j^2}\right)$ and linear output units $y_k(\mathbf{x}) = \sum_{j=1}^M w_{kj}\phi_j(\mathbf{x}) + w_{k0}$.

- (a) Adaptation of the centers of the prototype with a supervised or unsupervised training procedure and
- (b) adaptation of the output layer together with the setting of the width of the RBF functions.

In the following methods for these two phases, which were applied in this investigation are described together with a procedure for the selection of initial prototypes.

3.1 Selection of seed prototypes

Different methods for the selection or generation of seed prototypes are possible. One of them, a random selection of prototypes from the avail-

able training data does not fulfill the requirements of good seed prototypes in terms of their equal distribution among the classes and good classification properties within a reasonable neighbourhood of their location. We therefore used an algorithm for finding seed prototypes from the training data which selects if possible an equal number of prototypes from each class with the additional requirement of having within the K nearest neighbours a majority of the same class. If that is not possible the neighbourhood requirement is dropped. If that is still not possible a data point of any class is used as a seed prototype.

More formally we define:

$$d_j^\mu = \|\mathbf{x}^\mu - \mathbf{c}_j\|$$

for $j \in \{1, \dots, M\}$ and $\mu \in \{1, \dots, N\}$. With the sequence μ_1, \dots, μ_M and

$$d_j^{\mu_1} \leq d_j^{\mu_2} \leq \dots \leq d_j^{\mu_M},$$

$$\mathcal{N}^K(\mathbf{c}_j) = \{\mathbf{x}^{\mu_1}, \dots, \mathbf{x}^{\mu_K}\}$$

is the set of K nearest neighbours (data points) of prototype \mathbf{c}_j and

$$\mathcal{N}_k^K(\mathbf{c}_j) = \{\mathbf{x}^\mu \in \mathcal{N}^K(\mathbf{c}_j) \mid \text{class}(\mathbf{x}^\mu) = k\}$$

is the set of nearest neighbours (data points) of class k among the K nearest neighbours (data points) of prototype \mathbf{c}_j .

The procedure for the selection of seed prototypes is described in Algorithm 1.

3.2 Adapting the prototype location

Supervised: OLVQ1. Kohonen [55] proposed a supervised version of a vector quantization algorithm called Learning Vector Quantization (LVQ). From the basic LVQ1 version, the OLVQ1, LVQ2, and LVQ3 competitive training procedures have been derived. In the following the OLVQ1 algorithm, which is used in the experiments, will be briefly illustrated. After the initialization of the prototypes \mathbf{c}_j and upon the presentation of a feature vector from the training set \mathbf{x}^μ , the location of the winning prototype \mathbf{c}_{j^*} with

$$j^* = \underset{j}{\operatorname{argmin}} \|\mathbf{x}^\mu - \mathbf{c}_j\|$$

Algorithm 1 Selection of seed prototypes

Require: K, M
 $m = 0, C = \emptyset$
while $m < M$ **do**
 for $i = 1$ to l **do**
 for $\mu \in \{1, \dots, N\} \setminus C$ **do**
 $\Delta_\mu = |\mathcal{N}_k^K(\mathbf{x}^\mu)| - \max_{i \neq k} |\mathcal{N}_i^K(\mathbf{x}^\mu)|$
 end for
 $\mu^* = \operatorname{argmax}_\mu \Delta_\mu$
 $C = C \cup \{\mu^*\}$
 $\mathbf{c}_m = \mathbf{x}^{\mu^*}, \operatorname{class}(\mathbf{c}_m) = k$
 $m = m + 1$
 end for
end while
if $m < M$ **then**
 select $M - m$ prototypes at random
end if

is adapted according to the learning rule:

$$\Delta \mathbf{c}_{j^*} = \begin{cases} +\eta_{j^*}(t)(\mathbf{x}^\mu - \mathbf{c}_{j^*}), & \text{if } \operatorname{class}(\mathbf{c}_{j^*}) = \operatorname{class}(\mathbf{x}^\mu) \\ -\eta_{j^*}(t)(\mathbf{x}^\mu - \mathbf{c}_{j^*}), & \text{if } \operatorname{class}(\mathbf{c}_{j^*}) \neq \operatorname{class}(\mathbf{x}^\mu) \end{cases} .$$

In contrast to the LVQ1 algorithm, OLVQ1 exhibits an individual learning rate $\eta_j(t)$ for each prototype \mathbf{c}_j . Usually $\eta_j(t)$ is positively decreasing. The class labels of the prototypes are not changed during adaptation, they remain constant after initialization.

Unsupervised: Batch k-means. It is possible to adapt the prototype location in a non-trivial way without using any class information of the training patterns. This is done by utilizing the neighbourhood relationships within the data points imposed by a distance measure. The k-means clustering procedure [56–58] is among the most popular methods in cluster analysis [59]. After initialization of the prototypes \mathbf{c}_j their location is adapted after every epoch according to (batch k-means algorithm):

$$\mathbf{c}_j = \frac{1}{|\mathcal{C}_j|} \sum_{\mathbf{x}^\mu \in \mathcal{C}_j} \mathbf{x}^\mu$$

with

$$\mathcal{C}_j = \left\{ \mathbf{x}^\mu \in X \mid j = \underset{i}{\operatorname{argmin}} \|\mathbf{x}^\mu - \mathbf{c}_i\| \right\}$$

being the set of hits at prototype \mathbf{c}_j and $X = \{\mathbf{x}^1, \dots, \mathbf{x}^N\}$ being the set of feature vectors of the training set S . It is easily shown that this update rule minimizes the error function:

$$E(\mathbf{c}_1, \dots, \mathbf{c}_M) = \sum_{j=1}^M \sum_{\mathbf{x}^\mu \in \mathcal{C}_j} \|\mathbf{x}^\mu - \mathbf{c}_j\|^2.$$

Dynamic LVQ. The previous two approaches for adapting the hidden layer require the a priori setting of the number of prototypes. Here, we present a simple data driven codebook generation scheme, which is similar to [60, 61]. Adaptation of the prototype location is done by OLVQ1 (although k-means could be used as well). During training a statistic for each prototype \mathbf{c}_j is maintained, which contains the number of class specific hits h_{jk} with $k \neq \operatorname{class}(\mathbf{c}_j)$. Based upon this statistic new prototypes are inserted, and if after the temporarily insertion of the new prototype the classification error does not decrease, the prototype inserted last is pruned from the network, see Algorithm 2 and Figure 3.

For $j \in \{1, \dots, M\}$ and $k \in \{1, \dots, l\}$ we define

$$\mathcal{C}_{jk} = \{ \mathbf{x}^\mu \in \mathcal{C}_j \mid \operatorname{class}(\mathbf{x}^\mu) = k \}$$

and for $k \in \{1, \dots, l\} \setminus \operatorname{class}(\mathbf{c}_j)$ we set

$$h_{jk} = |\mathcal{C}_{jk}|.$$

Then, there exists a sequence $(j_1, k_1), \dots, (j_p, k_p)$ with $p = M(l - 1)$ such that

$$h_{j_1 k_1} \geq h_{j_2 k_2} \geq \dots \geq h_{j_p k_p}.$$

Candidate prototypes for the temporary insertion into the codebook are defined as follows:

$$\mathbf{c}_{j_\nu}^\nu = \frac{1}{h_{j_\nu k_\nu}} \sum_{\mathbf{x} \in \mathcal{C}_{j_\nu k_\nu}} \mathbf{x} \quad \text{and} \quad \operatorname{class}(\mathbf{c}_{j_\nu}^\nu) = k_\nu.$$

Algorithm 2 Dynamic LVQ

Require: R, ϵ, MAX

OLVQ1 training: $\mathbf{c}_1, \dots, \mathbf{c}_M$

$E = \text{error}(\mathbf{c}_1, \dots, \mathbf{c}_M)$

while $(E \leq \epsilon)$ or $(M \geq MAX - 1)$ **do**

 Calculate all h_{jk} for $j \in \{1, \dots, M\}$ and $k \in \{1, \dots, l\} \setminus \text{class}(\mathbf{c}_j)$

 Sort h_{jk} such that $h_{j_1 k_1} \geq h_{j_2 k_2} \geq \dots \geq h_{j_p k_p}$

$\nu = 1$

while $\nu \leq p$ **do**

$\mathbf{c}_{j_\nu}^\nu = \frac{1}{h_{j_\nu k_\nu}} \sum_{\mathbf{x} \in \mathcal{C}_{j_\nu k_\nu}} \mathbf{x}$

$\text{class}(\mathbf{c}_{j_\nu}^\nu) = k_\nu$

 OLVQ1 ($\mathbf{c}_1, \dots, \mathbf{c}_M, \mathbf{c}_{j_\nu}^\nu$)

if $E > (e = \text{error}(\mathbf{c}_1, \dots, \mathbf{c}_M, \mathbf{c}_{j_\nu}^\nu))$ **then**

$M = M + 1, E = e$

$\mathbf{c}_M = \mathbf{c}_{j_\nu}^\nu, \text{class}(\mathbf{c}_M) = k_\nu$

else

$\nu = \nu + 1$

end if

end while

OLVQ1 training: $\mathbf{c}_1, \dots, \mathbf{c}_M$

$E = \text{error}(\mathbf{c}_1, \dots, \mathbf{c}_M)$

end while

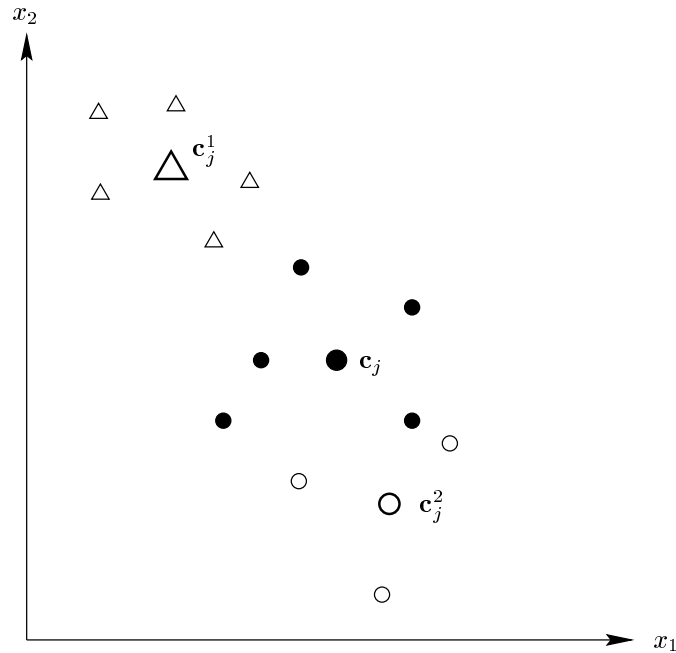


Figure 3. Illustration of prototype insertion with dynamic LVQ in a three class scenario (Δ , \circ , \bullet). The prototype c_j makes five misclassifications with class Δ and three with class \circ . At first a new prototype is added at the center of gravity c_j^1 . If the overall classification rate does not rise, location c_j^2 is tried.

3.3 Construction of the RBF network

So far we have only dealt with different vector quantization schemes for the adaptation of prototype locations. For the transition to a RBF network these prototype locations are used as centers of the radial basis functions. Additionally the initial setting and possible adaptation of the kernel widths σ_j of the radial basis functions and of the hidden to output layer connections w_{kj} is required. Here, we restrict ourselves to initializing the output weights to 1 for 'prototype-of-that-class' to 'output-of-that-class' connections and the others to small values (see Section 5).

3.3.1 Setting of the kernel widths.

The setting of the kernel widths is a critical issue in the transition to the RBF network. When the kernel width σ is too large the estimated probability density is over-smoothed and the nature of the underlying true density may be lost. Conversely, when σ is too small there may be an over-adaptation to the particular data set. In addition very small σ tend to cause numerical problems with gradient descent methods as their gradients vanish. We investigated three different schemes for the initial setting of the kernel widths in transition to the RBF network (γ is set heuristically):

1. The kernel width σ_j is set to the mean of the distance to the L nearest prototypes $\mathcal{P}_j(L)$ of prototype \mathbf{c}_j :

$$\sigma_j = \gamma \frac{1}{L} \sum_{\mathbf{c} \in \mathcal{P}_j(L)} \|\mathbf{c} - \mathbf{c}_j\|$$

2. Use the distance to the nearest prototype with a different class label for initialization:

$$\sigma_j = \gamma \min_{\substack{i \neq j \\ \text{class}(\mathbf{c}_i) \neq \text{class}(\mathbf{c}_j)}} \|\mathbf{c}_i - \mathbf{c}_j\|$$

3. All σ_j are set to the same value, which is proportional to the average minimal distance between all prototypes:

$$\sigma_j = \sigma = \gamma \frac{1}{M} \sum_{n=1}^M \min_{i \neq n} \|\mathbf{c}_i - \mathbf{c}_n\|$$

3.3.2 Gradient descent

The adaptation of the output weights and of the kernel widths was done with two gradient descent methods, i.e. plain back-propagation and back-propagation enhanced by Armijo line search.

Back-propagation. We give a brief summary of the use of error-back-propagation in the context of radial basis function network training, for a more detailed treatment see [54, 62, 63].

If we define as the error function of the network a differentiable function like the sum-of-squares error E ,

$$E = \frac{1}{2} \sum_{\mu=1}^N \sum_{k=1}^l (y_k^\mu - t_k^\mu)^2$$

with y_k^μ and t_k^μ as the actual and target output values respectively, and we consider a network with differentiable activation functions – which we have seen in Figure 2 – then a necessary condition for a minimal error is that its derivatives with respect to the parameters center location \mathbf{c} , kernel width σ and output weights \mathbf{w} vanish. The resulting equations are only solvable explicitly if the activation functions are linear and with a sum-of-squares error function. An iterative procedure for finding a solution to this problem is gradient descent. Here, the full parameter set $\mathbf{W} = (\mathbf{c}_j, \sigma_j, \mathbf{w}_j)$ is moved by a small distance η (the learning rate) – in their respective spaces – in the direction in which E decreases most rapidly, i.e. in the direction of the negative gradient $-\nabla E$:

$$\mathbf{W}^{(\tau+1)} = \mathbf{W}^{(\tau)} - \eta \nabla E(\mathbf{W}^{(\tau)}).$$

The update of the other parameters (\mathbf{c} , σ) is performed in a similar fashion. For the network of Figure 2 we obtain the following expressions for the derivatives of the error function with respect to the network parameters:

$$\begin{aligned} \frac{\partial E}{\partial \sigma_j} &= \sum_{\mu} \sum_k (y_k^\mu - t_k^\mu) w_{kj} \exp\left(-\frac{\|\mathbf{x}^\mu - \mathbf{c}_j\|^2}{2\sigma_j^2}\right) \frac{\|\mathbf{x}^\mu - \mathbf{c}_j\|^2}{\sigma_j^3} \\ \frac{\partial E}{\partial c_{ji}} &= \sum_{\mu} \sum_k (y_k^\mu - t_k^\mu) w_{kj} \exp\left(-\frac{\|\mathbf{x}^\mu - \mathbf{c}_j\|^2}{2\sigma_j^2}\right) \frac{x_i^\mu - c_{ji}}{\sigma_j^2} \\ \frac{\partial E}{\partial w_{kj}} &= \sum_{\mu} (y_k^\mu - t_k^\mu) z_j^\mu. \end{aligned}$$

Back-propagation with variable stepsize. Choosing the right learning rate or stepsize η is sometimes a critical issue in neural network training. If its value is too low convergence to a minimum is slow, conversely if it is chosen too high successive steps in parameter space overshoot the minimum of the error surface. This problem can be avoided by a proper

stepsize tuning. A procedure for obtaining such a stepsize was proposed by Armijo [64] in 1966. In the following very brief description of the method we draw heavily from the papers of Armijo [64] and Magoulas et al. [65], for details see the respective articles. Under mild conditions on the error function E , which are satisfied in our setting the following theorem holds:

Theorem (Armijo, 1966) If η_0 is an arbitrarily assigned positive number, $\eta_m = \frac{\eta_0}{2^{m-1}}$, $m = 0, 1, \dots$. Then the sequence of weight vectors $\{\mathbf{W}^{(\tau)}\}_0^\infty$ with

$$\mathbf{W}^{(\tau+1)} = \mathbf{W}^{(\tau)} - \eta_{m_\tau} \nabla E(\mathbf{W}^{(\tau)}), \quad \tau = 0, 1, 2, \dots$$

where m_τ is the smallest positive integer for which

$$E(\mathbf{W}^{(\tau)} - \eta_{m_\tau} \nabla E(\mathbf{W}^{(\tau)})) - E(\mathbf{W}^{(\tau)}) \leq -\frac{1}{2} \eta_{m_\tau} \|\nabla E(\mathbf{W}^{(\tau)})\|^2,$$

converges to the point \mathbf{W}^* which minimizes (locally) E .

Using Armijo's theorem Magoulas et al. [65] proposed a back-propagation algorithm with variable stepsize, see Algorithm 3.

Algorithm 3 Back-propagation with variable stepsize

Require: $E_{min}, \eta_{min}, \tau_{max}$

Epochs: $\tau = 0$

while $E(\mathbf{W}^{(\tau)}) > E_{min}$ & $\tau \leq \tau_{max}$ **do**

if $\tau = 0$ **then**

$\eta = 0.5$

else

$\eta = 0.5 / (\|\nabla E(\mathbf{W}^{(\tau)}) - \nabla E(\mathbf{W}^{(\tau-1)})\| / \|\mathbf{W}^{(\tau)} - \mathbf{W}^{(\tau-1)}\|)$

end if

while $\eta < \eta_{min}$ **do**

$\eta = 2\eta$

end while

while $E(\mathbf{W}^{(\tau)} - \eta \nabla E(\mathbf{W}^{(\tau)})) - E(\mathbf{W}^{(\tau)}) > -\frac{1}{2} \eta \|\nabla E(\mathbf{W}^{(\tau)})\|^2$ **do**

$\eta = \eta/2$

end while

$\mathbf{W}^{(\tau+1)} = \mathbf{W}^{(\tau)} - \eta \nabla E(\mathbf{W}^{(\tau)})$

$\tau = \tau + 1$

end while

4 Data

Subject groups. We compared a group of 51 healthy subjects (group A) with 44 cardiac patients at a high risk for malignant ventricular arrhythmias (group B, VT patients). All healthy volunteers (mean age 24.0 ± 4.1 years) had a normal resting ECG and a normal echocardiogram, and no cardiac symptoms or coronary risk factors. The patients with a high-risk for malignant ventricular arrhythmias (mean age 61.2 ± 8.9 years) were selected from our electrophysiologic database. Inclusion criteria were the presence of coronary artery disease, a previous myocardial infarction, a history of at least one symptomatic arrhythmia, and inducible sustained ventricular tachycardia (> 30 seconds) at electrophysiologic testing. Patients with bundle branch block or atrial fibrillation were excluded. All patients of group B underwent coronary angiography and programmed right ventricular stimulation due to clinical indications. Stimulation was done from the right apex and the right outflow tract. The stimulation protocol included up to 3 extrastimuli during sinus rhythm and at baseline pacing with a cycle length of 500 ms, and a maximum of 2 extrastimuli at baseline pacing with cycle lengths of 430 ms, 370 ms, and 330 ms. Group B consisted of 10 patients with single vessel disease, 17 patients with double vessel disease, and 17 patients with triple vessel coronary artery disease. Nineteen patients had a previous posterior infarction, 14 patients had a previous anterior infarction, and 11 patients had both a previous anterior and a previous posterior infarction. Mean left ventricular ejection fraction was $44.0\% \pm 14.9\%$. Forty-one patients had a documented episode of spontaneous, sustained ventricular tachycardia or ventricular fibrillation. Out of the remaining three patients, 1 patient had syncope and non-sustained ventricular tachycardias on Holter monitoring, and 2 patients had syncope of presumed cardiac origin.

Signal-Averaged ECG recordings. Ventricular late potential analysis (VLP) is a non-invasive method to identify patients with an increased risk for reentrant ventricular tachycardias and for risk stratification after myocardial infarction [66–68]. Techniques commonly applied in this purely time-domain based analysis are signal-averaging, high-pass filtering and late potential analysis of the terminal part of the QRS complex. The assessment of VLP's depends on three empirically defined limits of the

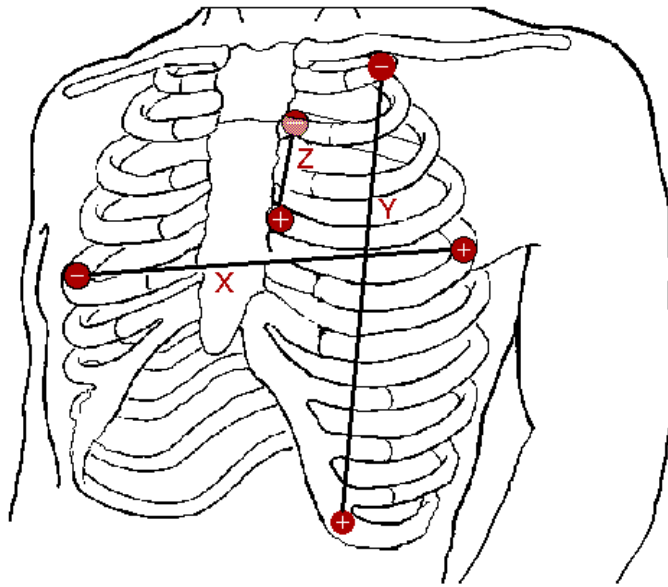


Figure 4. Placement of the three bipolar leads for the recording of high-resolution ECG's

total duration of the QRS and the duration and amplitude of the terminal low-amplitude portion of the QRS complex [69, 70].

High-resolution signal averaged electrocardiograms are recorded during sinus rhythm from three bipolar orthogonal X , Y , Z leads, see Figure 4. Before ECG recording antiarrhythmic drugs were stopped for at least four half-lives. The skin was carefully prepared and recordings were done with the subjects in reclining position in a Faraday cage, see Figure 5.

Sampling rate was 2000 Hz, A/D resolution was 16 bit, and an analog bandpass filter of 0.05-300 Hz was used (anti-aliasing). The ECG's were recorded with the Predictor system (Corasonix Inc., Oklahoma, USA). The three leads were averaged and combined into a vector magnitude signal $V = \sqrt{X^2 + Y^2 + Z^2}$ and bidirectionally filtered with a 4-pole Butterworth filter (40-250 Hz), see Figure 6.

From this vector magnitude signal V three features are extracted:

- QRSd (QRS duration):

$$QRSd := QRS_{offset} - QRS_{onset}$$



Figure 5. Picture of the high resolution ECG recording facilities. The subject is lying inside the Faraday cage.

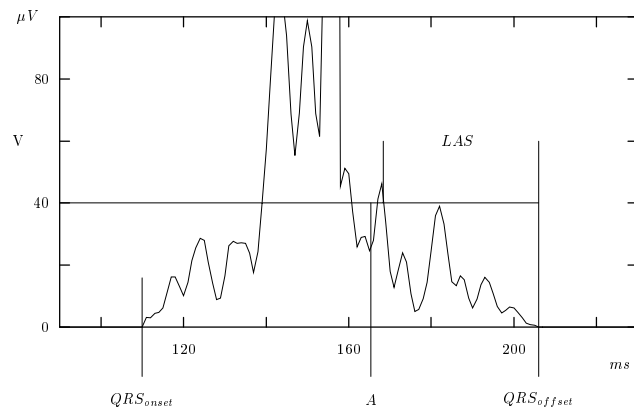


Figure 6. Signal averaged ECG: Example of the vector magnitude signal V of a patient with late potentials.

- RMS (Time A: $A := QRS_{offset} - 40ms$):

$$RMS := \sqrt{\frac{1}{QRS_{offset} - A} \sum_{i=A}^{QRS_{offset}} V_i^2}$$

- LAS (Duration of the low amplitude signal below $40\mu V$):

$$LAS := QRS_{offset} - \operatorname{argmax}\{i \mid V_i \geq 40\mu V\}$$

In standard late potential analysis a subject is termed “VLP positive” if 2 of the 3 following criteria are met: $QRS_{offset} > 115ms$, $RMS < 20\mu V$, $LAS > 38ms$, see Table 4. In this investigation these three features are used as inputs to a classifying RBF network, which is trained to predict the group status, see subject groups. Figure 7 (left) shows a 2-dimensional visualization of the complete dataset.

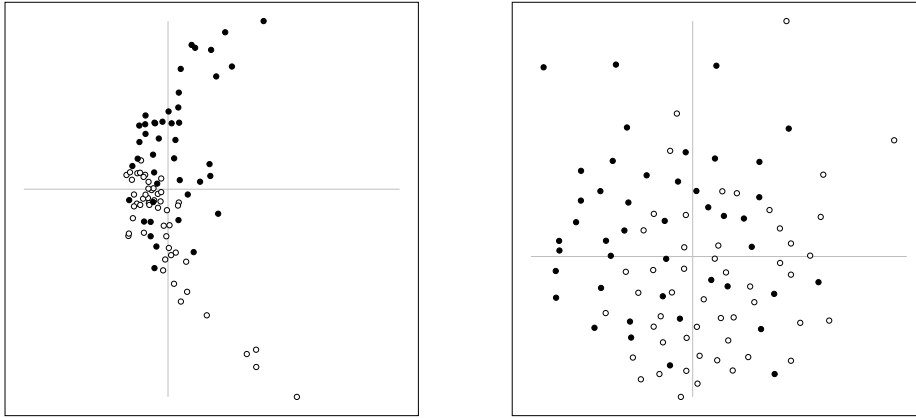


Figure 7. Mapping (Sammon mapping [71]) of the signal averaged ECG features (3D, left) and the beat-to-beat QRS variability features (14D, right) onto the 2 dimensional plane visualizing distance relations (using LVQ-Pak 3.1 Helsinki University of Technology). Samples from the healthy subject group are marked with \bullet those from the VT patient group with \circ .

Beat-to-beat ECG recordings. High-resolution beat-to-beat electrocardiograms of 30 min duration were recorded during sinus rhythm from bipolar orthogonal X , Y , Z leads using the same equipment as with

the signal-averaged recordings. Sampling rate was reduced to 1000 Hz. QRS triggering, reviewing of the ECG, and arrhythmia detection was done on a high-resolution ECG analysis platform developed by our group [72]. The three leads were summed into a signal $V = X + Y + Z$. From each recording 250 consecutive sinus beats preceded by another sinus beat were selected for subsequent beat-to-beat variability analysis. In a first step the signals were aligned by maximizing the cross-correlation function [73] between the first and all following beats. Prior to the quantification of signal variability the beats were pre-processed to suppress the main ECG waveform, bringing the beat-to-beat micro-variations into clearer focus. To achieve this, the individual signal was subtracted from its cubic spline smoothed version (spline filtering, spline interpolation through every seventh sample using the not-a-knot end condition) [74, 75], compare Figure 8. This method resembles a waveform adaptive, high-pass filtering without inducing phase-shift related artefacts. Next, for each individual beat the amplitude of the difference signal was normalized to zero mean and a standard deviation of $1 \mu\text{V}$. Beat-to-beat variation of each point was measured as the standard deviation of the amplitude of corresponding points across all 250 beats. For the QRS we used a constant analysis window of 141 ms which covered all QRS complexes of this series [17].

The resulting 141 dimensional variability vector was used as input for classification into subject group A or B. Figure 7 (right) shows a 2-dimensional visualization of the complete dataset.

5 Results

Different types of classification experiments with the Gaussian basis function networks of Section 3 were performed. The selection of seed prototypes was done with Algorithm 1 ($K = 5$).

The following taxonomy gives an overview of the experiments and the notation:

1. Adaptation of the prototype layer (see Section 3.2). Figures 9 and 10, and Tables 5 and 6 give the nearest neighbour classification results on the prototype layer.

OLVQ1 : Initial and maximal learning rate was set to 0.3.

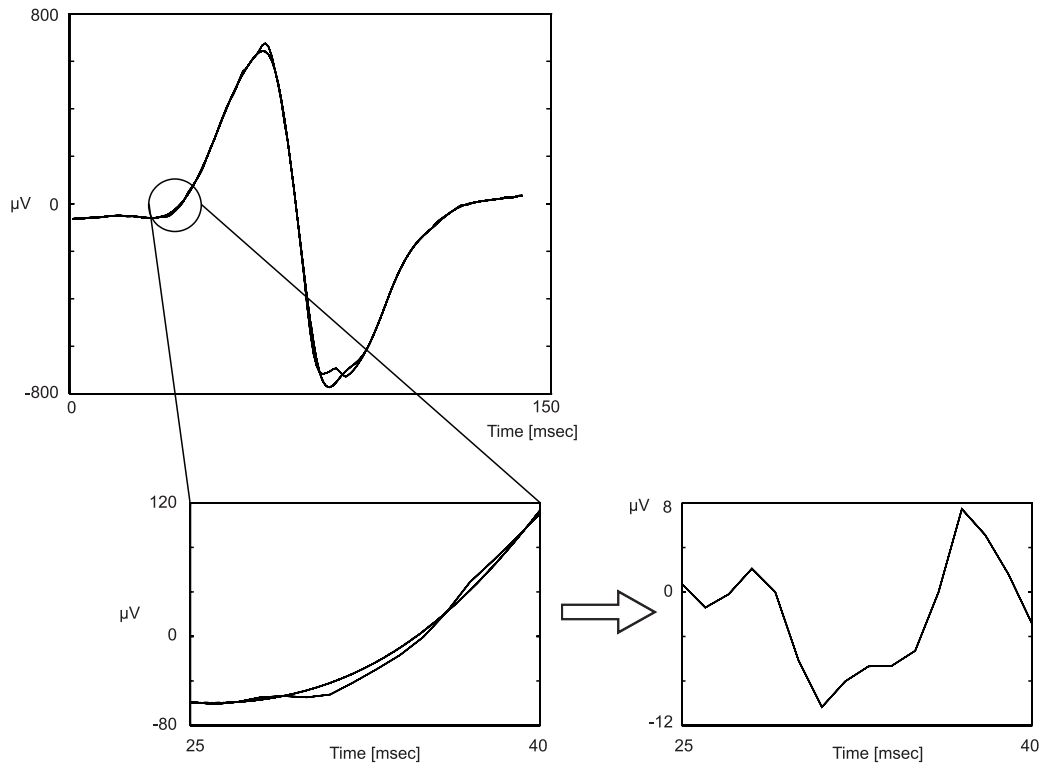


Figure 8. Diagram of the spline-filtering procedure. The upper panel shows both signals, the QRS-complex (sum of the three leads) and the cubic spline. A zoom-in makes the differences more apparent (lower left panel). The resulting signal difference is shown on the lower right panel (note the different scaling of the Y-axis).

Number of adaptation steps: $120M$.

Batch k-means : Training until the assignment of data points to prototypes does not change with each epoch. Class labels are not used in this adaptation step. After training class labels are assigned to prototypes by majority vote of those class labels of data points having minimal distance to the prototype.

Dynamic LVQ : List size p (see Algorithm 2) was set to 12. After every inserted prototype $20M$ OLVQ1 iterations were performed. Learning rate was set to 0.3. Following the growth process the prototypes were finally adapted with $120M$ OLVQ1 iterations. The initial number of prototypes was 2 in all cases and was bounded by the preset value. In some cases the preset number of prototypes was not attained. This occurred for $M \geq 44$ on the 3D data set and $M \geq 26$ on the 141D data set.

2. Training of the RBF network (see Section 3.3). Figures 11 and 12 give the classification results for the RBF networks with a re-training of the output layer only (online back-propagation). Kernel widths were set with the three methods given below. Figures 13 and 14, and Tables 7 and 8 show the classification results on the completely re-trained RBF network (σ initialization was done with method s_a) with Algorithm 3.

(a) Initialization of RBF widths (see Section 3.3.1)

γ was heuristically set to 2 in all cases.

s_a : σ_j were set to the average distance to the nearest three prototypes.

s_b : σ_j were set with the distance to the nearest prototype of a different class.

s_c : All σ_j were set to the same value, proportional to the average minimal distance between all prototypes.

(b) Initialization of the output weights w_{kj} : Output weights were always initialized to 1 for connections between prototypes of a class and their corresponding output unit, otherwise they were set to small random values (uniform distribution) in the range of $[-1E-6, +1E-6]$.

(c) Gradient descent (see Section 3.3.2)

Back-propagation on the output layer: Online back-propagation training, i.e. weight update after every presentation of a training pattern (learning rate $\eta = 0.04/\mu$ for every epoch μ). Every epoch consisted of $N = 95$ random presentations of a training pattern. A total of 100 training epochs was performed.

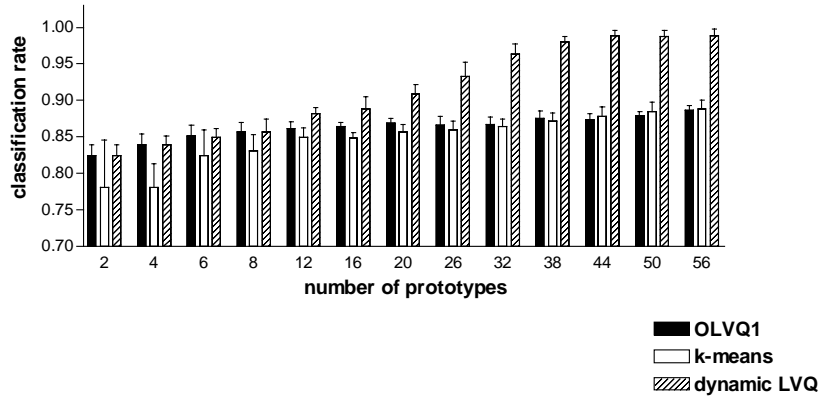
Re-training of the complete network: Complete re-training of the network with back-propagation with variable stepsize (Algorithm 3) for 80 epochs. Every epoch consists of a complete presentation of the training data.

The classification performance is given in terms of **re-validation** and **10-fold cross-validation** results. Re-validation means training and test on the whole data set. 10-fold cross-validation means partitioning the whole data set into 10 disjoint subsets and carrying out 10 training and test runs always using 9 subsets as the training set and testing on the remaining one. The results are those on the test sets. Each of these re-validation or 10-fold cross-validation simulations was performed 10 times. The difference between subsequent simulations was the initialization of the random number generator and the random permutation of the data set. The classification results of the different networks are summarized into five measures of performance.

		Actual Condition of Population	
		Patients with Disease	Patients without Disease
Classification Result	Positive	a (true-positives)	b (false-positives)
	Negative	c (false-negatives)	d (true-negatives)

$$\begin{aligned}
 \text{Accuracy (Acc)} &= (a + d) / (a + b + c + d) \\
 \text{Sensitivity (Sensi)} &= a / (a + c) \\
 \text{Specificity (Speci)} &= d / (b + d) \\
 \text{Positive predictive value (PPV)} &= a / (a + b) \\
 \text{Negative predictive value (NPV)} &= d / (c + d)
 \end{aligned}$$

**Nearest neighbour classification (prototype layer)
re-validation, 3D data**



**Nearest neighbour classification (prototype layer)
10-fold cross-validation, 3D data**

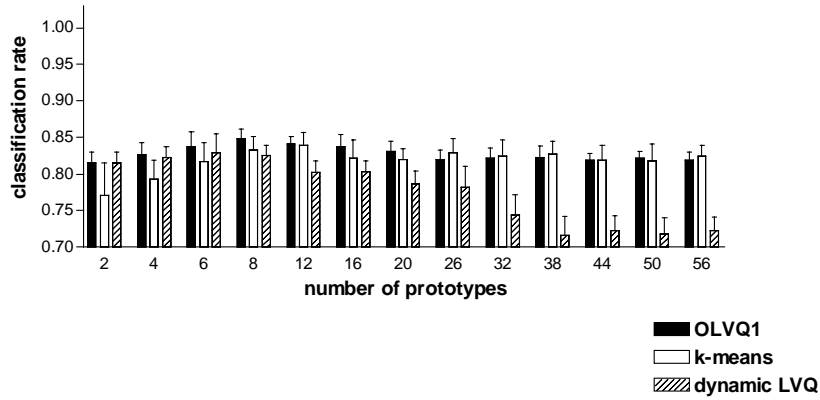
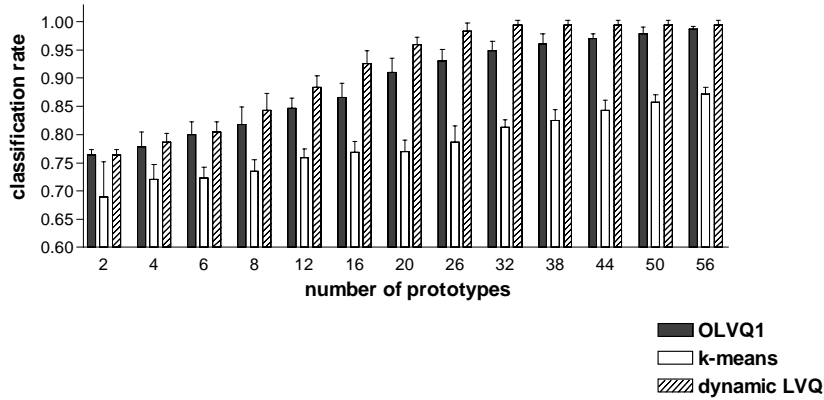


Figure 9. Signal-averaged 3D data: Re-classification (training and test set are the same) and 10-fold cross-validation results (accuracy) for the three prototype adaptation schemes (OLVQ1, k-means, dynamic LVQ) using the nearest neighbour rule on the hidden layer. Results are averages over 10 training and test runs for re-classification and over ten 10-fold cross-validation runs with their standard deviations.

**Nearest neighbour classification (prototype layer)
re-validation, 141D data**



**Nearest neighbour classification (prototype layer)
10-fold cross-validation, 141D data**

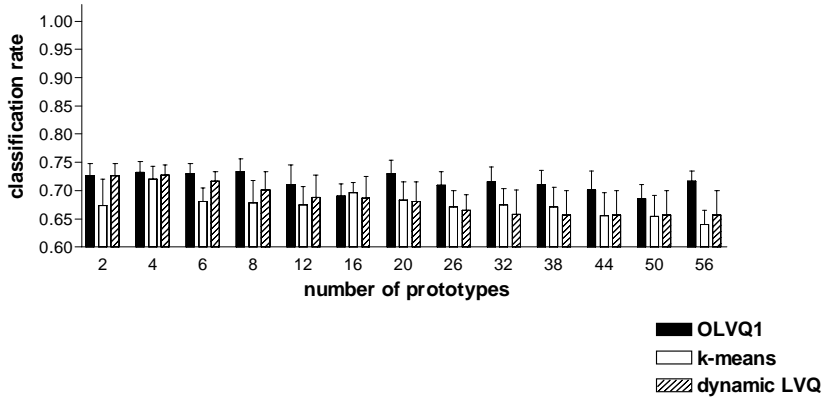


Figure 10. Beat-to-beat variability data: Re-classification (training and test set are the same) and 10-fold cross-validation results (accuracy) for the three prototype adaptation schemes (OLVQ1, k-means, dynamic LVQ) using the nearest neighbour rule on the hidden layer. Results are averages over 10 training and test runs for re-classification and over ten 10-fold cross-validation runs with their standard deviations.

Classification results on the prototype layer of the signal-averaged data (8 prototypes)

OLVQ1

	Acc	Sensi	Speci	PPV	NPV
Re-val	85.7% ± 1.3%	70% ± 3.2%	99.2% ± 1%	98.8% ± 1.6%	79.3% ± 1.6%
Cross-val	84.8% ± 1.3%	69.3% ± 2.9%	98.2% ± 1.4%	97.2% ± 2.3%	78.8% ± 1.5%

k-means

	Acc	Sensi	Speci	PPV	NPV
Re-val	83.1% ± 2.3%	67.7% ± 6.8%	96.3% ± 2.2%	94.3% ± 3.2%	77.7% ± 3.2%
Cross-val	83.3% ± 1.9%	67.7% ± 3.2%	96.7% ± 1.6%	94.6% ± 2.6%	77.7% ± 1.8%

dynamic LVQ

	Acc	Sensi	Speci	PPV	NPV
Re-val	85.7% ± 1.7%	71.4% ± 3.9%	98% ± 1.6%	97% ± 2.5%	79.9% ± 2.1%
Cross-val	82.5% ± 1.4%	69.8% ± 2.2%	93.5% ± 1.9%	90.4% ± 2.6%	78.2% ± 1.3%

Table 5. Detailed classification results of the three prototype adaptation schemes for the signal averaged data (3 dimensional, 8 prototypes): Re-classification and 10-fold cross-validation results are given (ten runs, mean ± standard deviation). Classification was performed on the prototype layer with the nearest neighbour rule.

Classification results on the prototype layer of the beat-to-beat variability data (8 prototypes)

OLVQ1

	Acc	Sensi	Speci	PPV	NPV
Re-val	81.8% ± 3.1%	71.8% ± 6.7%	90.4% ± 3.3%	86.7% ± 3.6%	79% ± 3.7%
Cross-val	73.3% ± 2.4%	62.7% ± 4.9%	82.4% ± 2.6%	75.4% ± 2.7%	72% ± 2.6%

k-means

	Acc	Sensi	Speci	PPV	NPV
Re-val	73.5% ± 2%	60.7% ± 8.9%	84.5% ± 6.8%	78.7% ± 8.6%	71.6% ± 3%
Cross-val	67.8% ± 4%	52% ± 4.5%	81.4% ± 5.6%	71% ± 6.8%	66.3% ± 2.8%

dynamic LVQ

	Acc	Sensi	Speci	PPV	NPV
Re-val	84.3% ± 3%	76.1% ± 5.2%	91.4% ± 3.5%	88.5% ± 4.2%	81.7% ± 3.4%
Cross-val	70.1% ± 3.3%	61.4% ± 2.6%	77.6% ± 5.8%	70.7% ± 6.3%	69.9% ± 1.9%

Table 6. Detailed classification results of the three prototype adaptation schemes for the beat-to-beat data (141 dimensional, 8 prototypes): Re-classification and 10-fold cross-validation results are given (ten runs, mean ± standard deviation). Classification was performed on the prototype layer with the nearest neighbour rule.

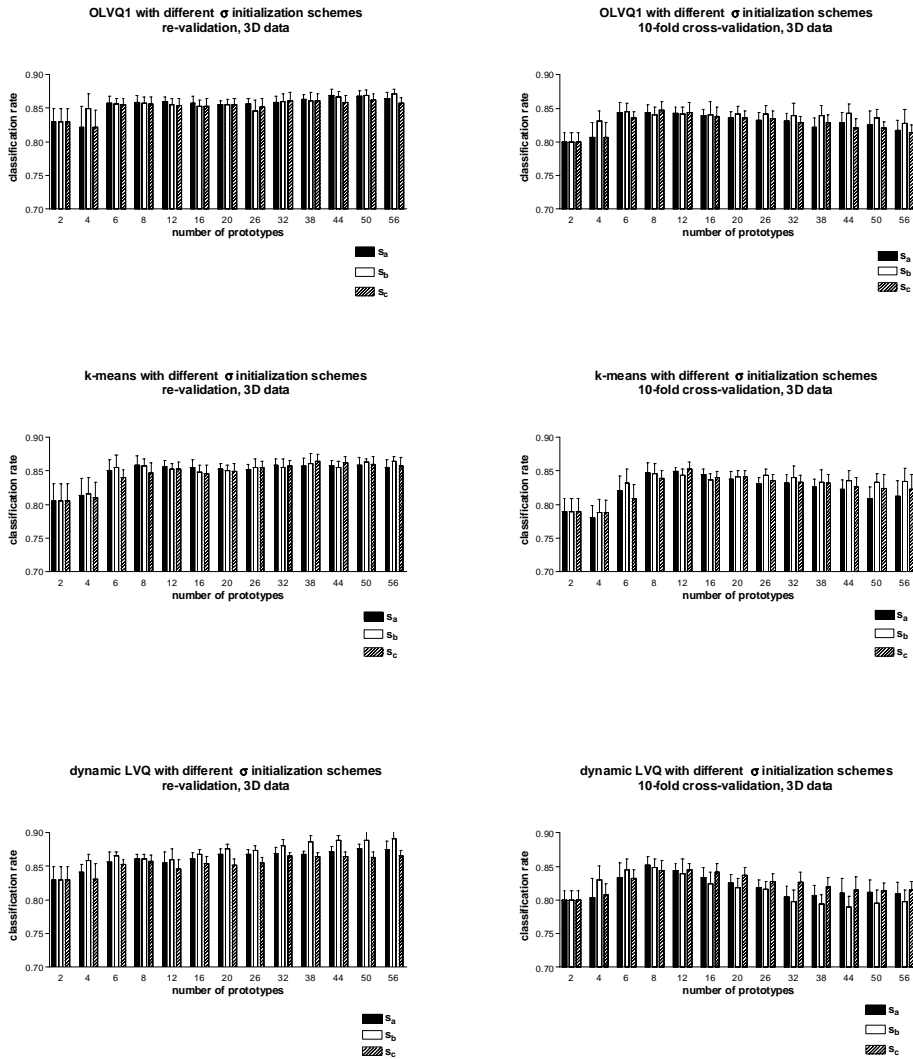


Figure 11. RBF results for the signal-averaged data. The figures show the results for the OLVQ1 (top row) prototype adaptation scheme, k-means (middle) and dynamic LVQ (bottom). Each figure gives the results for the three different σ initialization schemes (see text). Training of the output weights is done with back-propagation. Re-classification results are shown in the graphs on the left and cross-validation results in those on the right (accuracy, averages over ten runs and standard deviations are given).

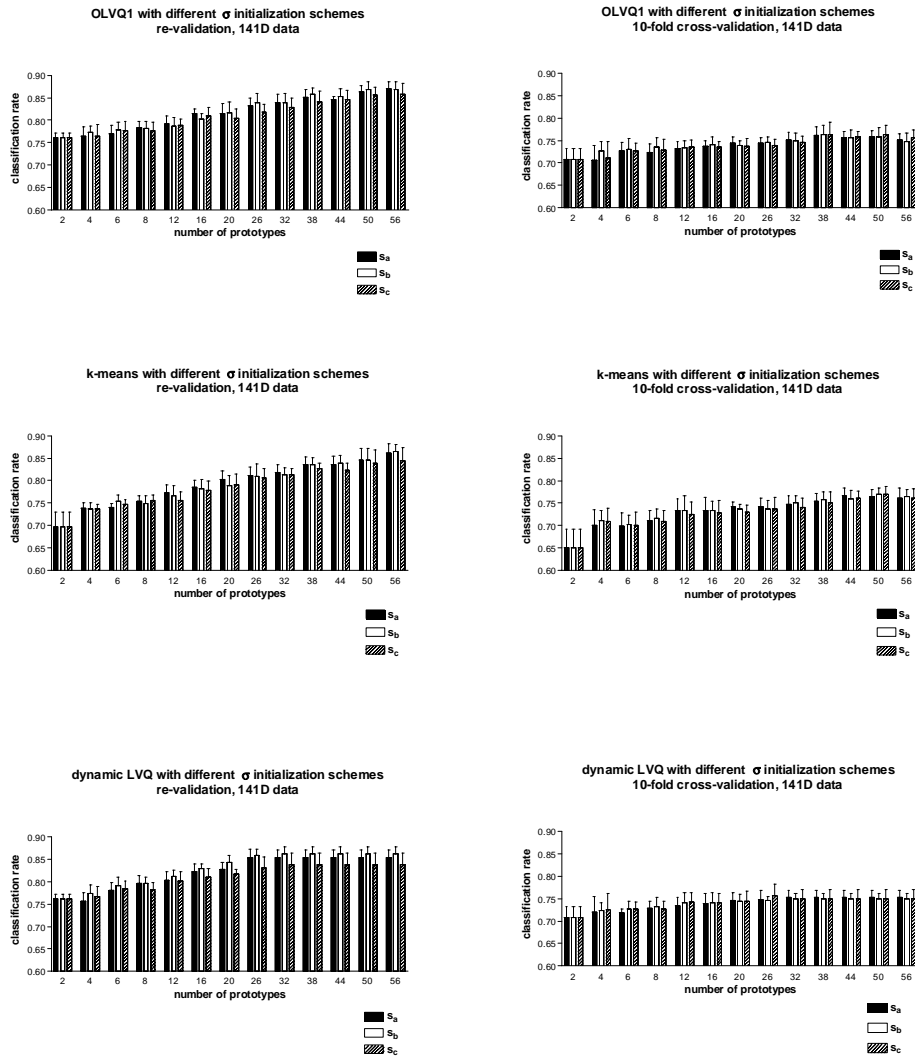


Figure 12. RBF results for the beat-to-beat variability data. The figures show the results for the OLVQ1 (top row) prototype adaptation scheme, k-means (middle) and dynamic LVQ (bottom). Each figure gives the results for the three different σ initialization schemes (see text). Training of the output weights is done with back-propagation. Re-classification results are shown in the graphs on the left and cross-validation results in those on the right (accuracy, averages over ten runs and standard deviations are given).

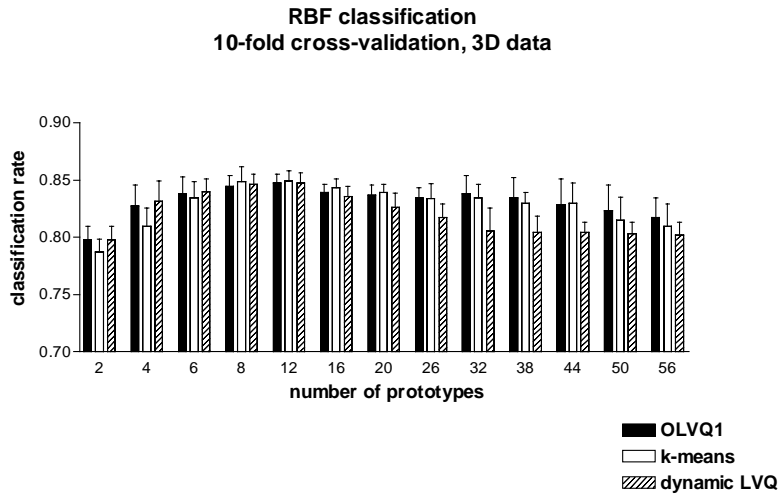
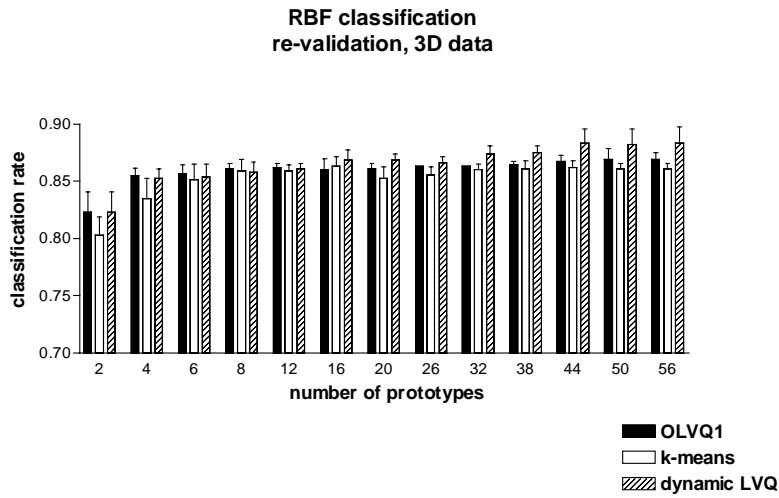
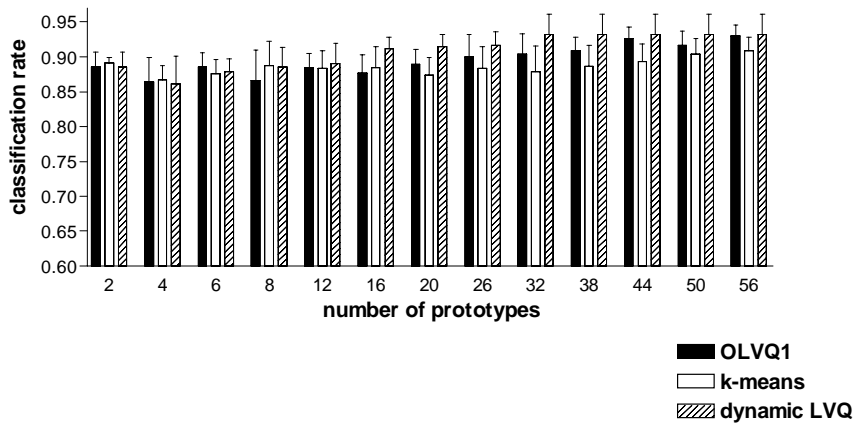


Figure 13. RBF results for the signal-averaged data. The graphs show the results for the OLVQ1 prototype adaptation scheme, k-means and dynamic LVQ for different numbers of hidden neurons. The widths of the RBF functions are initialized via their distance to the next three prototypes (see text). Output weights are initialized to 1 for connections between prototypes of one class and their corresponding output unit, otherwise small values. The complete network $(\sigma_j, \mathbf{c}_j, \mathbf{w}_k)$ is re-trained with back-propagation with variable stepsize. Re-classification results are shown in the figure on the left and cross-validation results on the right (accuracy, averages over ten runs and standard deviations are given)

**RBF classification
re-validation, 141D data**



**RBF classification
10-fold cross-validation, 141D data**

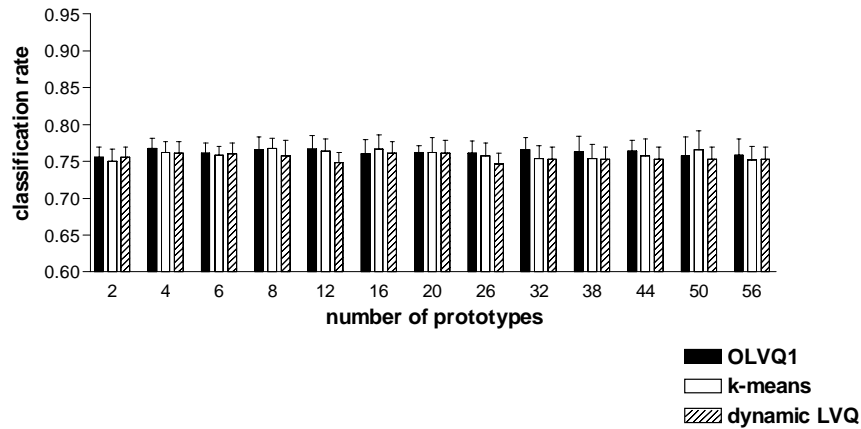


Figure 14. RBF results for the beat-to-beat variability data. The graphs show the results for the OLVQ1 prototype adaptation scheme, k-means and dynamic LVQ for different numbers of hidden neurons. The widths of the RBF functions are initialized via their distance to the next three prototypes (see text). Output weights are initialized to 1 for connections between prototypes of one class and their corresponding output unit, otherwise small values. The complete network (σ_j, c_j, w_k) is re-trained with back-propagation with variable stepsize. Re-classification results are shown in the figure on the left and cross-validation results on the right (accuracy, averages over ten runs and standard deviations are given)

RBF results for the signal-averaged data (8 hidden neurons)

OLVQ1					
	Acc	Sensi	Speci	PPV	NPV
Re-val	86.1% \pm 0.4%	70.9% \pm 1.8%	99.2% \pm 1.4%	98.8% \pm 2%	79.8% \pm 0.8%
Cross-val	84.4% \pm 1%	69.8% \pm 1.1%	97.1% \pm 1.9%	95.4% \pm 2.9%	78.8% \pm 0.6%

k-means					
	Acc	Sensi	Speci	PPV	NPV
Re-val	85.9% \pm 1%	70% \pm 1.4%	99.6% \pm 0.8%	99.4% \pm 1.4%	79.4% \pm 0.9%
Cross-val	84.8% \pm 1.3%	70.5% \pm 1.1%	97.3% \pm 2.1%	95.8% \pm 3.2%	79.2% \pm 0.8%

dynamic LVQ					
	Acc	Sensi	Speci	PPV	NPV
Re-val	85.8% \pm 0.9%	71.4% \pm 1.2%	98.2% \pm 1.4%	97.3% \pm 2.2%	79.9% \pm 0.7%
Cross-val	84.6% \pm 0.9%	70.2% \pm 0.7%	97.1% \pm 1.7%	95.4% \pm 2.4%	79.1% \pm 0.4%

Table 7. Detailed RBF classification results of the three prototype adaptation schemes for the signal averaged data (3 dimensional, 8 hidden neurons): Re-classification and 10-fold cross-validation results are given (ten runs, mean \pm standard deviation). The widths of the RBF functions are initialized via their distance to the next three prototypes (see text). The complete network ($\sigma_j, \mathbf{c}_j, \mathbf{w}_k$) is re-trained with back-propagation with variable stepsize.

RBF results for the beat-to-beat data (8 hidden neurons)

OLVQ1					
	Acc	Sensi	Speci	PPV	NPV
Re-val	86.6% \pm 4.3%	83.6% \pm 5.7%	89.2% \pm 3.7%	87% \pm 4.5%	86.4% \pm 4.5%
Cross-val	76.5% \pm 1.8%	69.8% \pm 3.2%	82.4% \pm 1.3%	77.3% \pm 1.7%	76% \pm 2%

k-means					
	Acc	Sensi	Speci	PPV	NPV
Re-val	88.7% \pm 3.5%	86.1% \pm 5.8%	91% \pm 3%	89.2% \pm 3.6%	88.5% \pm 4.2%
Cross-val	76.7% \pm 1.4%	70.7% \pm 2.5%	82% \pm 1.2%	77.2% \pm 1.5%	76.4% \pm 1.6%

dynamic LVQ					
	Acc	Sensi	Speci	PPV	NPV
Re-val	88.5% \pm 2.8%	85.9% \pm 4%	90.8% \pm 2.3%	88.9% \pm 2.8%	88.2% \pm 3.1%
Cross-val	75.8% \pm 2.1%	69.1% \pm 3.7%	81.6% \pm 1.4%	76.4% \pm 1.9%	75.4% \pm 2.4%

Table 8. Detailed RBF classification results of the three prototype adaptation schemes for the beat-to-beat data (141 dimensional, 8 hidden neurons): Re-classification and 10-fold cross-validation results are given (ten runs, mean \pm standard deviation). The widths of the RBF functions are initialized via their distance to the next three prototypes (see text). The complete network ($\sigma_j, \mathbf{c}_j, \mathbf{w}_k$) is re-trained with back-propagation with variable stepsize.

6 Concluding remarks

Several topics were touched in this investigation: The role of non-invasive risk assessment in cardiology, new signal processing techniques utilizing not only the three standard VLP parameters but processing sequences of beats, and the possible application of RBF networks in this assessment.

By using the more elaborate categorization methods of RBF networks compared to VLP assessment (see Section 4) on the 3 dimensional signal-averaged data an increase in accuracy of about 10% could be gained (VLP results: Acc = 72.6%, Sensi = 63.6%, Speci = 80.4%) in all cases of prototype based or RBF classification (see Tables 5 and 7 and Figures 9,11 and 13). The increase from prototype based to RBF network categorization was moderate but still visible and was accompanied by a reduction of variance. All network classification results show only a slight difference between re-validation and cross-validation for a small number of prototypes or radial basis functions. This substantiates the robustness of the methods on this data. Surprisingly does the accuracy stagnate at about 86% for the re-validation case on all simulations, only the dynamic LVQ makes an exception and shows an over-adaptation to the data set, which is reflected in its poor generalization ability at higher prototype numbers. Unfortunately the sensitivity of all methods on the 3D data is still too low to qualify as a single screening test (see Section 2, Current status). Another issue of using signal-averaged features solely as predictors of SCD are their moderate positive predictive value (see Table 4) and, as far as the simulations of this paper are concerned, a problematic prognostic significance of the PVS result for SCD (see Table 3). The positive and negative predictive values presented here should be treated with care as the study population does not represent the true prevalence of SCD.

In the case of the 141 dimensional best-to-beat variability data there is also a substantial (7% - 15%) increase in classification accuracy (see Tables 6 and 8 and Figures 10, 12 and 14) compared to categorization via a single cut-off value on the sum of the variability features (re-val: Acc = 73.7%, Sensi = 68.2%, Speci = 78.4% [17]; 10-fold cross-val (mean \pm stdev): Acc = 68.9% \pm 5%, Sensi = 66.1% \pm 8.7%, Speci = 71.4% \pm 16.8%). Compared to the 3D data classification results drop when switching to the RBF network and training the output layer only. Inter-

estingly, is the performance not influenced by the type of initialization of the kernel width in all cases on both data sets. Only on the high dimensional data a re-tuning of the complete network seems to be mandatory to recover from and even increase, the performance compared to a purely prototype based classification.

All networks end, after being re-trained, in a comparable range of performance (within their prototype number), although this is not the case on the prototype layer. The difference between the supervised training methods of OLVQ1 and dynamic LVQ and the unsupervised k-means is most pronounced with the variability data (Figure 10). This is not too surprising as the 2 dimensional projection (Figure 7) also points in that direction.

The presented results indicate that both types of features are of a supplementary nature (results on sensitivity). This calls for studies in which both are combined. Further investigations into the uncertainty of the target classification (long term follow-up) are needed to assess the diagnostic and prognostic value of the presented methods, their combination and their applicability to different patient groups.

Acknowledgement

An investigation like the above would not be possible without the support of many people, in particular we would like to thank André Müller, and especially Dr. Martin Höher and Professor Vinzenz Hombach for providing an inspiring research environment.

References

- [1] Buxton, A., K. Lee, J. Fisher, M. Josephson, E. Prystowsky, and G. Hafley (1999). A randomized study of the prevention of sudden death in patients with coronary artery disease. Multicenter Unsustained Tachycardia Trial Investigators. *N.Engl.J Med*, vol. 341(25), pp. 1882–1890.
- [2] Moss, A., W. Hall, D. Cannom, J. Daubert, S. Higgins, H. Klein, J. Levine, S. Saksena, A. Waldo, D. Wilber, M. Brown, and M. Heo (1996). Improved survival with an implanted defibrillator in patients with coronary disease at high risk for ventricular arrhythmia. *New Engl J Med*, vol. 335, pp. 1933–1940.

- [3] Moss, A. (1997). Update on MADIT: the multicenter autonomic defibrillator implantation trial. *Am J Cardiol*, vol. 79(6A), pp. 16–17.
- [4] Sherif, N. E., B. Scherlag, R. Lazzara, and R. Hope (1977). Re-entrant Ventricular Arrhythmias in the Late Myocardial Infarction Period. 1. Conduction Characteristics in the Zone. *Circulation*, vol. 55, pp. 686–702.
- [5] Sherif, N. E., R. Hope, B. Scherlag, and R. Lazzara (1977). Re-entrant Ventricular Arrhythmias in the Late Myocardial Infarction Period. 2. Patterns of Initiation and Termination of Re-entry. *Circulation*, vol. 55, pp. 702–719.
- [6] Ciaccio, E. (2000). Localization of the Slow Conduction Zone During Reentrant Ventricular Tachycardia. *Circulation*, vol. 102, pp. 464–469.
- [7] Antman, E. and E. Braunwald (1997). Acute myocardial infarction. In E. Braunwald, ed., *Heart disease - A textbook of cardiovascular medicine*, W.B. Saunders Company, Philadelphia-London-Toronto-Montreal-Sydney-Tokyo. pp. 1184–1288.
- [8] Lamas, G., G. Flaker, G. Mitchell, J. Smith, S.C., B. Gersh, C. Wun, L. Moye, J. Rouleau, J. Rutherford, and M. Pfeffer (1995). Effect of infarct artery patency on prognosis after acute myocardial infarction. The Survival and Ventricular Enlargement Investigators. *Circulation*, vol. 92(5), pp. 1101–1109.
- [9] Julian, D., A. Camm, G. Frangin, M. Janse, A. Munoz, P. Schwartz, and P. Simon (1997). Randomised trial of effect of amiodarone on mortality in patients with left-ventricular dysfunction after recent myocardial infarction: EMIAT. European Myocardial Infarct Amiodarone Trial Investigators [published errata appears in *Lancet* 1997 Apr 19;349(9059):1180 and 1997 Jun 14;349(9067):1776]. *Lancet*, vol. 349(9053), pp. 667–674.
- [10] Cairns, J., S. Connolly, R. Roberts, and M. Gent (1997). Randomised trial of outcome after myocardial infarction in patients with frequent or repetitive ventricular premature depolarisations: CAMIAT. Canadian Amiodarone Myocardial Infarction Arrhythmia Trial Investigators [published erratum appears in *Lancet* 1997 Jun 14;349(9067):1776] [see comments]. *Lancet*, vol. 349(9053), pp. 675–682.
- [11] The Acute Infarction Ramipril Efficacy (AIRE) Study Investigators (1993). Effect of ramipril on mortality and morbidity of survivors of acute myocardial infarction with clinical evidence of heart failure. *Lancet*, vol. 342(8875), pp. 821–828.

- [12] Amiodarone Trials Meta-Analysis Investigators (1997). Effect of prophylactic amiodarone on mortality after acute myocardial infarction and in congestive heart failure: meta-analysis of individual data from 6500 patients in randomized trials. *Lancet*, vol. 350(9089), pp. 1417–24.
- [13] Breithardt, G., M. Borggrefe, A. Martinez Rubio, and T. Budde (1989). Pathophysiological mechanisms of ventricular tachyarrhythmias. *Eur Heart J*, vol. Suppl.E, pp. 9–18.
- [14] Hombach, V., V. Braun, H. Hopp, D. Gil-Sanchez, H. Scholl, D. Behrenbeck, M. Tauchert, and H. Hilger (1982). The applicability of the signal averaging technique in clinical cardiology. *Clin.Cardiol.*, vol. 5, pp. 107–124.
- [15] Hombach, V., M. Hoehner, and M. Kochs (1993). Clinical significance of high resolution electrocardiography – sinus node, His bundle and ventricular late potentials. In J. Gomes, ed., *Signal averaged electrocardiography, concepts, methods and applications*, Kluwer Academic Publishers, Dordrecht-Boston-London. pp. 267–295.
- [16] Hoehner, M., J. Axmann, T. Eggeling, M. Kochs, P. Weismuller, and V. Hombach (1993). Beat-to-beat variability of ventricular late potentials in the unaveraged high resolution electrocardiogram - Effects of antiarrhythmic drugs. *Eur:Heart J*, vol. 14 (Suppl. E), pp. 33–39.
- [17] Kestler, H. A., J. Wöhrle, and M. Höher (2000). Cardiac vulnerability assessment from electrical microvariability of the high-resolution electrocardiogram. *Medical & Biological Engineering & Computing*, vol. 38, pp. 88–92.
- [18] Higham, P. and R. Campbell (1994). QT dispersion. *Br.Heart J*, vol. 71(6), pp. 508–510.
- [19] Algra, A., J. Tijssen, J. Roelandt, J. Pool, and J. Lubsen (1993). QT interval variables from 24 hour electrocardiography and the two year risk of sudden death. *Br.Heart J*, vol. 70(1), pp. 43–48.
- [20] Rosenbaum, D., L. Jackson, J. Smith, H. Garan, J. Ruskin, and R. Cohen (1994). Electrical alternans and vulnerability to ventricular arrhythmias. *New Engl J Med*, vol. 330, pp. 235–241.
- [21] Smith, J., E. Clancy, C. Valeri, J. Ruskin, and R. Cohen (1988). Electrical alternans and cardiac electrical instability. *Circulation*, vol. 77(1), pp. 110–121.

- [22] Kleiger, R., J. Miller, J. Bigger-JT, and A. Moss (1987). Decreased heart rate variability and its association with increased mortality after acute myocardial infarction. *Am J Cardiol*, vol. 59(4), pp. 256–262.
- [23] Malik, M., T. Farrell, T. Cripps, and A. Camm (1989). Heart rate variability in relation to prognosis after myocardial infarction: Selection of optimal processing techniques. *Eur.Heart J*, vol. 10, pp. 1060–1074.
- [24] Bigger Jr, J., J. Fleiss, R. Steinman, L. Rolnitzky, R. Kleiger, and J. Rottman (1992). Frequency domain measures of heart period variability and mortality after myocardial infarction. *Circulation*, vol. 85(1), pp. 164–171.
- [25] Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology (1996). Heart rate variability: standards of measurement, physiological interpretation and clinical use. [see comments]. *Circulation*, vol. 93(5), pp. 1043–1065.
- [26] Breithardt, G., L. Seipel, T. Meyer, and R. Abendroth (1982). Prognostic Significance of Repetitive Ventricular Response During Programmed Ventricular Stimulation. *Am J Cardiol.*, vol. 49, pp. 693–698.
- [27] Gonzalez, R., D. Arriagada, R. Corbalan, G. Chamorro, A. Fajuri, and J. Rodriguez (1988). Role of programmed electrical stimulation of the heart in risk stratification post-myocardial infarction. *Pacing Clin.Electrophysiol.*, vol. 11, pp. 283–288.
- [28] Hamer, A., J. Vohra, D. Hunt, and G. Sloman (1982). Prediction of sudden death by electrophysiologic studies in high risk patients surviving acute myocardial infarction. *Am J Cardiol*, vol. 50(2), pp. 223–229.
- [29] Richards, D., D. Cody, A. Denniss, P. Russell, A. Young, and J. Uther (1983). Ventricular electrical instability: a predictor of death after myocardial infarction. *Am J Cardiol*, vol. 51(1), pp. 75–80.
- [30] Marchlinski, F., A. Buxton, H. Waxman, and M. Josephson (1983). Identifying patients at risk of sudden death after myocardial infarction: value of the response to programmed stimulation, degree of ventricular ectopic activity and severity of left ventricular dysfunction. *Am J Cardiol*, vol. 52(10), pp. 1190–1196.
- [31] Waspe, L., D. Seinfeld, A. Ferrick, S. Kim, J. Matos, and J. Fisher (1985). Prediction of sudden death and spontaneous ventricular tachycardia in survivors of complicated myocardial infarction: value of the response to pro-

- grammed stimulation using a maximum of three ventricular extrastimuli. *J Am Coll Cardiol*, vol. 5(6), pp. 1292–1301.
- [32] Roy, D., E. Marchand, P. Theroux, D. Waters, G. Pelletier, and M. Bourassa (1985). Programmed ventricular stimulation in survivors of an acute myocardial infarction. *Circulation*, vol. 72(3), pp. 487–494.
- [33] Santarelli, P., F. Bellocchi, F. Loperfido, M. Mazzari, R. Mongiardo, A. Montenero, U. Manzoli, and P. Denes (1985). Ventricular arrhythmia induced by programmed ventricular stimulation after acute myocardial infarction. *Am J Cardiol*, vol. 55(4), pp. 391–394.
- [34] Bhandari, A., J. Rose, A. Kotlewski, S. Rahimtoola, and D. Wu (1985). Frequency and significance of induced sustained ventricular tachycardia or fibrillation two weeks after acute myocardial infarction. *Am J Cardiol*, vol. 56(12), pp. 737–742.
- [35] Kowey, P., H. Waxman, A. Greenspon, R. Greenberg, D. Poll, S. Kutalek, L. Gessman, L. Muenz, T. Friehling, R. Marinchak, K. Volosin, L. Beau regard, G. Kidwell, F. Kempf, and M. White (1990). Value of electrophysiologic testing in patients with previous myocardial infarction and nonsustained ventricular tachycardia. *Am J Cardiol.*, vol. 65, pp. 594–598.
- [36] Iesaka, Y., A. Nogami, K. Aonuma, J. Nitta, Y. Chun, H. Fujiwara, and M. Hiraoka (1990). Prognostic significance of sustained monomorphic ventricular tachycardia induced by programmed ventricular stimulation using up to triple extrastimuli in survivors of acute myocardial infarction. *Am J Cardiol.*, vol. 65, pp. 1057–1063.
- [37] Bourke, J., D. Richards, D. Ross, E. Wallace, M. McGuire, and J. Uther (1991). Routine programmed electrical stimulation in survivors of acute myocardial infarction for prediction of spontaneous ventricular tachyarrhythmias during follow-up: Results, optimal stimulation protocol and cost-effective screening. *J Am Coll. Cardiol.*, vol. 18, pp. 780–788.
- [38] Bhandari, A., R. Hong, A. Kotlewski, N. McIntosh, P. Au, A. Sankooikal, and S. Rahimtoola (1989). Prognostic significance of programmed ventricular stimulation in survivors of acute myocardial infarction. *Br. Heart J*, vol. 61, pp. 410–416.
- [39] Steinbeck, G., D. Andresen, P. Bach, R. Haberl, M. Oeff, E. Hoffmann, and E. Von Leitner (1992). A comparison of electrophysiologically guided antiarrhythmic drug therapy with beta-blocker therapy in patients with

- symptomatic, sustained ventricular tachyarrhythmias. *New Engl J Med*, vol. 327, pp. 987–992.
- [40] Denniss, A., D. Richards, D. Cody, P. Russell, A. Young, M. Cooper, D. Ross, and J. Uther (1986). Prognostic significance of ventricular tachycardia and fibrillation induced at programmed stimulation and delayed potentials detected on the signal-averaged electrocardiograms of survivors of acute myocardial infarction. *Circulation*, vol. 74, pp. 731–745.
- [41] Bhandari, A., J. Widerhorn, P. Sager, C. Leon, R. Hong, A. Kotlewski, J. Hackett, and S. Rahimtoola (1992). Prognostic significance of programmed ventricular stimulation in patients surviving complicated acute myocardial infarction: a prospective study. *Am Heart J*, vol. 124(1), pp. 87–96.
- [42] Gomes, J., S. Horowitz, M. Millner, J. Machac, S. Winters, and P. Barreca (1987). Relation of late potentials to ejection fraction and wall motion abnormalities in acute myocardial infarction. *Am J Cardiol*, vol. 59(12), pp. 1071–1074.
- [43] Denniss, A., D. Richards, D. Cody, and et al (1987). Correlation between signal-averaged electrocardiogram and programmed stimulation in patients with and without spontaneous ventricular tachyarrhythmias. *Am J Cardiol.*, vol. 59, pp. 586–590.
- [44] Kuchar, D., C. Thorburn, and N. Sammel (1987). Prediction of serious arrhythmic events after myocardial infarction: Signal-averaged electrocardiogram, Holter monitoring and radionuclide ventriculography. *J Am Coll. Cardiol.*, vol. 9, pp. 531–538.
- [45] Camm, A. and L. Fei (1995). Risk stratification following myocardial infarction: heart rate variability and other risk factors. In M. Malik and A. Camm, eds., *Heart rate variability*, Futura, Armonk, NY. pp. 369–392.
- [46] Zabel, M., T. Klingenhoben, M. Franz, and S. Hohnloser (1998). Assessment of QT Dispersion for Prediction of Mortality or Arrhythmic Events After Myocardial Infarction. Results of a Prospective, Long-term Follow-up Study. *Circulation*, vol. 97, pp. 2543–2550.
- [47] Fananapazir, L., D. McAreavey, and N. Epstein (1995). Hypertrophic cardiomyopathy. In D. Zipes and J. Jalife, eds., *Cardiac electrophysiology – From cell to bedside, 2nd edition*, W.B. Saunders Company, Philadelphia. pp. 769–779.

- [48] Mancini, D., H. Eisen, W. Kussmaul, R. Mull, J. Edmunds-LH, and J. Wilson (1991). Value of peak exercise oxygen consumption for optimal timing of cardiac transplantation in ambulatory patients with heart failure. *Circulation*, vol. 83(3), pp. 778–786.
- [49] Pedretti, R., M. Etro, A. Laporta, S. Braga, and B. Caru (1993). Prediction of late arrhythmic events after acute myocardial infarction from combined use of noninvasive prognostic variables and inducibility of sustained monomorphic ventricular tachycardia. *Am J Cardiol.*, vol. 71, pp. 1131–1141.
- [50] Zoni-Berisso, M., D. Molini, G. Mela, and C. Vecchio (1996). Value of programmed ventricular stimulation in predicting sudden death and sustained ventricular tachycardia in survivors of acute myocardial infarction. *Am J Cardiol*, vol. 77(9), pp. 673–680.
- [51] Ikeda, T., T. Sakata, M. Takami, N. Kondo, N. Tezuka, T. Nakae, M. Noro, Y. Enjoji, R. Abe, K. Sugi, and T. Yamaguchi (2000). Combined Assessment of T-Wave Alternans and Late Potentials Used to Predict Arrhythmic Events after Myocardial Infarction. *J Am Coll Cardiol*, vol. 35, pp. 722–730.
- [52] Anderson, K., V. Shusterman, S. Brode, V. Gottipaty, D. Schwartzman, and R. Weiss (1999). Noninvasive testing for selection of patients for electrophysiological study. *Annals of Noninvasive Electrophysiology*, vol. 4(4), pp. 434–442.
- [53] Schmidt, G., M. Malik, P. Barthel, R. Schneider, K. Ulm, L. Rolnitzky, J. Camm, J. Bigger, and A. Schmig (1999). Heart rate turbulence after ventricular premature beats as a predictor of mortality after acute myocardial infarction. *Lancet*, vol. 353, pp. 1390–96.
- [54] Bishop, C. (1995). *Neural Networks for Pattern Recognition*. Oxford University Press, Oxford, UK.
- [55] Kohonen, T. (1989). *Self-Organization and Associative Memory*. Springer-Verlag, Berlin.
- [56] Lloyd, S. P. (1957). Least Squares Quantization in PCM's. Tech. rep., Bell Telephone Laboratories Paper, Murray Hill, NJ.
- [57] Linde, Y., A. Buzo, and R. M. Gray (1980). An Algorithm for Vector Quantizer Design. *IEEE Transactions on Communication*, vol. 28, pp. 84–95.

- [58] MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*. vol. 1, pp. 281–297.
- [59] Backer, E. (1995). *Computer-assisted Reasoning in Cluster Analysis*. Prentice Hall, New-York.
- [60] Odorico, R. (1997). Learning Vector Quantization with Training Count (LVQTC). *Neural Networks*, vol. 10(6), pp. 1083–1088.
- [61] Kestler, H. A., A. A. Farschtschi, and P. H. Wein (1994). Untersuchung und Simulation Neuronaler Netze mit dynamischer Strukturänderung. *Automatisierungstechnische Praxis*, vol. 36(2), pp. 47–51.
- [62] Hertz, J., A. Krogh, and R. G. Palmer (1991). *Introduction to the Theory of Neural Computation*. Addison Wesley, New York.
- [63] Wasserman, P. (1993). *Advanced methods in neural computing*. Van Nostrand Reinhold, New York.
- [64] Armijo, L. (1966). Minimization of functions having Lipschitz continuous first partial derivatives. *Pacific Journal of Mathematics*, vol. 16(1), pp. 1–3.
- [65] Magoulas, G., M. Vrahatis, and G. Androulakis (1997). Effective Back-propagation Training with Variable Stepsize. *Neural Networks*, vol. 10(1), pp. 69–82.
- [66] Gomes, J., S. Winters, M. Martinson, J. Machac, D. Stewart, and A. Targonski (1989). The prognostic significance of quantitative signal-averaged variables relative to clinical variables, site of myocardial infarction, ejection fraction and ventricular premature beats. *JACC*, vol. 13, pp. 377–384.
- [67] Höher, M. and V. Hombach (1991). Ventrikuläre Spätpotentiale – Teil I Grundlagen. *Herz & Rhythmus*, vol. 3(3), pp. 1–7.
- [68] Simson, M. (1981). Use of Signals in the Terminal QRS Complex to Identify Patients with Ventricular Tachycardia after Myocardial Infarction. *Circulation*, vol. 64(2), pp. 235–242.
- [69] Breithardt, G. and M. Borggreffe (1986). Pathophysiological mechanisms and clinical significance of ventricular late potentials. *Eur Heart J*, vol. 7, pp. 364–385.

- [70] Breithardt, G., M. Cain, N. El-Sherif, N. Flowers, V. Hombach, M. Janse, M. Simson, and G. Steinbeck (1991). Standards for Analysis of Ventricular Late Potentials Using High Resolution or Signal-Averaged Electrocardiography. *Eur Heart J*, vol. 12, pp. 473–80.
- [71] Sammon, J. (1969). A nonlinear mapping for data structure analysis. *IEEE Transactions on Computers*, vol. C-18, pp. 401–409.
- [72] Ritscher, D. E., E. Ernst, H. G. Kammrath, V. Hombach, and M. Höher (1997). High-Resolution ECG Analysis Platform with Enhanced Resolution. *Computers in Cardiology*, vol. 24, pp. 291–294.
- [73] van Bommel, J. and M. Musen, eds. (1997). *Handbook of Medical Informatics*. Springer Verlag, Heidelberg / New York.
- [74] de Boor, C. (1978). *A Practical Guide to Splines*. Springer Verlag.
- [75] Kestler, H. A., M. Höher, G. Palm, M. Kochs, and V. Hombach (1996). Time Domain Variability of High Resolution Beat-to-Beat Recordings Classified by Neural Networks. In A. Murray and R. Arzbaecher, eds., *Computers in Cardiology*. IEEE Computer Society, pp. 317–320.