ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Analysis and Interpretation of Electrocardiogram Signals for the Detection of Hypoglycaemia

Charilaos Alexakis

A thesis submitted in partial fulfilment of the requirements of
Sheffield Hallam University
for the Degree of Doctor of Philosophy

May 2005

Collaborating Organisation: Northern General and Royal Hallamshire Hospitals

Abstract

Diabetes is a complication of metabolism where the glucose control system of the human body is impaired and cannot preserve the blood glucose levels in the normal range. This research investigated the relationship between abnormally low glucose levels (hypoglycaemia) and cardiac function in human subjects with Type 1 diabetes. The aim of the research was to detect the onset of spontaneous nocturnal hypoglycaemia indirectly through analysis of the subject's Electrocardiogram (ECG). The research hypothesis follows from previous studies, that suggested changes in ECG morphology, in particular prolongation of the QT interval and flattening of the T-wave, during hypoglycaemia.

The research methodology involved ECG feature extraction and classification of extracted features into euglycaemic (normal glucose levels) and hypoglycaemic categories. A number of time-domain ECG features were evaluated and a few ECG annotation algorithms were investigated for detection of onsets, peaks and offsets of the ECG components. Autoregressive (AR) modelling was also employed as a means of describing and characterising post-QRS ECG segments. ECG segment classification was carried out using Multi-layer Perceptron (MLP) neural networks. Statistical classifiers were also employed namely, Linear Discriminant Analysis (LDA) and the k-Nearest Neighbour (kNN).

This research proposed a new methodology for detection of spontaneous nocturnal hypoglycaemia by combining time-domain characterisation and classification of the post-QRS ECG segment. Two novel ECG features were introduced to characterise T-wave morphology. MLPs achieved better classification of ECG feature vectors compared to LDA. Also ECG representation by AR coefficients was marginally superior to individual ECG features, according to classification performance by LDA. Finally a Knowledge-Based System (KBS) was designed for ECG monitoring during the night. It was developed and tested on offline data in a manner that simulated an online monitoring scenario. The system was able to detect ECG abnormalities related to spontaneous nocturnal hypoglycaemia and to raise an alarm if necessary. In its optimal configuration, the system correctly monitored 30 out of the 32 recorded nights (originating from 19 patients) while there were 2 false alarms. This performance corresponds to accuracy, sensitivity and specificity of 93.75%, 100% and 91.30% respectively.

The main contribution to knowledge from this research was successful detection of the onset of spontaneous nocturnal hypoglycaemia indirectly, using solely ECG information. This result supports the hypothesis stating that spontaneous hypoglycaemia affects the cardiac function and is manifested on the ECG. A detailed analysis of the ECG signal for the detection of hypoglycaemia was carried out in the thesis. ECG features were extracted and assessed as predictors of the clinical condition. A number of approaches for ECG representation and classification (MLP, kNN, LDA) were examined and compared. Moreover, a KBS capable of achieving satisfactory monitoring performance on offline data from diabetic patients was designed. It was found that ECG changes in response to hypoglycaemia were short-time transients and incorporation of temporal information in the classification system caused significant improvement in performance. Successful continuation of this work may lead to a hypoglycaemia-detection system for the bedside.

Quote

Σε κάθε δύσκολη στιγμή απού θα συναντήσω,
Βάνω το μπέτη μου μπροστά και δε γιαγέρνω πίσω.
Ψαραντώνης – Ιδαίον Άντρον [1]

[1] The above two verses constitute a "mantinada" by the Cretan artist "Psarantonis". A mantinada, consisting of two 15-syllable verses, lies among the most characteristic forms of traditional Cretan poetic expression.

Dedication

Αυτή η διατριβή αφιερώνεται με πολλή αγάπη στην γιαγιά μου, Μαρία Πετράκη, για την ατέλειωτη και συνεχή αγάπη και προσφορά της και επειδή στάθηξε και στέκεται άξιο πρότυπο ανθρώπου.

(Dedicated to my beloved grandmother, Maria Petrakis, for her endless love and support and for being a model human being.)

Acknowledgements

When I had first started flicking through the pages of the dissertations and theses of my postgraduate friends, at the early years of my Bachelors' degree, the only thing I could comprehend from the material was the "acknowledgements" section! I believe it is common knowledge that this section is probably, and very sadly, the only passage of a thesis that could be of interest to those who are not aware of, or interested in the area of research! I hope it is an interesting read for them.

The journey towards a doctorate degree is a lonely one. Nevertheless, many people are watching you and it seems they can't do much to help you, but they actually have a big impact on you making it to the end.

First of all I would like to thank my initial director of studies Prof Henry Nyongesa for giving me the opportunity to undertake this research, and for his guidance and expert advice throughout the course of the project. I am also very grateful to Prof Marcos Rodrigues who agreed to act as my new director of studies, after Prof Nyongesa left Sheffield Hallam University. Although he joined later in the project his contribution was significant, especially in the final stages of thesis preparation and proof-reading. I would also like to thank him for successfully arranging a nine-month extension to my studentship, which was of much help. Last but not least, I would like to express my gratitude to my supervisor, Dr Reza Saatchi for being very supportive and for providing me with his guidance, advice and expertise throughout the course of the project. Moreover, for being very generous timewise and devoting long hours and energy for informal discussions on the research.

Many thanks should also be addressed to the medical collaborators of this project for directing me in the biomedical aspects of this work and also for providing the datasets. Dr Nigel Harris and Dr Simon Heller were very helpful throughout the course of the project and also very supportive and encouraging. They provided lots of advice, insights, and their medical expertise. Mrs Cath Davies provided a lot of help and direction with the datasets, data format, ECG annotating etc. Dr Celia Emery was also very helpful whenever I needed it. Finally I would like to thank Dr Rob Ireland for useful discussions on the research and mainly for his support regarding the ECGLAB toolbox.

The diabetic patients who provided the data used in this research are greatly acknowledged. Without their contribution this research would not have been possible.

I am grateful to Sheffield Hallam University for supporting me financially by means of a Digital Technologies Research Centre (DTRC) Studentship. I would also like to thank Medtronic MiniMed Inc for supporting financially the data acquisition in the form of glucose meters and transcutaneous probes.

I greatly appreciated all the discussions and informal chats I had with friends and colleagues in the lab and generally at Sheffield Hallam University. I would specifically like to thank Bala and Selvan, for their help, advice and constructive discussions during the past four years. Also George, who provided food for thought at various instances.

Special thanks should go to my friend Demetrios with whom I followed the same long route and shared similar problems for the past five and a half years. I am also grateful to Grigoris for his support and motivation in circumstances where I really needed it, and Myrto for her encouraging words and for insightful discussions.

The circle of friends in Sheffield was changing often due to the nomadic lifestyle of Masters' students who have to move on after 12 months or so. It is therefore extremely difficult to address my thanks by name, without leaving people out. I believe that all those not addressed here by name know who they are!

I am indebted to my friends back in Greece, who are many to mention, for being there when I needed them and for contributing towards very refreshing vacation periods when I was visiting Greece. I would especially like to thank George (Fotakis) for his friendship and for always keeping an eye on me during the course of this PhD, although he was miles away.

Finally I would like to thank my parents, Michalis and Athena, my beloved sister Maria and the super-nanny (!), my lovely grandmother Maria, for their continuous love and encouragement and for so many other little or not so little things that it would take ages to list here.

C Alexakis

List of Publications

Material from this thesis has appeared in the following publications, presented in reverse chronological order:

C Alexakis, R Saatchi, M Rodrigues, HO Nyongesa, ND Harris, C Davies, RH Ireland, C Emery, SR Heller, The use of a knowledge-based monitoring system for detection of abnormal cardiac repolarization during spontaneous hypoglycaemia, in patients with Type 1 diabetes. Diabetic Medicine, April 2005, Volume 22, Supplement 2, pp 39
(Awarded the Eli Lilly best poster award for the Clinical Science stream at the Diabetes UK Annual Professional Conference, 20-22 April 2005, SECC Glasgow.)

C Alexakis, M Rodrigues, R Saatchi, HO Nyongesa, ND Harris, SR Heller, "Detection Of Abnormal Electrocardiogram Traces, Related To Hypoglycaemia, Using A Knowledge-Based Approach", Postgraduate Research Conference in Electronics, Photonics, Communications & Networks, and Computing Science (PREP) 2005, University of Lancaster, 30 March - 1 April 2005

HO Nyongesa, C Alexakis, R Saatchi, M Rodrigues, ND Harris, SR Heller, "Classification Of SAECG By Auto-Regressive Modelling And Neural Networks", IEEE Africon 04, Gaborone, Botswana, 15 – 17 September 2004

C Alexakis, HO Nyongesa, R Saatchi, M Rodrigues, SR Heller, C Davies, C Emery, ND Harris, RH Ireland, "Classification Of Signal-Averaged Electrocardiogram Traces Modelled By Auto-Regressive Modelling", Postgraduate Research Conference in Electronics, Photonics, Communications & Networks, and Computing Science (PREP) 2004, pp 25-26, University of Hertfordshire, Hatfield, 5 – 7 April 2004

C Alexakis, HO Nyongesa, R Saatchi, ND Harris, C Davies, C Emery, RH Ireland, SR Heller, "Feature Extraction and Classification Of Electrocardiogram (ECG) Signals Related to Hypoglycaemia", Computers in Cardiology 2003, vol 30, pp 537-540, Thessaloniki, Greece, 21-24 September 2003

C Alexakis, HO Nyongesa, R Saatchi, ND Harris, C Davies, C Emery, SR Heller, "Detection Of Hypoglycaemia-Induced Cardiac Arrhythmias Using Neural and Statistical Classifiers", 5th International conference on Neural Networks and Expert Systems in Medicine and Healthcare - 1st International conference on Computational Intelligence in Medicine and Healthcare (NNESMED-CIMED 03), pp 17-22, Sheffield, England, 21-23 July 2003

C Alexakis, HO Nyongesa, R Saatchi, ND Harris, C Davies, C Emery, SR Heller, "Detection Of Hypoglycaemia-Induced Cardiac Arrhythmias Using Artificial Neural Networks", Eunite 2003 Symposium, pp 123-126, Oulu, Finland, 10-11 July 2003

C Alexakis, H Nyongesa, R Saatchi, N Harris, S Heller, C Davies, "Feature extraction and classification of ECG signals related to type 1 diabetic patients.", Postgraduate Research Conference in Electronics, Photonics, Communications & Networks, and Computing Science (PREP03), pp 151-152, University of Exeter, Exeter, England, 14-16 April 2003

C Alexakis, R Saatchi, ND Harris, C Davies, C Emery, SR Heller, HO Nyongesa, "The use of artificial neural networks to identify abnormal cardiac repolarization during clinical hypoglycaemia in patients with Type 1 diabetes.", Diabetic Medicine, April 2003, Volume 20, Supplement 2, pp 112
(presented in Diabetes UK Annual Professional conference held in Glasgow, Scotland on 19-21 March 2003)

List of Contents

Key to Abbreviations and Symbols

ADC	analogue-to-digital converter
aka	also known as
ANN	artificial neural networks
AR	Auto-Regressive
BSS	Blind Source Separation
CGMS	Continuous Glucose Monitoring System
dc	direct current
DOF	degrees of freedom
ECG	ElectroCardioGram
EMG	ElectroMyoGram
ES	Expert System
eugly	euglycaemia
FIS	Fuzzy Inference System
HB	Healthy Band
HP	Hewlett Packard
HRECG	High-Resolution ECG
hypo	Hypoglycaemia
i.v.	Intra-venous
KBS	Knowledge-Based System
kNN	k-Nearest Neighbour
LDA	Linear Discriminant Analysis
LHS	Left-Hand Side
max	maximum
MF	Membership Function
min	minimum
MLP	Multi-Layer Perceptron
msec or ms	milliseconds
mV	milliVolts
NGH	Northern General Hospital
NSB	Normal Sinus Beat
PC	Personal Computer
p.d.f.	probability density function
perf	performance
PVC	Premature Ventricular Contraction
rec	record
RHH	Royal Hallamshire Hospital
RHS	Right-Hand Side

RTapexcWS	Window Size for the RTapexc feature
SAECG	Signal-Averaged ECG
scRTapexc	scaling factor for setting the HB width of RTapexc
scTampl	scaling factor for setting the HB width of Tampl
sd (or SD)	standard deviation
SISO	Single-Input Single-Output
SSS	Sick Sinus Syndrome
std	standard deviation
TamplWS	Window Size for the Tampl feature
U.S.	United States of America
ver	version
vs	versus
WS	Window Size
+ve	positive
-ve	negative
$\aleph$	set of Natural numbers
$\Re$	set of Real numbers

Uppercase letters in **bold** denote matrices or vectors

Closed numeric intervals denoted by square brackets "[ ]", e.g. $[x_1 x_2]$

Open numeric intervals denoted by standard brackets "( )", e.g. $(x_1 x_2)$

Performance metrics of classifiers are always given in the order: accuracy, sensitivity and specificity when further clarification is not given.

Introduction

1.0 Introduction

The aim of this thesis was to investigate the relationship between hypoglycaemia and cardiac function and to attempt detection of the onset of nocturnal hypoglycaemia, in Type 1 diabetic patients, indirectly through analysis of their Electrocardiogram (ECG). The research focused on the development and use of feature extraction and signal classification techniques in order to analyse the ECG signals. This chapter presents the biomedical background to the study, the motivation behind this research and the detailed objectives set. It also presents the goals reached and outlines the structure of the thesis.

1.1 Biomedical Background

This section provides theoretical background on the medical condition of diabetes and the complication of hypoglycaemia. It also describes the architecture of the human heart and the processes of depolarisation and repolarisation of cardiac cells. It includes information about electrocardiography and the electrocardiogram and discusses the relationship between hypoglycaemia and cardiac function.

1.1.1 Diabetes and Hypoglycaemia

Diabetes is derived from the Greek word "diabainein" (διαβαίνειν) meaning "to pass" because in a specific form of diabetes (diabetes mellitus) glucose is passed through the body and out with the urine. Diabetes is a complication of metabolism where the glucose control system is impaired and is not able to maintain the blood glucose in the normal physiological range.

Insulin is the hormone that serves to lower glucose levels and is produced by the pancreas. The complication of diabetes and the root cause of most of the damage it does

is related to excess glucose in the blood. This happens because the cells of the pancreas do not produce sufficient amounts of insulin. Because of the insufficient amount of insulin the glucose cannot be controlled effectively. The only treatment to the absence of insulin is to replace it. This is done by injection or by infusion through an insulin pump.

Hypoglycaemia is the opposite complication i.e. not enough sugar in the blood, which is a dangerous situation since glucose is needed to maintain brain function. Severe hypoglycaemia can lead to coma and even death. There is strong circumstantial evidence that hypoglycaemia can cause overnight death in young adults and children, a syndrome known as "Dead in Bed" [Campbell 1991]. The mechanism of such deaths remains unclear but may be cardiac related.

The form of diabetes in which the urine contains glucose matter is called diabetes mellitus. There are two different types of diabetes mellitus:

- Type 1 diabetes or insulin-dependent diabetes is a severe, acute form of diabetes caused by lack of production of insulin. The disease, which typically appears in childhood or adolescence, is characterized by increased sugar levels in the blood and urine, excessive thirst, frequent urination, acidosis and wasting.
- Type 2 diabetes or non-insulin-dependent diabetes is a chronic form of diabetes that typically appears in late adulthood and is exacerbated by obesity and an inactive lifestyle. In this type of diabetes the patient develops insulin resistance and is not able to effectively use the insulin produced by the pancreas. At onset this disease often has no symptoms, is usually diagnosed by tests that indicate glucose intolerance, and is treated with changes in diet and an exercise regimen.

Diabetic patients need to monitor their glucose levels and act appropriately to keep them in the normal range, which lies between 4 and 8 mmol/lt approximately. Patients have the choice to keep their glucose levels as close as possible to this range or to allow their glucose to vary in a wider range than the normal. In order to perform tight glycaemic control, they need to monitor their glucose levels in the blood at regular intervals and to inject regular small quantities of insulin before meals. In Type 2 diabetes they can diet or use oral medication.

Quality of Glucose Control. It has been shown [DCCT 1991] that tight glycaemic control in patients with insulin-dependent diabetes mellitus has the advantage of reducing the frequency and progression of serious long term complications to the patient, compared to loose control; the drawback being that hypoglycaemia, a short term complication, is more frequent [Davis 1998]. The patients are therefore faced by the dilemma of whether (i) to set higher glucose targets and achieve reduced risk of hypoglycaemia but increase the risk of long term diabetic complications or (ii) to set lower glucose targets and reduce the risk of long term complications but increase the risk of hypoglycaemia.

Because many patients nowadays try to achieve tight control, there is a greater need for a continuous glucose-monitoring device or a hypoglycaemia-detection device. Such a device may also help diabetic patients who fail to achieve tight control for psychological reasons such as the fear of hypoglycaemic episodes. Patients often allow their glucose to run high because of fear of the other extreme, that of hypoglycaemia.

Clinical studies of hypoglycaemia. Spontaneous hypoglycaemia occurs naturally when the glucose of the diabetic patient drops to abnormally low levels and is difficult to study clinically because it is relatively infrequent. However, some studies have utilised experimental hypoglycaemia which is caused artificially by infusing controlled amounts of insulin to a non-diabetic subject and "clamping" the blood glucose at an abnormally low value (e.g. 2.5 mmol/lt) for a short period of time. Experimental hypoglycaemia has been found to introduce more prominent changes on the ECG [Marques 1997] compared to spontaneous hypoglycaemia. The spontaneous hypoglycaemic events used in this research were nocturnal and the subject was asleep whereas in experimental studies the subject is normally kept awake, in a supine and relaxed position.

1.1.2 Architecture of the heart

The cardiovascular system is responsible for the rapid transport of oxygen, water and nutrients around the body and the rapid washout of metabolic waste products like carbon dioxide. It also plays a vital role in temperature regulation, by delivering heat from the core of the body to the skin. It consists of: a pump (the heart), a series of distributing and collecting tubes and an extensive system of thin vessels that permit rapid exchange between the tissues and the vascular channels.

The heart organ is the muscular pump of the cardiovascular system that drives the blood around the body. The human heart consists of two muscular pumps in series, the right and left ventricles (Figure 1.1). Each pump is filled from a reservoir, the right or left atrium. The right ventricle propels blood through the lungs for exchange of oxygen and carbon dioxide (pulmonary circulation) and the left ventricle simultaneously propels blood to all the other tissues of the body (systemic circulation). Unidirectional flow through the heart is achieved by the appropriate arrangement of effective flap valves. Although the cardiac output is intermittent, continuous flow to the periphery occurs by distension of the aorta and its branches during ventricular contraction (systole) and elastic recoil of the walls of the large arteries with forward propulsion of the blood during ventricular relaxation (diastole).

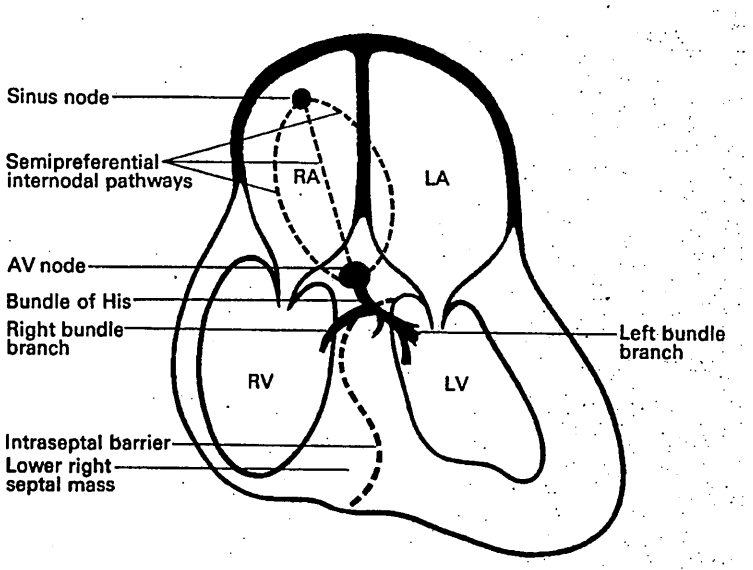


Figure 1.1: The left and right atria (LA, RA) and ventricles (LV, RV) of the heart [McLachlan 1981]

The outstanding feature of the heart's function is its ability to initiate and maintain a rhythmical beat. As with all muscle, the heart contracts in response to an electrical impulse which spreads throughout its surface, reaching the contractile muscle cells. It is the spread of this impulse which the electrocardiogram records, not the muscular contraction which follows [McLachlan 1981].

Depolarisation and repolarisation of myocardial cells

The depolarisation and repolarisation of the heart muscle will be described in a cellular level. The two terms will be defined for an isolated cell of the cardiac muscle. Consider the resting muscle cell being in a state of electrical equilibrium with positive charges on the outer surface of its membrane and negative charges on the inner surface. In this resting state the cell is said to be polarised and remains stable until it is stimulated.

In response to a stimulus the cardiac cell begins to depolarise. The charges on the outer surface of the cell become negative and those on the inner surface become positive. This change in polarity does not happen instantly for the whole cell surface. It starts from the point where the stimulus was received and propagates along the cell as seen in Figure 1.2. The advancing boundary of change is called the activation front. Once the electrical charge has changed polarity, throughout the cell, then depolarisation has finished and the cell is said to be depolarised.

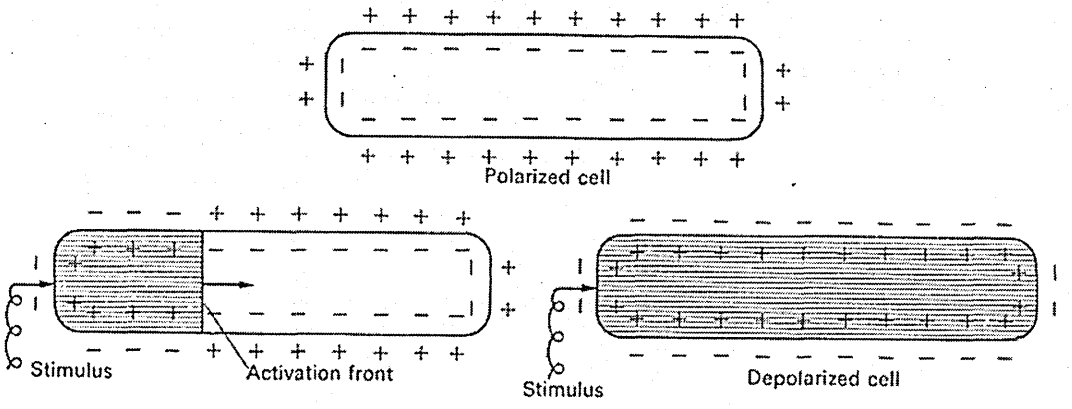


Figure 1.2: Illustration of the depolarisation process [McLachlan 1981]

The inverse process is called repolarisation and metabolic energy is required for it to take place. Once the cell is completely depolarised it must be returned to its polarised state in preparation for the next depolarising stimulus. In the case of repolarisation the advancing boundary of change is called repolarisation front. The repolarisation of the ventricles is the cardiac process of interest in this thesis. As will be mentioned in more detail later in the chapter, hypoglycaemia has been observed to cause a delay in ventricular repolarisation which has been associated with the risk of cardiac arrhythmias.

1.1.3 Electrocardiography and the electrocardiogram

Electrocardiography is used as a method for studying the action of the heart muscle. It can be defined as the procedure of making graphical records of the variations in electrical potential caused by electrical activity of the heart muscle and detected at the body surface.

The ECG lies among the most widely used signals in biomedical practice. It describes the electrical activity of the heart. The standard ECG consists of 12 leads but the 3-lead ECG was used instead, as it was sufficient for this work (Section 3.2.1). The latter consists of three leads in an anatomically orthogonal configuration which are referred to as XYZ leads, similar to the coordinate axes used in geometry [Berbari 2000].

The utilisation of only three leads was probably due to the fact that technology used by early investigators was limited in the number of low-noise amplifiers available and the relatively slow speed of computers [Berbari 2000]. Reducing the number of leads from 12 to the orthogonal (XYZ) set of three leads seemed to be a reasonable compromise from both computing and physiological standpoints [Berbari 2000]. The anatomical locations for the XYZ leads are presented in Figure 1.3.

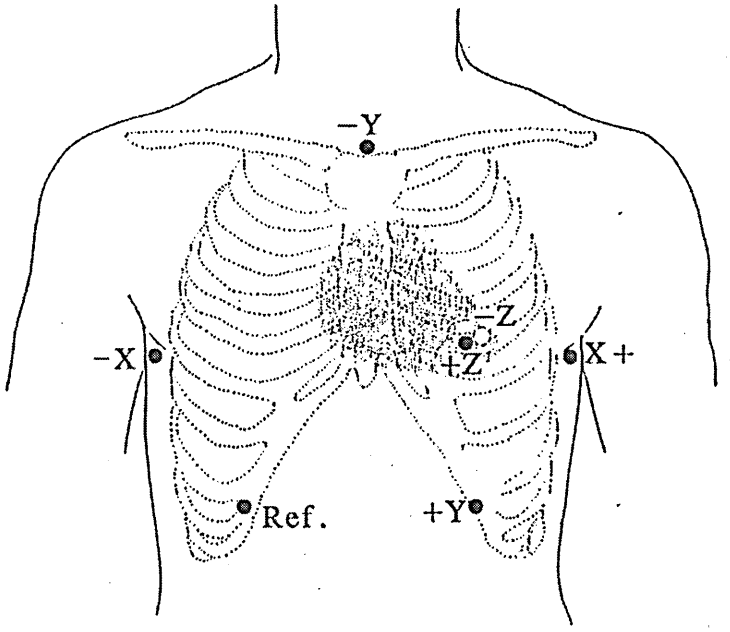


Figure 1.3: Anatomical locations for the XYZ leads [AMA 1988]

The figure depicts an idealised male torso with the 6 electrodes (X+, X-, Y+, Y-, Z+, Z-) forming the orthogonal lead set, plus a reference electrode. It is obvious from the graph that the three electrode pairs form three axes. The Z+ electrode is located at the same level as the X lead and is placed on the anterior chest. The Z- electrode is placed on the reflection of the Z+ electrode on the patient's back.

Morphology of the electrocardiogram

A typical ECG cycle consists of three waves namely the P wave, the QRS complex and the T wave as shown in Figure 1.4. A U wave following the T wave is sometimes present on ECG cycles. The R peak is the peak with the highest amplitude in an ECG cycle.

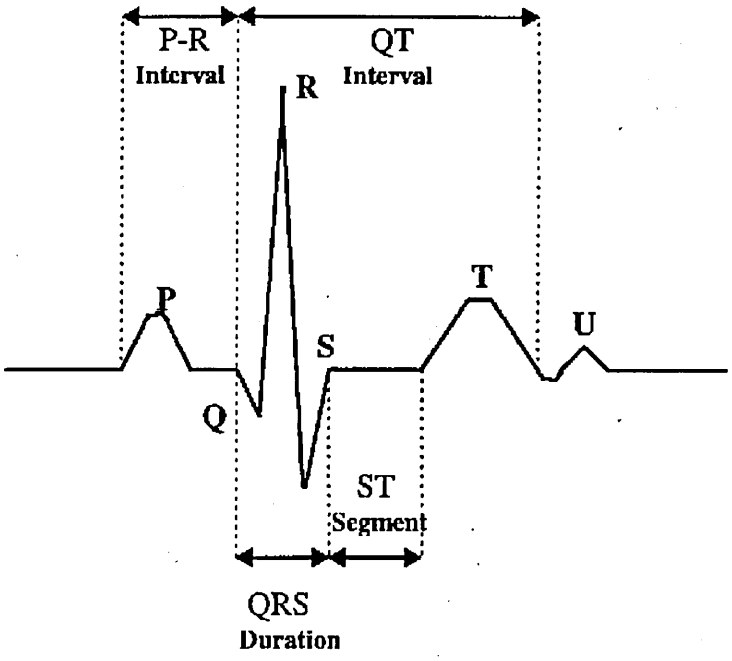


Figure 1.4: A typical ECG cycle [Gholam-Hosseini 1998]

The time duration between two successive R peaks denotes the instantaneous heart rate. The P wave corresponds to the atrial depolarisation, the QRS complex to the ventricular depolarisation and the T wave is due to the ventricular repolarisation of the myocardium. Figure 1.5 shows the mapping of each wave to the corresponding cardiac event and the corresponding regions in the heart.

A number of key time-intervals can also be identified in Figure 1.4. The time-intervals related to this research are: the QRS duration, the ST segment and most importantly the QT interval. The latter is very widely used in the biomedical community and is defined as the interval from the Q point to the end of the T wave.

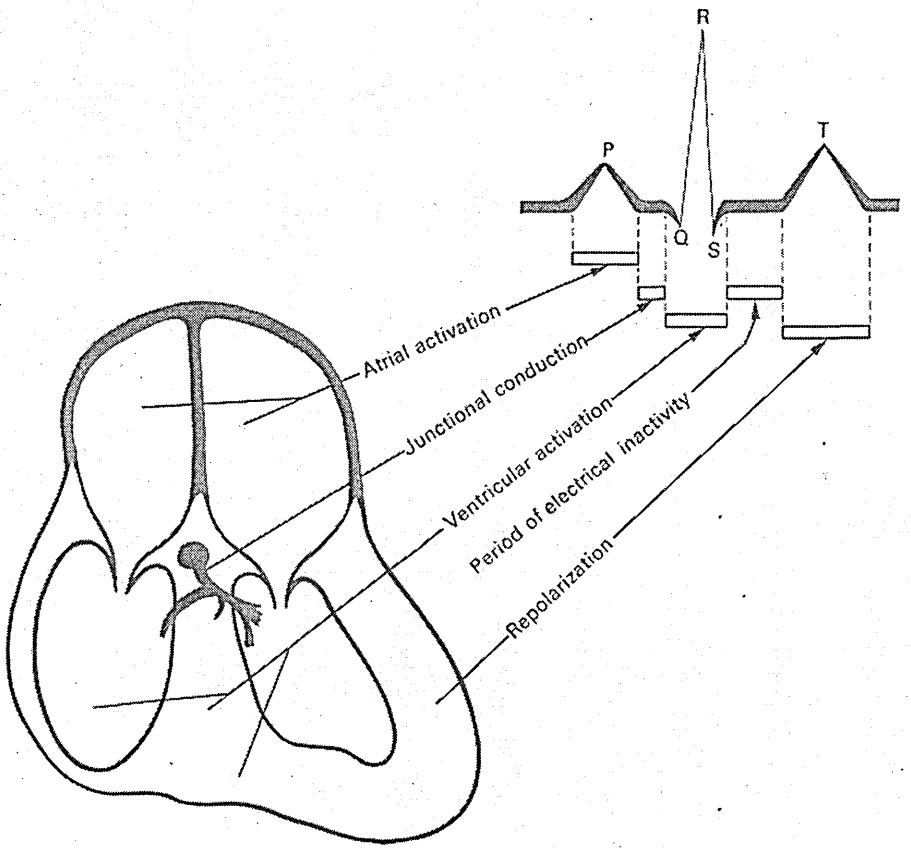


Figure 1.5: Relationships between the cardiac events and the waves of the ECG cycle [McLachlan 1981]

1.1.4 Effect of potassium and adrenaline on the cardiac function

Cardiac function is affected by both variations in potassium and adrenaline, beyond the normal range. The fundamental importance of potassium ions in the cellular mechanism of muscle contraction is that variations in their concentration beyond the normal range in the intracellular fluid may be expected to influence the clinical ECG [McLachlan 1981]. When cellular potassium levels are depleted, ST segment depression takes place and also a prominent U wave blended with a flat T wave gives a false impression of a prolonged QT interval [McLachlan 1981]. Under hypoglycaemia serum potassium often falls to low levels due to a combination of high insulin concentrations and the effects of sympathoadrenal activation [Harris 2000].

Adrenaline causes lengthening of the QTc interval and a decrease in plasma potassium (K^+), when infused into healthy subjects at concentrations comparable to those seen during moderate hypoglycaemia [Cryer 1980]. Although adrenaline causes a drop in plasma potassium, Lee et al [Lee 2003] have shown that disturbed cardiac repolarisation as a result of increases in circulating adrenaline, released due to hypoglycaemia, occurs independently of extracellular potassium.

1.1.5 Relationship between hypoglycaemia and cardiac function

During hypoglycaemia, the counter-regulatory responses of the human body cause a fall in potassium and the release of adrenaline, which delays the Ventricular Repolarisation process (VR) in the heart. These changes may be reflected on the ECG by changes in T wave morphology. A hypothesis has been formulated stating that there is a relationship between abnormally low glucose levels and cardiac function. In more detail, hypoglycaemia can cause abnormal cardiac repolarisation and an attendant risk of cardiac arrhythmia [Robinson 2004].

This hypothesis is based on preliminary evidence from a number of studies [Heinemann 1995, Marques 1997]. Marques et al [Marques 1997] have shown that experimental hypoglycaemia prolongs the ventricular repolarisation of the heart, causing the development of cardiac arrhythmias. Although the hypothesis is supported by experimental results, sufficient evidence is necessary to demonstrate the hypothesis for spontaneous events of nocturnal hypoglycaemia. Early studies demonstrate agreement with the hypothesis for spontaneous nocturnal events. Robinson et al [Robinson 2004] have recently shown that the QTc interval lengthens significantly during spontaneous nocturnal hypoglycaemia but the increase is generally less than that observed during experimental episodes.

Since hypoglycaemia is believed to affect the cardiac function, then its onset is expected to be reflected on the ECG. Hypoglycaemia can affect the ST segment and the T wave i.e. the ventricular repolarisation of the heart. More specifically the T wave becomes flattened and prolonged under hypoglycaemia and often another wave, the U wave, is present. Figure 1.6 illustrates potential changes on the ECG due to dropping glucose. The illustration in the figure originates from real data. Each pair of graphs corresponds to a different time instant and the graphs are successive in time from top left to bottom right. At each time instant the left-hand-side graph demonstrates a truncated T wave and the right-hand-side graph depicts the glucose level up to the current time instant. The portion of the ECG cycle prior to the T wave is not shown. It is obvious from the graph that dropping glucose to abnormal levels (below 3 mmol/lt approximately) causes significant changes on the T wave. The T wave amplitude drops significantly while the symmetry of the same wave changes slightly. Moreover the temporal location of the T peak is slightly moved to the right. The end of the T wave is also shifted slightly to the right and its exact position becomes more ambiguous compared to the initial normal

trace. The shifting of the end of the T wave to the right also causes a proportional increase in QT interval since its end point is defined by the end of the T wave.

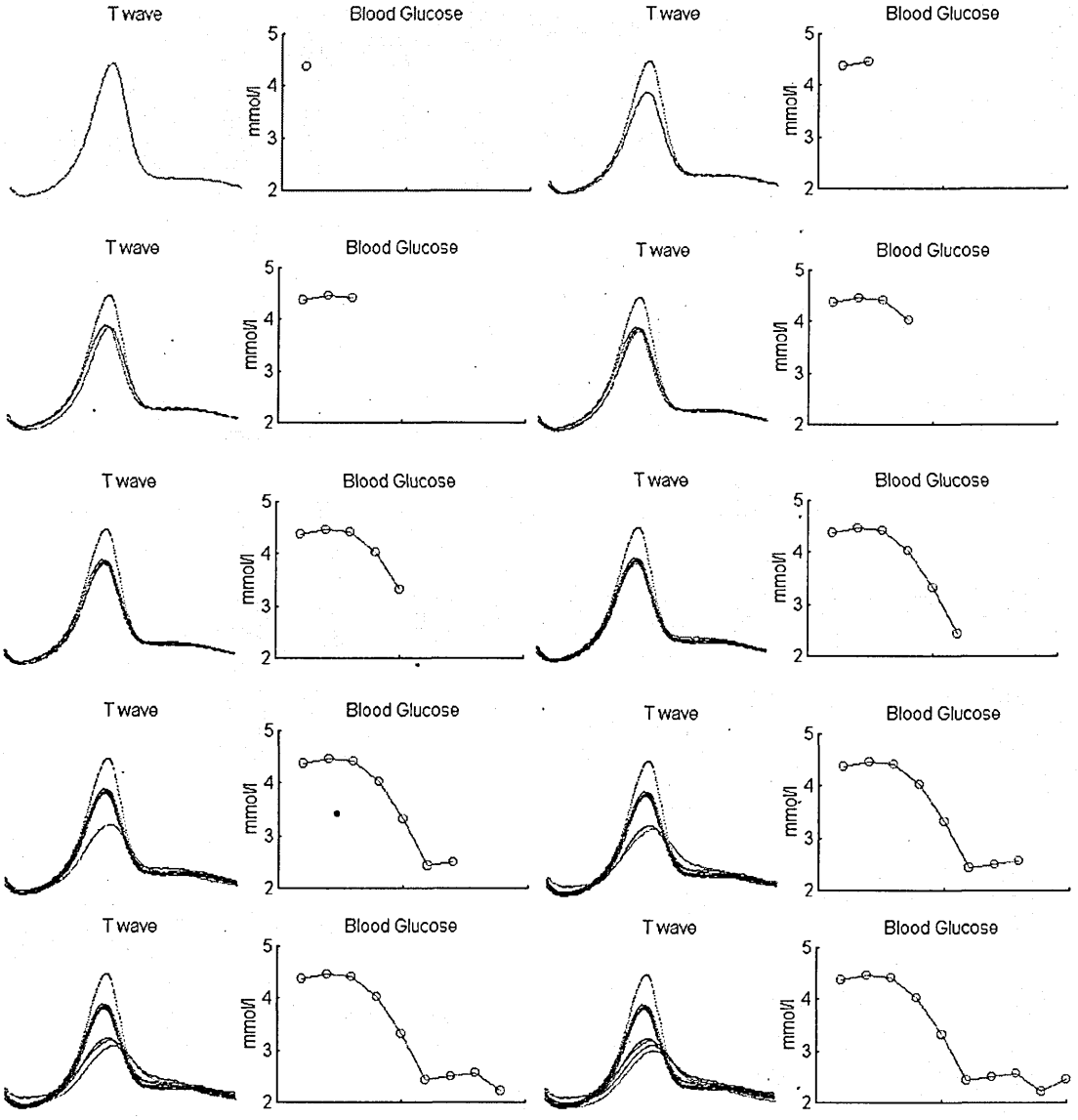


Figure 1.6: Illustration of changes in T wave with dropping glucose [URL 1].

Since hypoglycaemia may be reflected on the ECG and may introduce cardiac arrhythmias, the investigation of this biomedical signal is proposed as an indirect means of detection of this medical condition. If the changes in the ECG can be automatically identified, this may provide a warning of hypoglycaemia or of a potentially pro-arrhythmogenic condition.

It is emphasised that the manifestation of hypoglycaemia on the ECG is more likely to be due to the high circulating adrenaline affecting the cardiac function rather than due to the abnormally low glucose affecting it. Therefore this research is practically attempting to detect the cardiac changes caused by the counterregulatory responses triggered by

hypoglycaemia. This is a secondary effect. The direct effect would be if abnormally low glucose was directly affecting the cardiac function.

1.2 Motivation behind the research

Diabetes is one of the most common causes of severe morbidity and disability in the United Kingdom, affecting approximately 1.5 million people. In spite of improvements in insulin therapy around 40% of patients develop serious long-term complications. It has been shown that the incidence of these complications can be significantly reduced if blood glucose control is tightly maintained [DCCT 1993]. However, patients who attempt intensive insulin therapy are three times more likely to suffer a severe hypoglycaemic episode, which can lead to coma or death. Consequently, fear of hypoglycaemia is probably the main reason why patients fail to achieve tight control of blood glucose. Furthermore, the symptoms of hypoglycaemia cannot be recognised by young patients or if the patient is asleep and this has been implicated in the sudden overnight death of patients with diabetes, perhaps due to the development of cardiac arrhythmias [Tattersall 1991].

While early diagnosis of diabetes is known to be critical to its treatment, good management remains the only means to avoid complications of the disease. Since diabetes is patient managed, information, decision support and alarm tools are particularly valuable to diabetic patients. The ideal solution for detection of hypoglycaemia would be the development of a transcutaneous or non-invasive sensor, which could monitor glucose concentrations continuously. Although recent technological advances offer the promise of a transcutaneous sensor in the medium to long term, there are still formidable technological problems to overcome [Ireland 2000] and the initial cost may prohibit widespread use.

In context with the aforementioned points, this research work investigates the possible manifestation of hypoglycaemia on the patient's electrocardiogram. Such a manifestation may contribute in demonstrating a clear relationship between spontaneous nocturnal hypoglycaemia and delayed ventricular repolarisation and hence it may aid the explanation of the mechanisms behind the "Dead in Bed" syndrome. The possibility of producing an alarm system for hypoglycaemia, based solely on secondary responses caused by dropping of glucose to abnormally low levels is also examined. Until an affordable and robust transcutaneous or non-invasive glucose sensor can be

manufactured, an alarm system based solely on monitoring the ECG may partly solve the problem of detection of the onset of nocturnal hypoglycaemia.

1.3 Detailed objectives

The detailed objectives of this research programme are given below:

1. To formulate a hypoglycaemia detection methodology based solely on information from the ECG. The methodology will involve an ECG preprocessing, a feature extraction and a classification/inference stage.
2. To extract ECG features that will be able to quantify the suspected cardiac changes introduced by hypoglycaemia. This involves the re-use and fine-tuning of existing features besides the introduction of new ones. Emphasis is given on the introduction of features for assessment and characterisation of T wave morphology.
3. To select optimal feature combinations to be used for classification of normal and hypoglycaemic ECG traces.
4. To compare both neural and statistical classifiers in the task of discriminating between normal and hypoglycaemic ECG traces.
5. To design and develop a classification system that can distinguish normal ECG waveforms from ECGs reflecting hypoglycaemia-induced cardiac arrhythmias. The system should address overnight monitoring of patients for detection of possible hypoglycaemic events.
6. To design a Knowledge-Based System (KBS) to be used as the classification engine of the patient monitoring system, making use of existing knowledge from human experts. To compare the KBS against the neural and statistical classifiers, regarding their ability to carry out the classification/inference task within the patient monitoring system.
7. By combining all research findings to attempt an assessment of the hypothesis suggesting a relationship between hypoglycaemia and cardiac function.
8. To assess the feasibility of detection of hypoglycaemic events and hypoglycaemia-induced arrhythmias using a specialised ECG monitoring system.

1.4 Contribution to Knowledge

The main contribution to knowledge from this research was successful detection of spontaneous nocturnal hypoglycaemia in some Type 1 diabetic patients. This result supports the hypothesis stating that hypoglycaemia affects the cardiac function (Section 1.1.5). There was strong evidence of such a relationship for the case of experimental

hypoglycaemia. The results from this research support the hypothesis for the case of spontaneous hypoglycaemia where more subtle ECG changes occur and their detection is more challenging.

During the process of achieving this result, we have proposed a methodology for implementing a diagnostic system to be used in hypoglycaemia monitoring. The system consists of an ECG representation stage in cascade with a classification stage, discussed in Chapter 4. In the ECG representation stage, each ECG cycle is characterised by the parameters extracted from it. In the classification stage these parameters are distinguished into those corresponding to euglycaemia (normality) and hypoglycaemia. The proposed system was implemented as an early software prototype. It was tested on offline data from Type 1 diabetic patients experiencing spontaneous nocturnal hypoglycaemic episodes.

Detailed analysis of the ECG signal was carried out, for examination of the above relationship. In more detail:

- A number of time-domain ECG features were extracted for describing the cardiac changes occurring under hypoglycaemia. Novel ECG features were introduced besides reusing and modifying existing ones.
    - Two novel ECG features, inspired from the third and fourth central moments used in statistical theory, were introduced for the evaluation of T wave symmetry and morphology.
    - The concept behind an existing feature that has been used, among other features, as a predictor of the Long QT Syndrome [Benhorin 1990] was borrowed and modified accordingly to produce a third feature for assessing T wave symmetry. The new feature was based on the ratio of the two areas under the T wave to the left and right of the T peak.
- A comparative study of geometric methods for marking the T wave end was carried out using data from Type 1 diabetic patients.
- AutoRegressive (AR) coefficients were employed for the description of the post-R segment. This approach was chosen as an alternative ECG representation technique to that of segmenting the ECG cycle by extraction of ECG features.
- Data analysis on the available dataset highlighted the existence of high inter-patient variability. This suggested that the performance of a monitoring system would be boosted by allowing customisation to the specific patient to be monitored.

Significant differences in ECG behaviour were also observed among different night-recordings of the same patient (day-to-day variability) which suggests that a robust monitoring system should also be made adaptive to ECG changes as time elapsed. Hypothesis testing (Student's t-test) on all ECG features extracted, proposed that different features may be robust predictors of hypoglycaemia among different patients. The occurrence of hypoglycaemia may not be sufficiently described using the same features on all patients.

- Investigation of the ECG and glucose profiles also indicated that the ECG responses to hypoglycaemia are expressed in the form of transient events and hence, the incorporation of a temporal dimension in a classification system would be essential for robust detection of the condition.
- Multi-Layer Perceptron (MLP) Artificial Neural Networks (ANN) were assessed for the classification of extracted ECG features. Both approaches of producing global classifiers, to be used on large groups of patients, and also producing classifiers customised for individual patients were followed. Patient-oriented customisation yielded a significant improvement in performance.
- Statistical classifiers, namely the k-Nearest Neighbour (kNN) and Linear Discriminant Analysis (LDA) were also assessed regarding their ability to classify the extracted ECG features.
- A Knowledge-Based System (KBS) was designed for monitoring and classifying offline data from diabetic patients in a manner that simulated an online patient-monitoring scenario. Two versions of the KBS were produced, namely a version based on Crisp Set Logic and a version based on Fuzzy Logic. The two systems were using the same rule-base. The Fuzzy Inference System (FIS) was introduced because of its ability to provide a degree of certainty when raising hypoglycaemic alerts. Employing a knowledge-based approach yielded the highest performance among all the classification techniques considered. This system was able to accurately monitor patients that were consistent with the initial hypothesis i.e. exhibiting QT prolongation and T wave flattening during hypoglycaemia. A significant difference of the KBS compared to the neural and statistical classifiers was that it incorporated temporal information, while the latter were performing static[2] pattern classification.

[2] i.e. no time-stamps

- A methodology was proposed for making the monitoring system adaptive as time elapsed. Adaptivity in our case meant that the definitions of QT prolongation and T wave flattening (and also the definitions of normal QT and T wave) were changing as time elapsed.
- By means of the performance of the KBS, it was demonstrated that two ECG features were sufficient for detecting the condition on those patients who manifested both QT prolongation and T wave flattening under hypoglycaemia. More features could be used in the future for achieving more robust monitoring across the patient population.
- Development of the KBS also contributed in formulating the vague knowledge of the basic ECG changes under hypoglycaemia in the form of rules of natural language. This was informative for medical researchers and provided feedback to the clinical experts who formulated the initial hypothesis and contributed the initial guidelines for the knowledge-base.

To the best of our knowledge, this work lies among the first studies attempting to detect the onset of spontaneous nocturnal hypoglycaemia indirectly through analysis of the patient's ECG. Novel datasets are used that consist of both the ECG traces and the accompanying glucose data. Online ECG databases, available for research purposes in the World Wide Web contain ECG data only, without glucose information. A number of researchers have been carrying out studies [Marques 1994, 1997], [Ireland 1998, 2000], [Harris 2000], [Lee 2003], [Robinson 2004], with specialised ECG-glucose datasets but the analysis was carried out from a clinical viewpoint.

A lot of engineering studies have been carried out focusing on ECG signal processing and classification for cardiac diagnostics and arrhythmia detection[3] but this is among the first ones looking at the quantification of ECG changes related to hypoglycaemia-induced arrhythmias for the detection of the symptomatic status of hypoglycaemia.

1.5 Structure of the thesis

The structure of the thesis, following this first chapter, is outlined below:

[3] A small sample of such studies is: [Xiong 1983], [Shibahara 1984], [Yang 1994], [Kundu, Nasipuri 1994], [Kennedy 1997], [Hedén 1997], [Simon 1997], [Hu 1997], [Al-Fahoum 1999], [Lagerholm 2000], [Acharya 2004], [Kundu 2000].

Chapter 2 presents the necessary theoretical background to support this work. Firstly it covers ECG signal processing (including pre-processing and noise reduction) and then discusses standard and abnormal ECG morphologies. Following these, ECG feature extraction and significant point detection is presented. Autoregressive modelling is also presented. Chapter 2 also provides theoretical background on the classification techniques employed. This includes brief Neural Network theory, theory on Linear Discriminant Analysis and the k-Nearest Neighbour classifier, background on Knowledge Based Systems and finally Fuzzy Logic theory.

Chapter 3 presents the biomedical resources used in this research. This mainly includes the patient data on which the studies were based. Moreover, the data acquisition equipment and data acquisition protocol are presented.

Chapter 4 presents the overall methodology used in this work followed by the feature extraction results. The ECG-specific software tools used for algorithmic development are presented. Moreover the way of combining the various techniques for producing the overall system for hypoglycaemia detection is given. The remaining of the chapter presents the results from the feature extraction of ECG signals. A number of automatic algorithms designed for ECG annotation are presented. A subset of these algorithms is utilised in a comparative study of four geometric methods for T end annotation. The ECG features produced, including two novel ones introduced, are presented and assessed. AR modelling, as a means of ECG representation is also included.

Chapter 5 presents the static pattern classification results produced. Classification is based on either individual ECG features or AR coefficients. ANN, LDA and kNN results from a number of approaches are presented. Both the approaches of producing global and patient-oriented classifiers are outlined.

Chapter 6 presents a knowledge-based monitoring system for hypoglycaemia detection during the night. The architecture of the knowledge based system used, including its rule base is presented. Results are produced both when applying crisp and fuzzy logic in the system.

Chapter 7 summarises and concludes the thesis. The contribution of this study is presented and the unsolved problems are discussed. Also some long-term research challenges and unsolved problems that impeded this study are discussed. Finally the way forward is considered and recommendations for further work are given.

Theoretical Background and Review of the Literature

2.0 Introduction

In this chapter the theoretical background relevant to the techniques applied in this study, is provided. It presents detailed information on Electrocardiography, extending the discussion of Chapter 1, and also presents ECG signal processing including noise reduction, automatic ECG annotation and ECG features. Next, a review of relevant biomedical equipment used for ECG and glucose monitoring is provided. Moreover, theory on intelligent systems is presented, including neural networks and knowledge-based systems. Statistical classifiers are also discussed.

2.1 Noise considerations related to ECG signals

Electrocardiogram recordings suffer greatly from various sources of noise. One of these is 50 Hz powerline interference, and its harmonics, together with electronic noise inherent in all electronic devices. Muscle noise (the signal known as ElectroMyoGram (EMG)) is another noise source. It is introduced by the movement of the skeletal muscles of the chest wall during respiration but may also be due to movement of other peripheral muscles [Marques 1994].

The first type of noise mentioned can be reduced by performing a good skin preparation before placing the electrodes, by using shielded cables and by avoiding proximity to the sources of 50 Hz, e.g. fluorescent lights, monitors etc. Good amplifier design with modern electronic components is usually adequate for limiting inherent noise. Remaining noise due to powerline interference can be reduced using a notch filter set at 50 Hz.

2.2 The Signal Averaged ElectroCardioGram (SAECG)

Even when being cautious to perform good data acquisition, noise problems will still exist, mainly due to the EMG signal. Averaging is often necessary and is used to further reduce the noise levels. Successive ECG records are aligned and then averaged to produce the Signal Averaged ECG (SAECG). The averaging process attenuates the stochastic components and amplifies the deterministic components in the signal. As each beat is added, the noise is reduced in the signal-averaged recording. This is the primary reason for using the signal-averaging method because very low-level signals are usually masked in noise and standard ECG techniques are not adequate for recording these very low-level signals.

Theoretically the square root of the number of beats averaged will be the factor by which the noise is decreased [Berbari 2000]. If 100 beats are averaged, then the noise will be reduced by a factor of 10. In practice this is only approximate because the characteristics of the noise may vary over time. Although averaging helps to reduce noise, one of the disadvantages of averaging successive ECG records is that distortion may be introduced and useful information lost. This happens because the noise-free ECG is different from cycle to cycle since it is a dynamic signal. The signal-averaging process would not introduce distortion in the ideal case where the ECG signal would be stationary and contaminated by random noise.

The way signal averaging is carried out is demonstrated in Figure 2.1. The diagram represents each ECG cycle by a piece-wise linear representation (only the QRS is represented and the P & T waves are not drawn). An exaggerated noise component is drawn on each ECG cycle. Successive ECG cycles are collected and aligned and then they are added and averaged. The averaged signal at the output of the summation operator, in Figure 2.1, is drawn with a smaller noise component in the post-QRS region to demonstrate the reduction in noise levels.

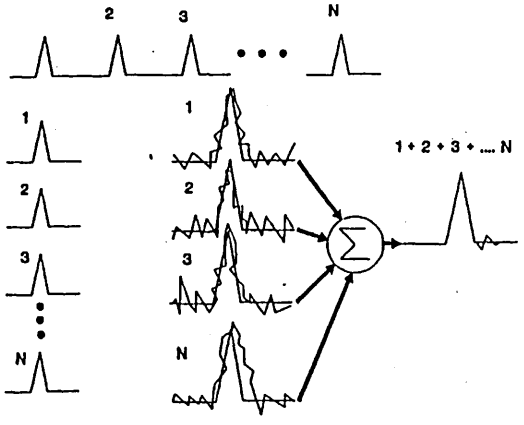


Figure 2.1: Demonstration of signal averaging process [Berbari 2000]

In order to carry out the above alignment and averaging process, the QRS complex of each cycle must be detected. Detecting QRS complexes is generally easy since it is the component of the ECG cycle with the highest amplitude. But simply using an amplitude criterion to detect the QRS may lead to false detections of premature ventricular contractions (PVC) or T waves with high amplitude, as QRS complexes. This introduces the need for QRS selection in order to discard false detections of premature ventricular contractions or T waves with high amplitude. An example of a PVC beat occurring between two Normal Sinus Beats (NSB) is shown in Figure 2.2 (top graph).

Alignment is done using a fiducial point on the QRS complex upslope (Figure 2.2 bottom graph). The fiducial point is a detection point that acts as a time reference in the alignment process. The shape of the QRS is used to determine the fiducial point. A more simplified approach would be to use the QRS amplitude or the magnitude of the first derivative at the R peak [Berbari 2000] but these do not contribute to a highly accurate alignment process since other components can be mistaken for the QRS as mentioned earlier. Using the shape of the QRS is superior because it eliminates all other ECG components (PVCs, high T waves), noisy beats and motion artefacts [Berbari 2000]. Moreover it allows fine alignment of the QRS complexes for the averaging process.

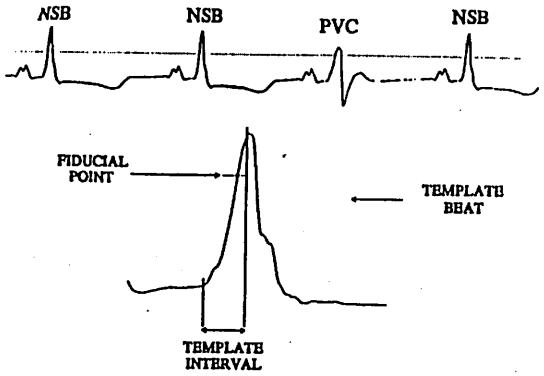


Figure 2.2: PVC between two normal beats (top graph) and fiducial point (bottom graph) [Berbari 2000]

A number of assumptions are made and a number of conditions need to be met for a successful averaging process [Marques 1994]:

1. The ECG signal and the noise should be independent.
2. The noise probability distribution function must remain constant throughout the recording period, i.e. it must exhibit stationarity.
3. The contaminating noise must have a Gaussian distribution.
4. The signal of interest must be periodic and associated to a timing reference.

The averaging technique is effective in removing EMG noise since the respiratory and cardiac cycles are not synchronous [Marques 1994]. Electronic noise is random and may be reduced or eliminated by employing narrow bandwidths. However, the 50 Hz noise may sometimes not be completely eliminated, since its probability density function is not Gaussian and therefore will never average to zero [Marques 1994]. Use of a notch filter before the averaging process may solve this problem.

2.3 ECG components and the morphology of the ECG

As mentioned in Chapter 1, a typical ECG cycle consists of three main components namely the P wave, the QRS complex and the T wave. Besides the above three, more components can be identified on a typical cycle. A full list of them (including the above three) is given below:

1. P wave which corresponds to the atrial depolarisation of the myocardium.
2. QRS complex defined by the following points on the ECG: Q, R and S. The QRS complex corresponds to the ventricular depolarisation of the myocardium. The R peak is the peak with the highest amplitude in a normal ECG cycle. The time duration between two successive R peaks denotes the instantaneous heart rate.
3. ST segment, which is a segment of electrical inactivity post the QRS.
4. J point (starting point of the ST segment).
5. J80 point (ending point of the ST segment, located at 80msec interval from the J point).
6. R104 point located at 104 msec interval from the R peak.
7. T wave which corresponds to the ventricular repolarization of the myocardium.
8. U wave which is not part of a typical ECG cycle and is only present under certain circumstances. Presence of U waves has been reported under hypoglycaemia [Lazarra 1992], [Ireland 2000].

2.3.1 Types of T wave morphology

A few different types of T wave morphology can be encountered and will be presented in this section. Different morphologies may have different clinical significance. Robust feature extraction algorithms must be able to annotate records even when non-standard morphologies occur. The shape of a normal T wave is given in Figure 2.3. Under certain circumstances, inverted T waves occur. They are similar in shape to the normal T waves with the difference that they are inverted, i.e. the orientation of the T wave is opposite to that of the P wave and QRS complex. ECG cycles containing inverted T waves are

presented in Figure 2.4. Finally biphasic T waves can also be encountered. These consist of a positive and a negative peak in succession. Either the positive (characterised as biphasic +ve/–ve) or the negative wave (biphasic –ve/+ve) can occur first. The combination of these two waveforms produces an undulation as seen in Figure 2.5.

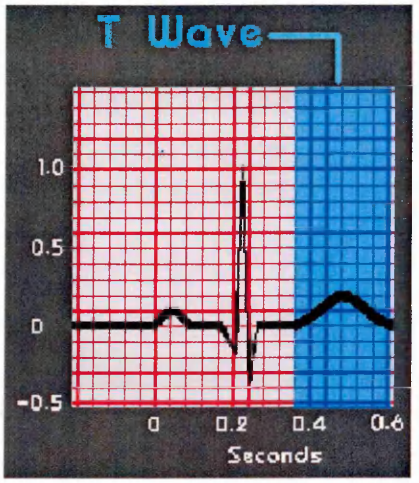


Figure 2.3: Normal T wave

Apart from the above morphologies, a U wave is sometimes present after the T wave as mentioned briefly in Chapter 1. This wave may exist as a separate entity or may be fused with the T wave. Detection and characterisation of this wave is desirable in ECG processing. An ECG waveform exhibiting an exaggerated U wave is presented in Figure 2.6.

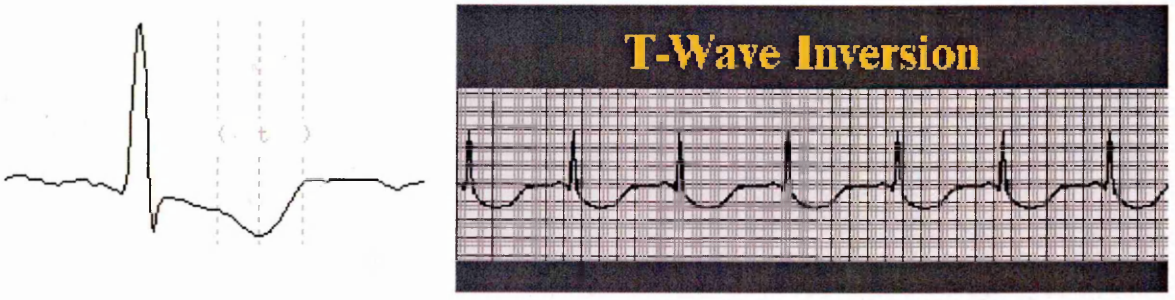


Figure 2.4: A single inverted T wave (left) and a train of inverted T waves (right)

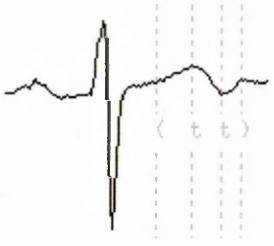


Figure 2.5: Biphasic (+ve/-ve) T wave

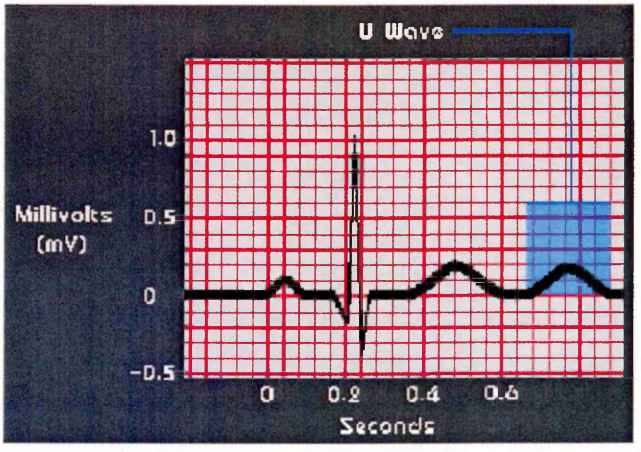


Figure 2.6: U wave present after the T wave

2.3.2 ECG characteristic points

Besides the previously discussed ECG components, a number of ECG characteristic points (sometimes referred to as "ECG significant points") need to be detected on each cycle. Such points are mainly the onsets, offsets and peaks of the three main waves (P, QRS, T). These characteristic points are necessary for the process of extracting ECG features as it will be discussed later. The characteristic points that need to be detected in the context of this research are: the R peak, the QRS onset and offset, the T peak and the T wave onset and offset. Moreover, detection of the peak, onset and offset of the U wave, if present, are useful. This type of wave is not always present or if it is, it can be fused with the T wave so its detection is not always possible.

Detecting the end of the T wave has, for many years, been a big problem in the research community. Under noise conditions, existence of U waves or other abnormalities, the end of the T wave can be very ambiguous and even manual marking by clinical experts can be difficult.

Manual vs Automating ECG Marking. The marking of time intervals on ECG cycles is done manually by cardiologists. ECGs on paper were used in the past, but nowadays computer screens are mostly used where the cardiologists set cursors to mark certain features and intervals. (ECGs were recorded in the past using a heated stylus that was marking strips of propagating paper.) Manual measurement is essential but not sufficient. Some of the problems associated with it are that: it is labour intensive, it introduces the problem of inter-observer variability and it does not give systematic results compared to automatic methods. Automatic algorithms have the following advantages: they produce systematic measurements, they need less resources (time, researchers) to be carried out, and they allow the design and use of automatic, online monitoring systems where the presence of a medical expert is not possible.

2.3.3 Automatic detection of the ECG characteristic points

A number of algorithms for detection of the ECG characteristic points will be presented below. Some of these focus only on detection of a specific component (mainly the R and T peaks) while others can be applied for detection of more than one component.

Although the end of the T wave is the last point in the sequence of significant points, it will be reported first since it is among the most important points to mark for detection of

hypoglycaemia and a challenge to detect especially in the presence of noise or other abnormalities. A review of a number of algorithms for detecting the end of the T wave will be discussed in the following section.

Automatic detection of the end of the T wave

Annotating the end of the T wave on an ECG cycle is necessary for measuring time intervals such as the QT. This can be done both manually and automatically. The manual method is the current "gold standard". A number of automatic methods also exist. Marking the end of the T wave is not always an easy task. Very often the point where the T wave ends is not very clear. The presence of a fused U wave is one of the cases where the detection of the T wave end is difficult.

Probably the simplest and most obvious approach for marking the end of the T wave is the one that uses the cross-section of the T wave downslope with the isoelectric line (flat baseline) to mark T_{end}. The weakness of this method is that in some cases, there is no crossing of the T wave downslope with the isoelectric line. Various methods for marking the end of T have been investigated in a number of papers. The most important algorithms are covered in this section:

1. Tangent Method (or MSI: **M**aximum **S**lope **I**ntercept) [Ireland 1998, 2000], [McLaughlin 1995, 1996]

   This method finds the point of the T wave downslope having the steepest tangent and marks the end of the T wave at the point where the steepest tangent line meets the isoelectric line. The characteristic of this method is that it relies only on a single point on the ECG trace for deciding where the T ends.

2. Peak Slope Intercept (PSI) [McLaughlin 1995, 1996]

   This method uses the peak of the T wave and the point of the T wave having the steepest tangent to define a straight line. The intersection of this line with the isoelectric line marks the end of the T wave. The way this method works is illustrated in Figure 2.7. The peak of the T wave and the steepest tangent point are marked with circles on the

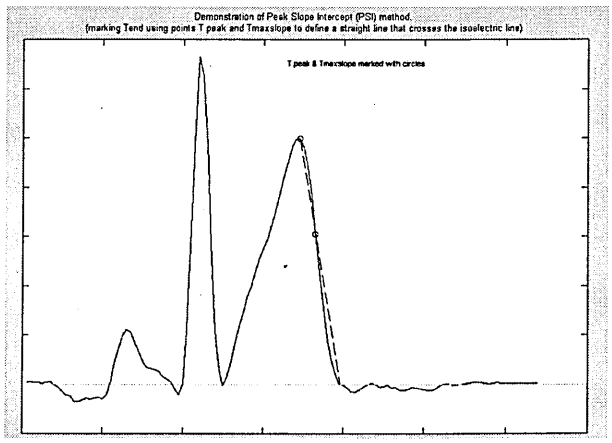


Figure 2.7: Demonstration of the PSI method

graph.

3. Fitting Method [McLaughlin 1995, 1996]

   The third method fits a 1st order polynomial on the downslope of the T wave. Again the intersection of the fitted straight line, with the isoelectric line, marks the end of the T wave. The line fitted by this method on the T downslope can be seen in Figure 2.8.

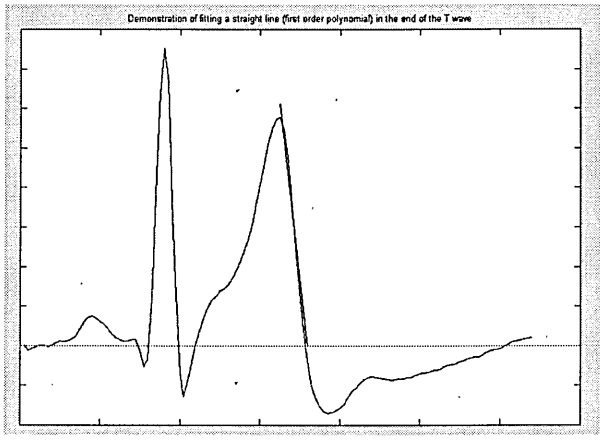


Figure 2.8: Demonstration of the fitting method

4. Threshold method (TH) [McLaughlin 1995, 1996]

   This method marks the end of the T wave by using a threshold level set at a fraction (e.g. 0.1) of the T wave amplitude. The intersection of the threshold level with the T wave marks the T end. This concept is illustrated in Figure 2.9 (top graph) where the threshold level is depicted by a dashed horizontal line.

5. Differential Threshold method (DTH) [McLaughlin 1995, 1996]

   The Differential Threshold method works similarly to the Threshold method with the difference that the signal used is the 1st derivative of the ECG trace instead of the trace itself. In both algorithms, threshold crossing points are determined using a left to right scan of the data from the T peak. Moreover the T peak and the threshold level are calculated relative to the isoelectric level. Figure 2.9 (bottom graph) illustrates the 1st derivative of the ECG trace, the threshold level chosen and the location where T end is identified.

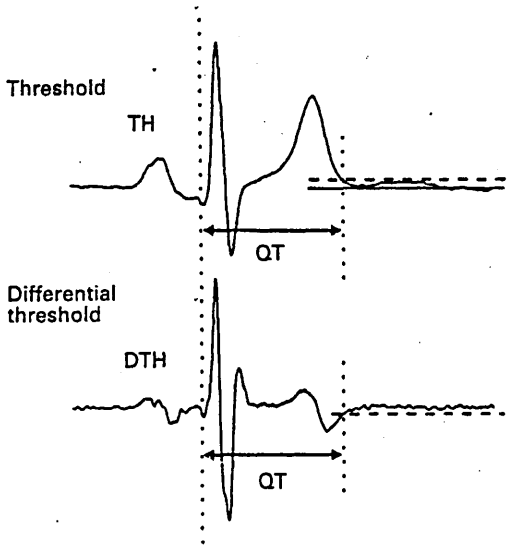


Figure 2.9: Illustration of TH (top) and DTH (bottom) [McLaughlin 1995, 1996]

6. Wings algorithm

In 1999 Daskalov and Christov [Daskalov 1999] presented their algorithm named as "Wings algorithm" for detection of the end of the T wave. The algorithm is also able to detect the T peak. It works by calculating the "wings" function obtained at each successive sample of the search interval of the current ECG cycle. The "wings" function is given by the product of two adjacent segments forming "wings" each 40 msec in length. The wings are calculated in sequence at each signal sample of the search interval. The search interval starts at the isoelectric point at the end of QRS[4] and reaches the end of the record. The minimum value of the wings function corresponds to the T peak. The algorithm is also capable of detecting biphasic T waves. The T wave end is chosen as the point having the smallest angle between wings of 10 msec length in the post-T_{peak} region.

7. Algorithm by Vila et al

   In 2000 Vila et al [Vila 2000] introduced a new TU complex detection and characterisation algorithm. Their work builds on previous research on modelling T waves using Action Potentials (APs) [Wohlfart 1987, Malik 1989, Padrini 1995]. According to the authors the most complete approach for such modelling was proposed by Padrini [Padrini 1995] which is capable of modelling combinations of T and U waves. The authors extend on Pardrini's work by employing a two-stage process for TU wave detection and characterisation. The first stage is modelling using Action Potentials differences and the second stage involves annotations on the modelled signal, instead of using the original one, using classic threshold detection algorithms such as those mentioned above.

8. Neural Network algorithm by Bystricky and Safer

   In 2002 Bystricky and Safer [Bystricky 2002] developed a neural network algorithm to mark the T wave end. It is based on a 2-layer Multilayer Perceptron (MLP) trained on the Physionet QT database. As far as we are aware this is the only neural-network-based algorithm for T end annotation.

Automatic Detection of the R peak

A number of R peak detection algorithms have been developed. Among the established algorithms, the one by Pan and Tompkins [Pan 1985] with some modifications [Laguna 1990] was also used for this research work. Other well-known R detection algorithms are the Balda algorithm [Balda 1977] which is included in a commercial ECG analysis

[4] The above isoelectric point is defined as the rightmost post-QRS sample satisfying a linearity and a slope condition defined in their paper [Daskalov 99].

system by HP, and the ARISTOTLE algorithm [Moody 1982]. Li et al [Li 1995] presented a wavelet algorithm for detecting the QRS complex while the algorithm is also able to detect other components such as P and T. In 2002 Saxena et al [Saxena 2002] developed a new wavelet for detection of the QRS complex.

At the early stages of this research program a basic R detection algorithm was developed and used as part of the feature extraction process.

"ecgpuwave" - Automatic Threshold Based Detector (TD) of Waveform limits

The "ecgpuwave" algorithm was designed by Laguna et al [Laguna, Thakor 1990], [Laguna 1994] and took its name from the name of the command that invokes it in a UNIX shell. It is able to detect the onsets and ends of the P, QRS and T waves. It can work on all leads of the standard 12-lead ECG and also in all 3 leads of the orthogonal ECG. It also classifies QRS complexes as normal or abnormal and T waves as normal, inverted, monophasic and biphasic. The QRS detection is carried out using the Pan and Tompkins algorithm [Pan 1985] with modifications by Laguna as mentioned earlier.

The detection of the ECG significant points is a two-stage process. Firstly the algorithm uses a differentiated and low-pass filtered version of the ECG signal to detect each beat and then the waveform boundaries are located in each lead. This algorithm has been written in FORTRAN, is open-source and is distributed under the GNU General Public License (GPL) [URL 2]. It is available as part of the Physio toolkit from Physionet.

2.4 ECG features

A number of ECG features can be extracted in the time domain, some of them being related to or based on the ECG components that were described above. The QT interval, the time interval from the onset of the QRS complex to the end of the T wave, is probably the most important parameter for indicating the onset of hypoglycaemia. It describes the duration of ventricular depolarization and repolarization. Several studies on the effect of hypoglycaemia on the QT interval have been carried out [Harris 2000, Ireland 1998, 2000]. An illustration of the QT interval is given in Figure 2.10. QT intervals longer than 440 msec in lead V_5 are considered abnormal (Long QT syndrome) [Benhorin 1990].

Other useful intervals are: the QRS duration, which describes the duration of ventricular depolarization, the PR interval, the RT interval, the ST segment etc. Although the QT interval can be an indicator of spontaneous hypoglycaemia, more sophisticated parameters will have to be extracted from the ECG for better detection of this medical condition. In the time domain the parameters that are proposed are describing the morphology and shape of the T wave. Some of the parameters that are proposed, besides the QT interval, are the following:

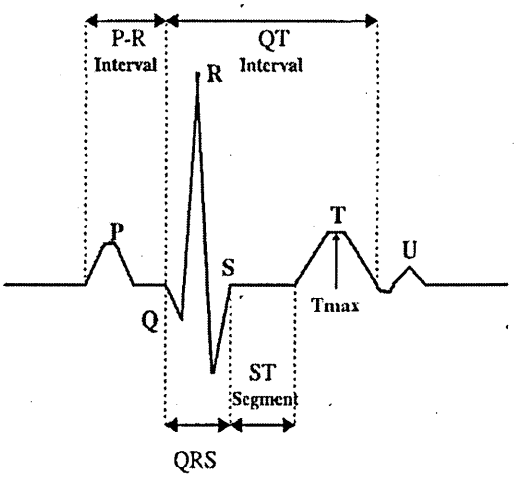


Figure 2.10: ECG features

- T wave amplitude, also shown in Figure 2.10 above as Tmax.
- Baseline T wave amplitude (at the beginning of measuring period) to current T wave amplitude.
- Area under the T wave.
- Symmetry of the T wave. The 3rd and 4th moments (skewness & kurtosis) are proposed for the evaluation of symmetry.
- Presence of U wave following the T wave.
- Ratio of R peak to T wave peak.
- Time to peak T wave amplitude.
- Rate of change of T wave amplitude.
- Baseline drifts between records or within a record.

2.4.1 QT and other time-interval correction

The QT interval may vary because of changes in the RR interval (the instantaneous heart rate). We are only interested in QT variation due to hypoglycaemia or other arrhythmias and not in variation due to changes in heart rate. To cater for this, many investigators normalize the QT interval to make it independent of HR. The most commonly used correction formula is Bazett's [Bazett 1920] formula which is given below:

$$QTc = QT / \sqrt{RR} \qquad \text{eq}^{n} (2.1)$$

According to this formula the heart-rate-corrected interval (QTc) is derived by dividing QT by the square root of the instantaneous HR. A few other methods of QT correction exist [Puddu 1988, Rautaharju 1993, Ahnve 1985]. Correction for heart rate is usually applied on the QT interval but it could also be used for other time intervals such as the RT, the ST, etc as it will be seen later.

2.4.2 Benhorin's features for detection of Long QT syndrome

In 1990 Benhorin et al [Benhorin 1990] had introduced seven new ECG features to identify patients with known long QT syndrome (QT>440msec in lead V_5). The features quantify various components of the post-QRS ECG segment on all leads of the standard ECG. The features were used to distinguish between normal subjects and Long QT Syndrome sufferers. Although this study was very significant in introducing novel features for classification between patients and normal subjects, it did not achieve detection of the symptomatic status of Long QT Syndrome sufferers, i.e. detection of the onset of QT prolongation. The latter problem needs to be solved in order to achieve monitoring of patients.

The features (with feature names given in brackets) are:

1) Early duration (SoTmc): HR corrected S wave offset (So) to T wave absolute maximal amplitude (Tm) interval.
2) Late duration (TmTo): Tm to T wave offset (To) interval.
3) Rate (t.A25-75): the time to accumulate the mid-50% of total absolute repolarisation area from its 25% to its 75% value.
4) Total Area (Atot): Total absolute repolarisation area from So to the end of repolarisation signal or to the next P onset (whichever occurred first).
5) Symmetry Ratio (SR): T wave area symmetry ratio; the ratio between the integrated area over SoTm and TmTo intervals, SR= SoTm/TmTo eq^n (2.2)
6) Late phenomena (%A@To): % of Total area (Atot) accumulated at To. This is the ratio of the repolarisation area from So to To upon the repolarisation area from So to the end of the ECG cycle: %A@To = 100 * (Area from So to To) / Atot eq^n (2.3)
7) Heterogeneity (SoTm_sd): the standard deviation of SoTm interval in the precordial leads. This feature describes the dispersion of the SoTm feature among the precordial leads.

All variables were measured in lead V_5 except SoTm_sd which was calculated from measurements in precordial leads 4 to 6. Heart rate correction was carried out using Bazett's formula.

2.5 Review of Glucose and ECG Monitoring Equipment

This section includes the review of the literature on the biomedical equipment related to this research. This includes glucose monitoring devices and ECG acquisition and analysis equipment.

2.5.1 Hypoglycaemia Detection and Glucose Monitoring Equipment

Two types of devices are discussed in this section: hypoglycaemia detection devices and glucose monitors. Hypoglycaemia detection devices are designed to detect the onset of hypoglycaemia but cannot necessarily measure the glucose levels of a patient. On the other hand, glucose monitors can produce glucose measurements and, if suitable alarms are programmed into them, they can also be used to detect the onset of hypoglycaemia and hyperglycaemia.

Hypoglycaemia detection devices

One of the approaches used in hypoglycaemia detection devices is to utilise the peripheral physiological responses to falling blood glucose (e.g. sweating), and then develop methods of measurement and the software to recognize these patterns. "Sleep Sentry" [Hansen 1993, URL 3] was an early commercial device that was detecting such physiological responses by monitoring skin conductance and temperature as an index of diaphoresis and skin blood flow but was not widely established because of a high number of false alarms. Although not widely established, Sleep Sentry remains in the market [URL 3].

Hastings et al [Hastings 1998] also presented a hypoglycaemia detector that was using peripheral physiological responses to falling blood glucose. This detector was presented as a prototype software engine for a hypoglycaemia detector. It was using skin conductance and heart rate as inputs while two additional inputs (snoring and ECG) were proposed for future versions of the system.

Glucose monitoring devices

Early measurement of glucose was achieved using urine-reactive strips. Currently, urine glucose testing is not recommended as the sole method for monitoring blood glucose [Kirk 1998]. The desire to improve glycaemic control has led to the use of blood-reactive strips to measure capillary blood glucose levels. Initially, these blood glucose strips were interpreted by colorimetry through visual readings and optionally by reflectance photometry [ADA 1995]. The first blood glucose monitor using reflectance photometry was the Ames Reflectance Meter (ARM) [Bernstein 2002]. The first patent for this meter was issued in September 1971. An early ARM prototype is depicted in Figure 2.11.

Figure 2.11: Early ARM prototype

All current home glucose monitoring systems use either reflectance photometry (first-generation systems) or an electrochemical process (second-generation systems) [Fleming 1994]. In both types of systems, an enzyme that catalyzes the glucose reaction within the test strip is used.

The introduction of first-generation blood-reactive meters led to more accurate readings while the devices became easier to use. Some disadvantages that later devices were challenged to solve are the long processing time for glucose measurement, the

requirement of a large drop of blood, the large size of the device and the limited memory features [Kirk 1998], [Foster 1999].

Second-generation blood glucose monitors measure an electrical charge generated by the glucose-reagent reaction. The electrochemical principle employed can be either amperometry or colorimetry. Amperometry meters did not solve the problem of using a small blood sample, since only a small portion of the blood sample is utilised. However, colorimetry meters solved this problem by utilising all of the sample glucose and converting it to an electrochemical charge that is measured. The latter approach is also insensitive to temperature and haematocrit variations while amperometry devices were suffering from such variations. Finally, an invaluable contribution by 2nd-generation colorimetry monitors is the possibility of using alternative sites (e.g., arm or thigh) to obtain blood samples. At these sites, capillaries and nerve endings are less numerous; therefore, a more sensitive measurement technology was necessary to provide virtually painless blood glucose testing [Mehta 2002].

Interstitial fluid sampling. A later approach to glucose sensing is to sample the interstitial fluid (ISF), i.e. the fluid that exists among the tissue cells. A few devices have been produced that follow this approach. MiniMed Inc [URL 4] has produced the MiniMed Continuous Glucose Monitoring System (CGMS) [Gross 2000, Steil 2000] which is an invasive sensor that measures glucose levels by sampling the glucose in the interstitial fluid. A probe is inserted in the subcutaneous tissue of the tummy area and measures glucose in the tissue. This sensor is used for data acquisition in this research and will be presented in greater detail in Chapter 3.

The limitation of the MiniMed CGMS system is that it employs an invasive approach which could, depending on the size of the probe used, be a degrading factor in the quality of life of the patient using it. Another device, the GlucoWatch [URL 5] by Cygnus Inc [URL 6] achieves non-invasive monitoring. It monitors glucose levels non-invasively by reverse iontophoresis (electroosmosis) [Rao 1995, Tamada 1995]. It is worn like a wristwatch and applies a small electric current to the skin, which is used to collect glucose molecules that exist between cells, through the skin in gel disks inside the watch. The glucose levels are then translated into electrical signals.

The next step in blood glucose monitoring is the development of non-invasive (3rd-generation) meters, in which the sample is obtained without direct interaction with body tissues [Mehta 2002]. These future meters will perform measurement remotely using various characteristics (e.g., spectral, optical, thermal, electromagnetic). The most promising prototypes use radiation technologies i.e. Near Infra-Red (NIR) spectroscopy, Far Infra-Red (FIR) spectroscopy, radiowave impedance, and optical rotation of polarized light.

2.5.2 ECG Acquisition and Analysis Equipment

Numerous devices have been developed for acquisition, analysis and monitoring of ECG signals. Modern electrocardiographs can capture ECG data from all 12 leads of the standard ECG. They perform appropriate pre-processing to remove noise and artefacts from various sources. Many of them also offer interpretation of ECG traces and hence can diagnose various cardiac rhythms or arrhythmias. A selection of the latest devices is presented in this section.

ECG monitoring devices could be categorised as trolley-based, computer-based and portable devices. Seca is among the companies producing trolley-based ECG monitoring equipment. It produces 3 ECG monitors (CT3000i, CT6i, CT8000P) [URL 7]. The Seca CT6i, incorporating interference filters and being capable of producing 12 Channel interpretive ECG, is depicted in Figure 2.12 [URL 7]. CARDIOVIT and GE Medical Systems are also manufacturing a number of trolley-based electrocardiographs.

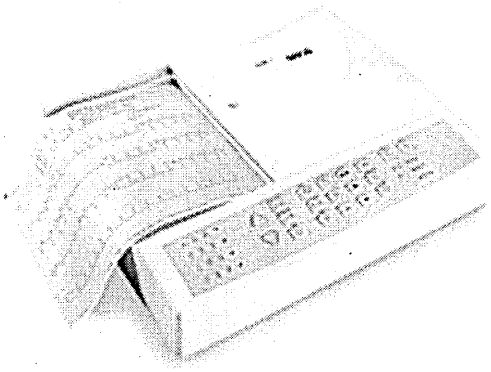


Figure 2.12: Seca CT6i ECG monitor

Computer-based ECG equipment as the name implies is based on a personal computer (PC). The recorded ECG is transferred, in real-time, to the PC via an appropriate interface and the relevant software is used to display, store and process the ECG traces.

An example of an interfacing link is the CardioView ECG-PC Link from Micromedical Industries which connects to the PC Serial Port. It is depicted in Figure 2.13 next to a PC.

Advancing handheld-computer technology has allowed the realisation of portable ECG monitors based on palmtop computers. Such an example is the PocketView ECG [URL 8, 9] shown in Figure 2.14. Biolog 3000i (Figure 2.15) is another handheld electrocardiograph, though not based on a palmtop computer. Besides the standard features that modern electrocardiographs posses it is able to display and record instant 12 lead ECG by placing the Biolog directly on to the patient's chest (i.e. no leads involved). MINISCOPE MS-3 [URL 10] is another portable device recording Emergency 1-Channel ECG.

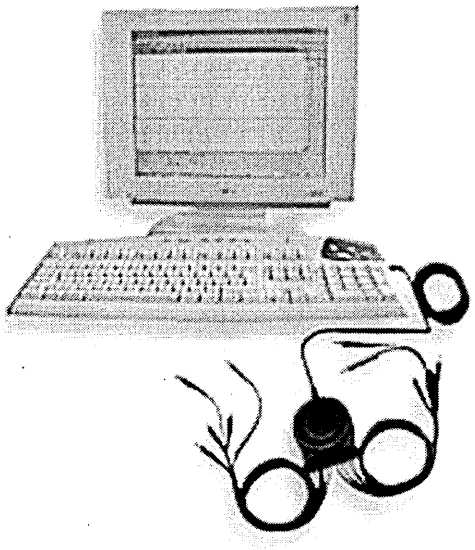


Figure 2.13: CardioView ECG-PC Link

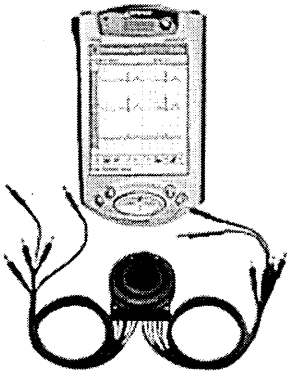


Figure 2.14: PocketView

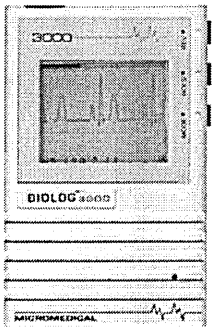


Figure 2.15: Biolog

CardioSoft and CardioView 3000 are software packages that can be run on a PC and can perform collection, analysis, review, and printing of ECGs. Besides the above two more generic packages, some more specialised software packages have been developed by GE Medical Systems, that contain software algorithms for in-depth processing of ECGs. Such packages are: 12SL ECG Analysis Program, Acute Cardiac Ischemia Time-Insensitive Predictive Instrument (ACI-TIPI) and Signal-Averaged High Resolution P-Wave (Phi-Res) Analysis.

2.6 Statistical Classifiers

Statistical classifiers were considered in order to allow comparisons with neural classifiers, presented in later sections. The statistical classifiers used were Linear Discriminant Analysis (LDA) and k-Nearest Neighbour (kNN). LDA works by minimising the Mahalanobis distance [MathWorks Statistics] which is a multivariate measure of the separation of a data set from a point in space. The Mahalanobis distance is a very useful way of determining the "similarity" of a set of values from an "unknown" sample to a set of values measured from a collection of "known" samples. This distance measure was introduced in 1936 by P. C. Mahalanobis, hence the name of this statistic.

The statistical distance or Mahalanobis distance between two points $x = (x_1,..., x_n)^T$ and $y = (y_1,...,y_n)^T$ in the n-dimensional space $\Re^n$ is defined as:

$$d(x, y) = ((\mathbf{x} - \mathbf{y})^T \mathbf{S}^{-1} (\mathbf{x} - \mathbf{y}))^{1/2} \qquad \text{eq}^n (2.4)$$

where **S** represents the within-group covariance matrix, lowercase letters in **bold** denote vectors or matrices and the superscript "T" denotes the transpose operation.

Another measure of distance that can be used for classification is the Euclidean distance.

The Euclidian distance between two points $x = (x_1,..., x_n)^T$ and $y = (y_1,...,y_n)^T$ in the n-dimensional space $\Re^n$ is defined as:

$$d(x, y) = ((x_1 - y_1)^2 + \ldots + (x_n - y_n)^2)^{1/2} = ((\mathbf{x}- \mathbf{y})^T (\mathbf{x}- \mathbf{y}))^{1/2} \qquad \text{eq}^n (2.5)$$

The Euclidean distance is a geometric distance measure as opposed to a statistical distance measure and in certain application domains can be inferior to the Mahalanobis distance when used as a classifier. In such applications the Mahalanobis distance is superior because it takes the distribution of the points into account.

Some of the advantages of the Mahalanobis distance are:

1. It takes into account not only the average value but also the variance and the covariance of the variables measured.
2. It accounts for ranges of acceptability (variance) between variables.
3. It compensates for interactions (covariance) between variables.
4. It is dimensionless.

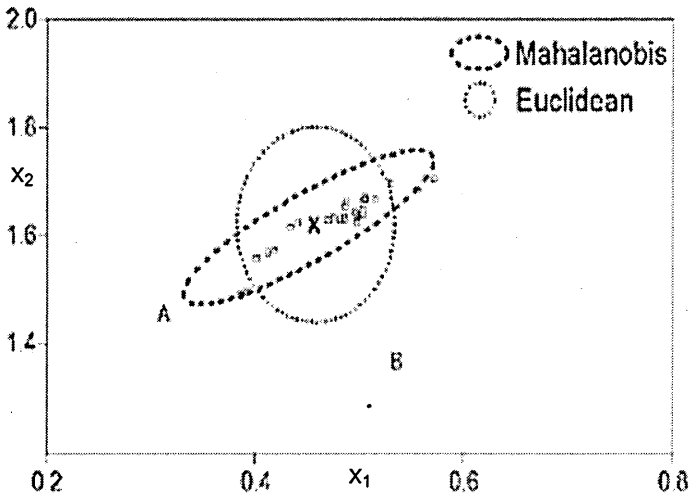


Figure 2.16: Comparison of Mahalanobis vs Euclidean classifier

The difference between the Mahalanobis and the Euclidean distance is illustrated graphically in the 2-dimensional scatter diagram presented in Figure 2.16. It is obvious that the cluster of data points on the scatter diagram is best described using the Mahalanobis distance which considers an elliptical boundary compared to the circle around the mean point X used by the Euclidean distance. Let us consider the example where the data-points with mean X (that occur within the two boundaries) comprise the training set of a classification problem and the points A and B shown in the figure are two unknown samples. The Euclidean distance of points A and B from the mean X is approximately the same. This means that point A is just as likely to be classified as belonging to the group, as point B. However, the Mahalanobis distance will classify point A to be more likely to belong to the known group since it lies on the trend-line of the known group, as described by the major axis of the ellipsoid. The limitation of the Euclidean classifier is that it does not take into account the variability of the values in all dimensions.

The kNN classifier mentioned earlier uses the Euclidean distance metric. Let us consider an n-dimensional space $\Re^n$ and a finite set $S \subset \Re^n$. S is the training set, i.e. the set of known sample points. The class of an unknown point (query point) q, $q \in \Re^n$ is determined according to its neighbouring points in the training set. The parameter k of the kNN classifier can, in theory, be set to any positive integer. It defines the number of neighbouring points from the training set, to be used in determining the class of a query point q. For example, for k=1 the class of q will be determined by the class of the point in the training set having the shortest Euclidean distance from q. Expressing this mathematically, the 1-NN classifier will find the element $p \in S$ such that:

$d(p,q) \leq d(r,q), \forall\, r \in S, r \neq p,$ eq^n (2.6)

where $d(x,y)$ is the Euclidean distance between two points x and y, with $x, y \in \Re^n$.

For k=3, three neighbouring points will be used in determining the class of q and the chosen class will be determined by majority voting. For a two-class classification problem, the operation of the kNN classifier can be illustrated in Figures 2.17 and 2.18. Figure 2.17 depicts the 2-dimensional scatter diagram for the variables x and y. One class is marked by circles and the other by squares.

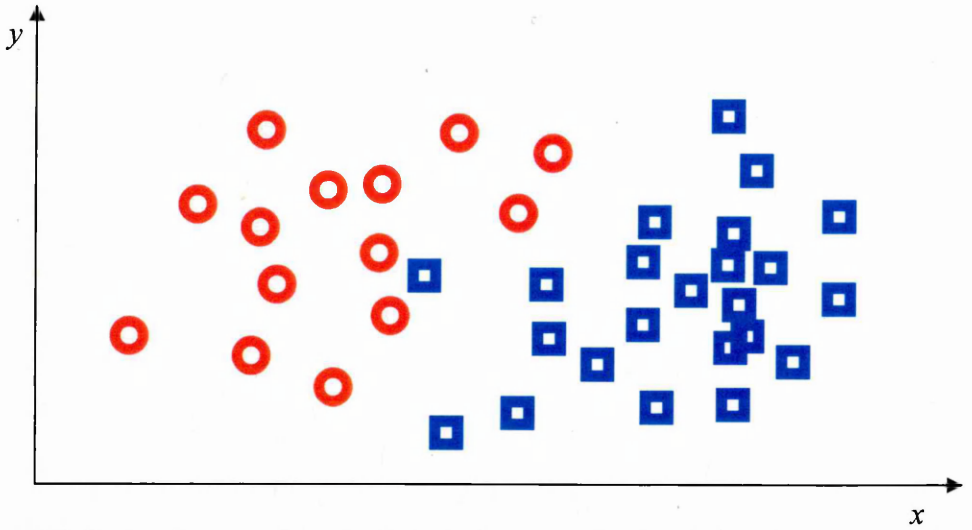


Figure 2.17: Scatter diagram of y vs x with each class marked with either circles or squares [URL 11]

Figure 2.18 depicts a section of the above scatter diagram after zooming in. The figure illustrates how the class of the unknown query point marked with a "*" is determined by the kNN classifier in the cases of k=1 (LHS graph) and k=3 (RHS graph).

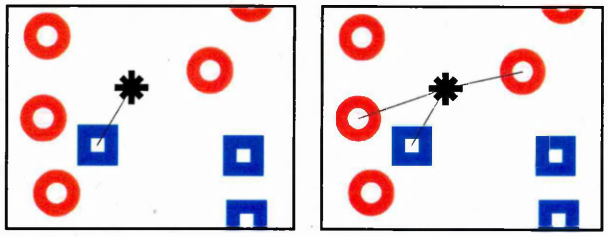


Figure 2.18: 1-NN classification (LHS) and 3-NN classification (RHS) [URL 11]

In the specific example presented above, the classification result is different for the two values of k. Using 1 neighbour, the query point is classed as being a member of the class of squares while using 3 nearest neighbours, it is classed as belonging to the class of circles. Choice of the number of nearest neighbours to be used can be a very important factor in classification performance.

When using an even number of neighbours a tie can occur, i.e. there are equal numbers of neighbours in each class. In order to get an output from the kNN classifier, the tie needs to be broken. One way to break the tie is to choose randomly between the classes

in the tie. Alternatively the tie can be broken by taking in account extra information from the nearest neighbour (i.e. shortest Euclidean distance) to the query point. In that case the nearest neighbour is used to determine the class. Using odd values of k is another way to avoid ties between classes.

An advantage of the kNN classifier is that the decision boundary can be arbitrarily complex. For instance, the classifier can operate not only under circumstances where non-linear decision boundaries exist but also in cases where a class exists within another class etc. A drawback of the kNN classifier is that it does not construct a generalised representation of the learnt classes. Instead, all the training examples are kept in memory and classifying any new point can be very computationally expensive. This is because the Euclidean distance of any new point from all points in the training set must be calculated.

In our case, the computational cost of the kNN was not an issue. The feature vectors fed as inputs and also the length of the datasets used were small and the algorithm executed in short time.

2.7 Artificial Intelligence

Artificial Intelligence (AI) is a branch of computer science concerned with the design and implementation of programs which are capable of emulating human cognitive skills such as problem solving, visual perception and language understanding [Jackson 1990].

Artificial Neural Networks (ANN) and Knowledge-Based Systems (KBS), including Fuzzy Inference Systems (FIS), constitute the Artificial Intelligence techniques considered in this work; for the classification of ECG traces. Theoretical background on the above techniques is given in the following sections.

2.7.1 Artificial Neural Networks (ANN)

The main classifiers considered in this research were artificial neural networks. A number of definitions have been proposed that describe what constitutes a neural network but there is no convergence to a single generally accepted definition. One of the eloquent definitions available is given overleaf.

According to Haykin [Haykin 1994] a neural network can be defined as a massively parallel distributed processor that has a natural propensity for storing experiential knowledge and making available for use. It resembles the brain in two respects:

1. Knowledge is acquired by the network through a learning process.
2. Interneuron connection strengths known as synaptic weights are used to store the knowledge.

Artificial neural networks are also referred to as connectionist models, and their field of study is alternatively termed parallel distributed processing, neurocomputing or artificial neural systems [Rumelhart 1986, Simpson 1990]. The inspiration of artificial neural networks originates from observations of the operation of biological nervous systems. The Central Nervous System (CNS) in the human body consists of the brain and the spinal cord. The Peripheral Nervous System (PNS) serves for communication with the rest of the body. The brain is the central information processing and control unit. The smallest processing element of the brain is the neuron. The human brain consists of an estimated 10^{11} nerve cells, or neurons [Gurney 1997]. A biological neuron is depicted in Figure 2.19. It consists of the cell body, synapses, dendrites and axon. The cell body is the core of the neuron. The dendrites are the inputs to the neuron while the axon is its only output. The output (axon) of a given cell serves as an input to other neurons in the brain. A synapse is the point where the axon from one neuron connects to a dendrite of another. The synapse determines the strength of such a connection (synaptic strength) and varies with time and through the learning process.

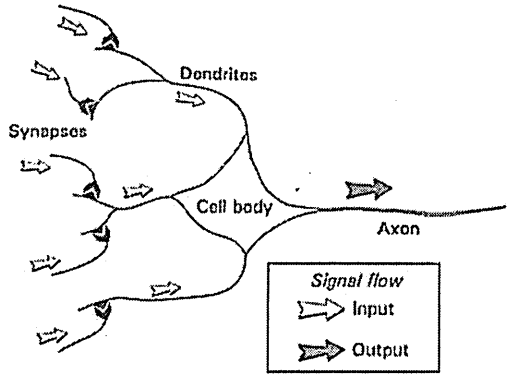


Figure 2.19: A biological neuron [Gurney 1997]

The computational neuron

The biological neuron can be abstracted into the computational neuron as depicted in Figure 2.20.

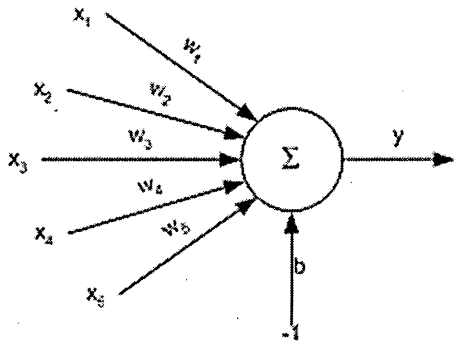


Figure 2.20: Computational neuron

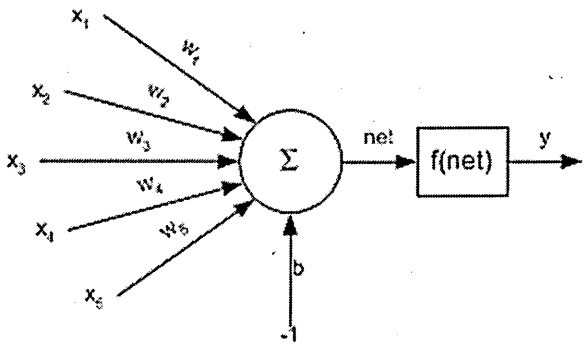


Figure 2.21: Neuron with activation function

The neuron in Figure 2.20 receives 5 inputs (x_1-x_5). Each input is multiplied by a scaling factor, the network weight. The network weights simulate the effect of the synaptic strengths encountered in biological neurons. The weighted versions of the inputs are summed to produce a resultant (net) input to the network. For the neuron depicted in the figure, the output equals the net input. Alternatively a mathematical function can be used to translate the net input to an output value. The function used to perform such a mapping is termed "activation function" and can be seen as the neuron transfer function. A neuron incorporating an activation function (*f(net)*) is depicted in Figure 2.21.

Typical activation functions are: step, linear, saturated linear, logistic sigmoid, hyperbolic tangent, etc. The activation function is also referred to as squashing function because most activation functions "squash" the value of the net input to the interval [0 1] or [-1 1], depending on the span of the function values on the y-axis.

The first formal definition of an artificial neuron was proposed in 1943 by McCulloch and Pitts [McCulloch 1943]. It is known as the McCulloch-Pitts neuron (MCP) or threshold-logic unit (TLU). It receives binary inputs and has a step activation function. The drawback of this architecture is that its weights need to be adjusted by the user, for a given task to be performed. They also have to be known for a given classification problem, i.e. the mapping from the problem domain to the neuron internal parameters must be known. The MCP neuron is not capable of learning by example and adapting itself in an automated manner.

In 1962 Rosenblatt introduced his Perceptron [Rosenblatt 1962], a single-neuron neural network that was able to learn by experience and adjust its weights autonomously. It was similar to a MCP neuron in many ways and its innovative feature was that it was able to learn by experience. The differences between the Perceptron and the MCP neuron are that the Perceptron is not restricted to binary inputs and that it incorporates a learning algorithm. Apart from these they are very similar and the single-neuron network depicted in Figure 2.21 can illustrate both a MCP and a Perceptron providing that the activation function is a step function.

Consider a linearly-separable two-dimensional classification problem as shown in Figure 2.22. A two-class classification problem is termed as linearly-separable if the two classes can be separated by a straight line in 2D, a plane in 3D or an (N-1)-dimensional hyperplane in the case of an N-dimensional classification problem. A Perceptron is always capable of classifying linearly-separable data by finding a linear decision boundary. It is guaranteed to find a solution in finite time if one exists, i.e. if data is linearly-separable (Perceptron Convergence Theorem [Minsky 1969, 1988]).

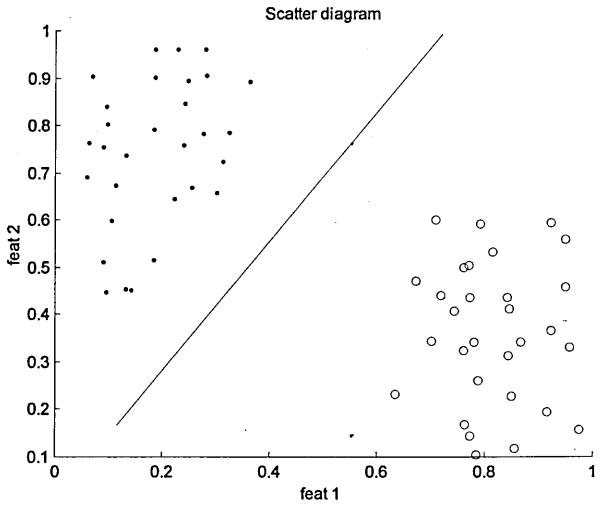


Figure 2.22: 2-class linearly separable classification

The Perceptron learning rule (weight-update rule) is given by the equation:

$w_i(t+1) = w_i(t) + \eta\, e(t)\, x_i(t)$ eq^n (2.7)

where:

- η is the learning rate ($0<\eta\leq 1$)
- $w_i(t)$ is the network weight of the i^{th} input at time instant t
- $e(t)$ is the error defined as: $e(t) = x_i^a(t) - x_i^d(t)$ eq^n (2.8)
  superscripts *a* and *d* stand for actual and desired network outputs respectively. The permitted values for e(t) are {-1,0,1}
- $x_i(t)$ is the input value received at the i^{th} input at time instant t

The MultiLayer-Perceptron

The MultiLayer-Perceptron (MLP) is a neural network architecture consisting of multiple layers of simple perceptrons incorporating non-linear activation functions.

X_1	X_2	Y
0	0	0
0	1	1
1	0	1
1	1	0

$Y = X_1 \oplus X_2$

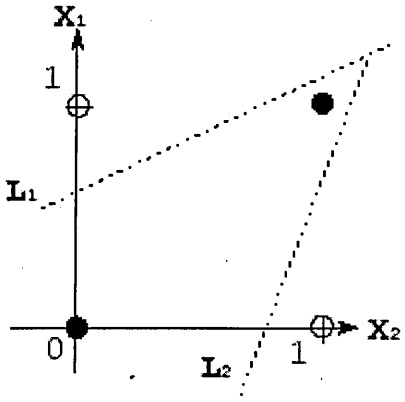


Figure 2.23: Illustration of the XOR problem (including truth table)

In 1969 Minsky and Papert published a book entitled "Perceptrons: an introduction to computational geometry" [Minsky 1969] where they stressed that a single-neuron neural network could not solve the XOR problem, depicted in Figure 2.23. The book had a negative influence and impeded neural network research for some years [Wilkes 2001, URL 12]. Research actions in solving the XOR problem lead to the introduction of multilayer networks consisting of more than one neuron organised in different layers. The XOR problem can be solved using three perceptrons, each one trained separately, as seen in Figure 2.24.

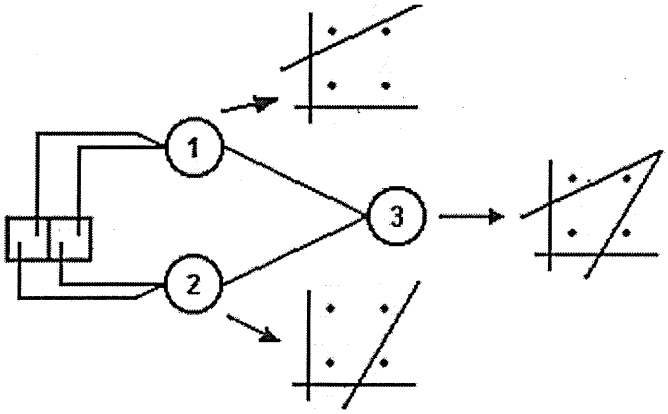


Figure 2.24: Three-neuron network solving the XOR problem

Perceptrons 1 and 2 are classifying the two sub-problems constituting the XOR problem using two straight lines and neuron 3 is combining the tasks of the first two neurons to achieve classification of the XOR problem. Using such a configuration to solve the XOR problem is not practical since each network has to be trained separately to solve each sub-problem and there is no algorithm achieving automated learning for the network as a whole. Perceptrons 1 and 2 are in principle separate neural networks performing a pre-processing task for the 3rd neural network.

One of the architectures introduced to solve the above problem was the MLP. The requisites for an MLP are:

- ❖ Each processing element must have a non-linear activation function.
- ❖ Each activation function must be differentiable throughout its range.
- ❖ A method of credit assignment is necessary to distribute the error at the output throughout the network.

An MLP having N inputs and n_o outputs is depicted in Figure 2.25. It is a two-layer network where w_{jk} are the weights between the inputs and the hidden layer (1st layer) and w_{ij} are the weights between the hidden layer and the output layer (2nd layer). It must be noted that the left-most layer (input layer) depicted is not an active layer and is only used as a buffering layer to transfer the inputs to the hidden layer. This is why the network depicted is a two-layer and not a three-layer network. Multilayer Perceptrons are trained using the backpropagation method which is based on a generalisation of the Perceptron learning rule. Standard backpropagation is a gradient descent algorithm. There are a number of variations on the basic algorithm which are based on various standard optimisation techniques.

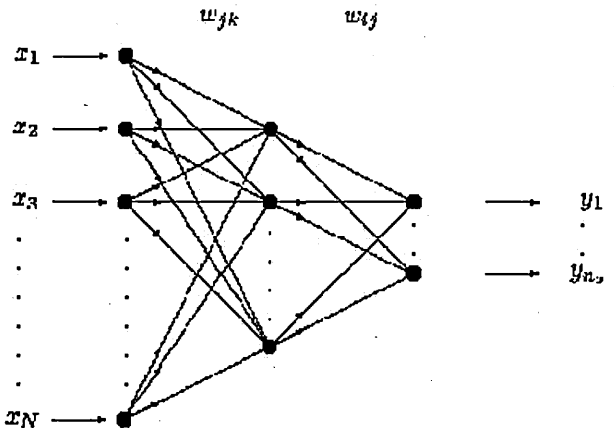


Figure 2.25: Two-layer multilayer perceptron

MLPs are alternatively termed feedforward neural networks because of the forward-only flow of information from input to output, i.e. no feedback loops. They are also backpropagation neural networks because of the backpropagation of error from output to input, as dictated by their learning algorithm.

Neural Network Learning

There exist three paradigms of neural-network learning namely supervised learning, unsupervised learning and reinforcement learning.

Supervised learning is characterised by supervision by a teacher. The teacher possesses the knowledge required in a specific problem domain and gives feedback to the neural network regarding its performance. Such feedback (error signal) is used by the network

during the training process in order for the network parameters to be adjusted for optimal performance. Examples of supervised neural networks are the simple and multi-layer perceptrons presented earlier.

In contrast to the supervised learning approach, the other two paradigms do not involve a teacher in their learning process. In reinforcement learning the learning of an input-output mapping is performed through continued interaction with the environment in order to minimise a scalar index of performance [Haykin 1994]. Although there is no teacher, an external critic is used to convert a primary reinforcement signal received from the environment into a higher quality reinforcement signal.

Unsupervised or self-organised learning does not require a teacher or critic to oversee the learning process. Learning is based on a task-independent measure of the quality of the representation that the network has to achieve. Such networks are used for data-mining.

Review of the use of ANNs for Biomedical Applications

One of the great motivations behind moving towards automated ECG interpretation and diagnostics is that manual analysis of long-term (24 h) ECGs is labour-intensive and prone to inter-observer variability. Computer techniques have been developed in order to facilitate visual analysis, e.g. by condensed printouts of various signals and trends [Lagerholm 2000]. With this type of presentation the operator usually can analyze a 24-hour recording in 20–40 minutes provided that no complex arrhythmias exist [Lagerholm 2000]. The use of automated systems for detection of arrhythmias considerably reduces the amount of time the operator needs to spend. Several commercial systems are available for long-term ECG analysis [Lagerholm 2000]. However, their performance deteriorates markedly when noise and artefacts are present and, as a consequence, an excessive number of beat classes is created which require considerable manual editing [Lagerholm 2000]. Artificial neural networks have significantly contributed in the process of automating ECG diagnostics. They have been widely used for characterisation of cardiac signals and a few such studies are reported in this section.

Kennedy et al [Kennedy 1997] employed ANNs (MLPs) for detection of Acute Myocardial Infarction. The diagnosis of acute myocardial infarction in patients with

chest pain is one of the challenges of emergency medicine while early diagnosis is critical. In their paper [Kennedy 1997] they report overall accuracy, sensitivity and specificity of 91.8%, 91.2% and 90.2% respectively when ANNs were tested on unseen data. Linear Discriminant Analysis classification results on the same dataset were 81%, 77.9% and 82.6%. Using another test dataset from another hospital yielded ANN results of 73.6%, 52.4%, 80% and LDA results of 65.1%, 28.5% and 76.9%. Examples of other researchers that used ANNs for detection of myocardial infarction are [Yang 1994] and [Hedén 1997].

Besides the MLP, other ANN architectures have been employed for ECG diagnostics. Simon and Eswaran [Simon 1997] designed an ECG classifier using Decision-based neural networks. The system was aimed at detecting the following cardiological conditions: Right Bundle Branch Block (RBBB), Left Bundle Branch Block (LBBB), Anterior wall Myocardial Infarction (AMI), Posterior wall Myocardial Infarction (PMI) and normal ECG. Al-Fahoum et al [Al-Fahoum 1999] used radial basis function neural networks, combined with wavelet transformations, for classifying cardiac arrhythmias. Unsupervised neural networks have also been employed. Hu et al [Hu 1997] used Self-Organising Maps (SOM) and Learning Vector Quantisation (LVQ) to construct a patient adaptable ECG classifier. Lagerholm et al [Lagerholm 2000] have also used SOM to perform clustering of ECG complexes.

Apart from ECG trace classification, neural networks have also been used for noise removal from ECG data [Paul 1997] and for ECG characteristic point detection [Bystricky 2002], e.g. for the detection of the T wave end.

A number of researchers have also investigated the Heart Rate Variability (HRV) signal for diagnosis of certain conditions; and ANNs have been employed in such studies. Acharya et al [Acharya 2004] have used MLPs for Heart Rate Variability (HRV) analysis with the aim to classify cardiac beats into eight categories (normal sinus rhythm (NSR), left bundle branch block (LBBB), pre-ventricular contraction (PVC), atrial fibrillation (AF), ventricular fibrillation (VF), complete heart block (CHB), ischaemic/dilated cardiomyopathy and sick sinus syndrome (SSS)).

To conclude it is emphasised that neural networks have been very successful and widely used for ECG diagnostics. It has been shown that ANNs for specific issues can perform

better than both experienced cardiologists and ruled-based criteria [Lagerholm 2000], e.g., in detecting acute myocardial infarction from the ECG [Hedén 1997]. A first generation of ANNs have also been implemented in commercial electrocardiographs [Yang 1994].

2.7.2 Knowledge-Based Systems

In the 1960's, early researchers in AI believed that the best approach to problem solving was the development and use of general purpose problem solvers, that would be able to offer solutions in a wide variety of fields [Patterson 1990]. The limitations of such systems, that were based only on a few laws or axioms, led to the introduction of knowledge-based systems. It was realised that the use of systems incorporating specialised expert knowledge in a specific domain was a much more powerful way to tackle complex problems. Feigenbaum [Feigenbaum 1977] emphasised that the real power of an expert system comes from the knowledge it possesses rather than the specific inference schemes and other formalisms it employs.

A *knowledge-based system* is any system which performs a task by applying rules of thumb to a symbolic representation of knowledge, instead of employing more algorithmic or statistical methods [Jackson 1990].

An *expert system* is a computer program that represents and reasons with knowledge of some specialist subject with a view to solving problems or giving advice [Jackson 1990]. Expert systems form a subset of the broader family of knowledge-based systems. In contrast to expert systems, a knowledge based system, although based on human knowledge, does not necessarily incorporate any expertise in the specific domain of application.

Two more definitions, presented below, are useful as part of the discussion of KBS:

- *Knowledge acquisition* is the transfer and transformation of potential problem-solving expertise from some knowledge source to a program [Buchanan 1983].
- *Knowledge representation* is concerned with the way in which information might be stored in the human brain, and the possibly analogous ways in which large bodies of knowledge can be formally described for the purposes of symbolic (non-numeric) computation [Jackson 1990].

The most common form of architecture used in knowledge-based and expert systems is the rule-based system [Patterson 1990]. According to this approach, the knowledge representation is achieved using rules of natural language. Rules consist of an antecedent ("IF" part) and a consequent ("THEN" part). Examples of such IF-THEN rules are given below:

IF *temperature* is **HOT**, AND *moisture* is **HUMID** THEN *comfort* is **VERY LOW**.
IF *Blood Pressure* is **VERY HIGH** THEN *raise alarm* AND *prescribe medication*

Linguistic values are presented in **bold**. Variable names are given in *italic*. The first rule involves two variables in the antecedent combined by an "AND" operator, while in the output only one variable is used. This example could be used for modelling the comfort of a human being in a specific location. The second rule includes only one variable in the antecedent but its consequent gives two outputs. The output of the rule is related to actions to be taken and resembles more a decision-making/control scenario rather than a modelling scenario. Significantly more complex rules, having a large number of variables in the antecedent and consequent parts can be employed.

Brief History of KBS and Expert Systems

A basic block diagram of an Expert System is depicted in Figure 2.26:

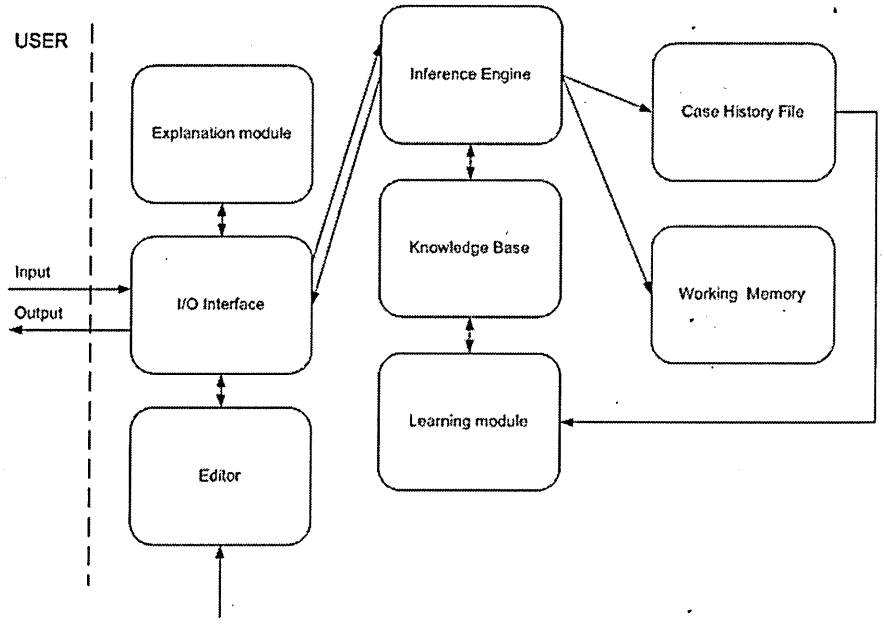


Figure 2.26: Illustration of an Expert System [Patterson 1990 pp331]

The first expert system was DENDRAL, developed at Stanford University in the late 1960s [Patterson 1990]. This system was capable of determining the structure of

chemical compounds. DENDRAL used heuristic knowledge obtained from experienced chemists. During tests, DENDRAL discovered a number of structures previously unknown to expert chemists.

Shortly after DENDRAL was completed, the development of MYCIN began at Stanford University [Patterson 1990]. MYCIN was an expert system for diagnosing infectious blood diseases and determining a recommended list of therapies. MYCIN's performance improved over several years as more knowledge was incorporated in the system. Tests had indicated that MYCIN's performance had reached or exceeded that of experienced physicians.

Two other early expert systems used in medical diagnostics were PUFF and INTERNIST. PUFF [Aikens 1983] was a diagnostic expert system for pulmonary diseases based on MYCIN. INTERNIST [Pople 1975], developed in the 1970s, was a medical diagnosis tool that contained nearly 100,000 relationships between symptoms and diseases. More recent expert systems used in biomedicine were BTDS, RESAC and a system for detection of breast cancer. BTDS (Brain Tumours Diagnostic System) [Wang 1990] was developed to aid in diagnosing the causes of brain tumours from computed tomography pictures. RESAC (Real Time Expert System for Advice and Control) [Linkens 1990], [Greenhow 1993] provided interactive advice and control during surgery. It focused specifically on the control of anaesthesia and had been embraced by human anaesthetists as they were confident to follow its suggestions. Finally, Morio et al [Morio 1989] developed an expert system for early detection of breast cancer. The system could undertake a conversation with a woman who was anxious about breast cancer. After listening to her symptoms, the system would present its conclusions and suggest courses of action to be taken.

Examples of early expert systems in other fields included PROSPECTOR, a system for assisting geologists in the discovery of mineral deposits, and R1 (aka XCON), a system used by the Digital Equipment Corporation to select and configure components of complex computer systems [Patterson 1990].

Expert Systems have been very useful in formulating the knowledge of human experts. In many fields there are only a few human experts and because of their rarity, the workload can be too much for them. For instance, in the case of industrial plant fault

diagnostics, a human expert may have to overlook many plants in different geographical locations. Coding the knowledge of such an expert to an ES allows duplication of the human expert's knowledge to many locations. The ES can also be used to train new experts and will be a tool that will never retire as is the case with human experts. Upon retirement of a human expert, the knowledge will be lost if not passed on to other humans.

Expert Systems and more broadly Knowledge-Based systems have also been used for ECG interpretation in order to achieve accurate diagnosis through modelling of the physician's ability in diagnosing ECGs. A small sample of such studies is: [Stockman 1976], [Xiong 1983], [Mylopoulos 1983], [Shibahara 1983], [Tsotsos 1987], [Kundu 1993], [Kundu, Nasipuri 1994], and [Kundu 1994].

2.7.3 Fuzzy Logic Theory

Fuzzy sets and fuzzy logic theory were introduced by Zadeh [Zadeh 1965]. The strong point about fuzzy systems is that they can combine human expertise together with sensory measurements and mathematical models. This section will present background and terminology of the field of Fuzzy Logic.

Classical (crisp) Sets and Fuzzy Sets

If Ω is the universal set, or universe of discourse, containing all the possible elements involved in a particular problem, a crisp set A within Ω is a set that has clear boundaries. Its elements have a well defined property. There are three main ways of defining a crisp set A. These are:

a. The list method where A is defined by listing all of its members.
b. The rule method where A is defined by specifying the properties that each of its members must have.
c. The membership method where a membership function $\mu_A(x)$ is introduced to denote whether an element x belongs to A or not. In crisp sets, this membership function is a discrete one and can only take two values, either 0 or 1. This is illustrated below:

$$\mu_A(x) = \begin{cases} 1 \text{ if } x \in A \\ 0 \text{ if } x \notin A \end{cases} \qquad \text{eq}^{n} (2.9)$$

Because of this crisp membership, crisp sets are unable to describe certain situations such as when sets do not have clear boundaries i.e. an element may belong to more than one set. To overcome this, fuzzy sets were introduced.

A fuzzy set A is a set where the membership of any element x of the universal set in A is described by a continuous membership function taking values in the interval [0, 1]. For every element x the pair $(x, \mu_A(x))$ can be formed containing the element and its corresponding membership degree. In general for a fuzzy set A we can write:

$$A = \{(x,\mu_A(x)) / x \in \Omega\} \qquad \text{eq}^n (2.10)$$

The value of the membership function $\mu_A(x)$ determines if the element belongs to the fuzzy set A and to what extent. It determines a degree of certainty i.e. degree of truth. This should not be confused with the probability of x belonging to the set A.

Types of Membership Functions

Three types of membership functions will be considered in this project. Two of them, the triangular and trapezoidal, are piecewise linear functions. The third type is the Gaussian membership function which is used for extra smoothness but is more complex in shape.

a) Triangular membership function

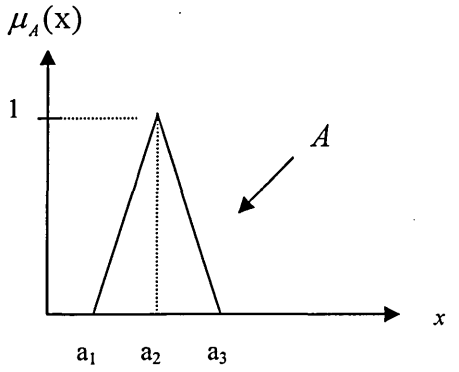


Figure 2.27: Shape of a triangular membership function

The shape of the triangular membership function, as seen in Figure 2.27, is described by the following straight line equations:

$$\mu_A(x) = \begin{cases} 0 & x < a_1 \\ \dfrac{x - a_1}{a_2 - a_1} & a_1 \le x < a_2 \\ \dfrac{a_3 - x}{a_3 - a_2} & a_2 \le x < a_3 \\ 0 & x \ge a_3 \end{cases} \qquad \text{eq}^n (2.11)$$

b) Trapezoidal

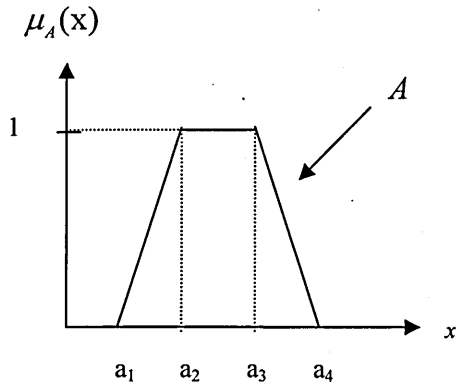


Figure 2.28: Shape of a trapezoidal membership function

The shape of the trapezoidal membership function, as seen in Figure 2.28, is described by the following straight line equations:

$$\mu_A(x) = \begin{cases} 0 & x < a_1 \\ \dfrac{x - a_1}{a_2 - a_1} & a_1 \le x < a_2 \\ 1 & a_2 \le x < a_3 \\ \dfrac{a_4 - x}{a_4 - a_3} & a_3 \le x < a_4 \\ 0 & x \ge a_4 \end{cases} \qquad \text{eq}^{\text{n}} (2.12)$$

c) Gaussian

It is given by the formula: $\mu_A(x) = \exp(-\dfrac{(x - c)^2}{2\sigma^2})$ $\qquad \text{eq}^{\text{n}} (2.13)$

where c denotes the centre of the bell-shaped curve and σ denotes the standard deviation.

Crossover Point

Consider Figure 2.29 below containing two trapezoidal (left-most and right-most) and a triangular membership function (middle). This figure is used to illustrate the concept of the crossover point. The points where the membership functions meet (at X = {25, 75}) are called crossover points. The degree of membership at the crossover points is, in this case 0.5 but this is not necessarily the case in fuzzy systems.

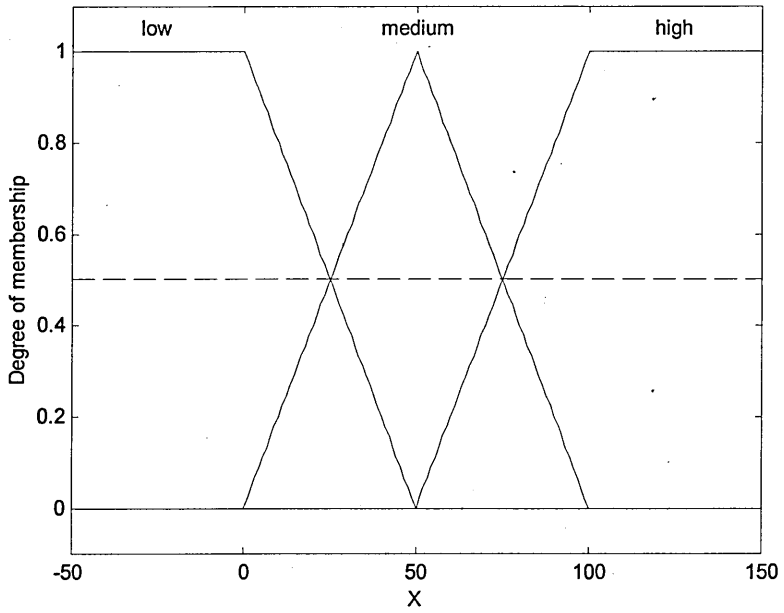


Figure 2.29: Illustration of the crossover point

The importance of the crossover point is that it marks the point at which, for a particular membership function, the certainty of belonging changes. For membership degrees higher than the crossover point the certainty of belonging to a particular membership function is higher than the certainty of not belonging. For membership degrees lower than the crossover point the opposite happens for the certainty of belonging.

Logical Operators

Boolean Logic operators can be applied on crisp sets. The standard and most widely used ones are the intersection, union, and complement operators. These operators are defined below in terms of Set Theory. If we consider two crisp sets A and B and the universal set Ω then the definitions are of the form:

1. The Intersection operator (AND)

$$A \cap B = \{x \mid x \in A \text{ and } x \in B\} \qquad \text{eq}^{n} (2.14)$$

2. The Union operator (OR)

$$A \cup B = \{x \mid x \in A \text{ or } x \in B\} \qquad \text{eq}^{n} (2.15)$$

3. The Complement operator (NOT)

$$\overline{A} = \{x \mid x \notin A, x \in \Omega\} \qquad \text{eq}^{n} (2.16)$$

For two binary variables A and B the truth tables for the AND, OR and NOT operators are given below:

Table 2.1: Truth table for AND operator

A	B	**AND**
0	0	0
0	1	0
1	0	0
1	1	1

Table 2.2: Truth table for OR operator

A	B	**OR**
0	0	0
0	1	1
1	0	1
1	1	1

Table 2.3: Truth table for complement operator

A	**NOT**
0	1
1	0

Boolean Logic operations can also be used to manipulate and combine fuzzy sets. For two fuzzy sets $\underline{A}$ and $\underline{B}$ the above three operators take the form:

1. Intersection (or conjunction) operator (AND)

$$\mu_{\underline{A}\cap\underline{B}}(x) = \mu_{\underline{A}}(x) \wedge \mu_{\underline{B}}(x) \qquad \text{eq}^{n}\ (2.17)$$

2. Union (or disjunction) operator (OR)

$$\mu_{\underline{A}\cup\underline{B}}(x) = \mu_{\underline{A}}(x) \vee \mu_{\underline{B}}(x) \qquad \text{eq}^{n}\ (2.18)$$

3. Complement operator (NOT)

$$\mu_{\underline{\bar{A}}}(x) = 1 - \mu_{\underline{A}}(x) \qquad \text{eq}^{n}\ (2.19)$$

A fuzzy intersection operator can be implemented by a "minimum" or a "product" operator. For two fuzzy sets A and B, the expression "A ∩ B" (or "A **AND** B") can be implemented as min(A,B) or prod(A,B). Similarly the union operator (A ∪ B) can be implemented using a "maximum" operator max(A,B).

Figure 2.30 illustrates the three operators for the cases of two-valued logic (crisp logic) and multivalued logic (fuzzy logic). It can be seen that in the latter case, the sets A and B are represented as membership functions and the figure depicts how the classical operators are applied on fuzzy membership functions.

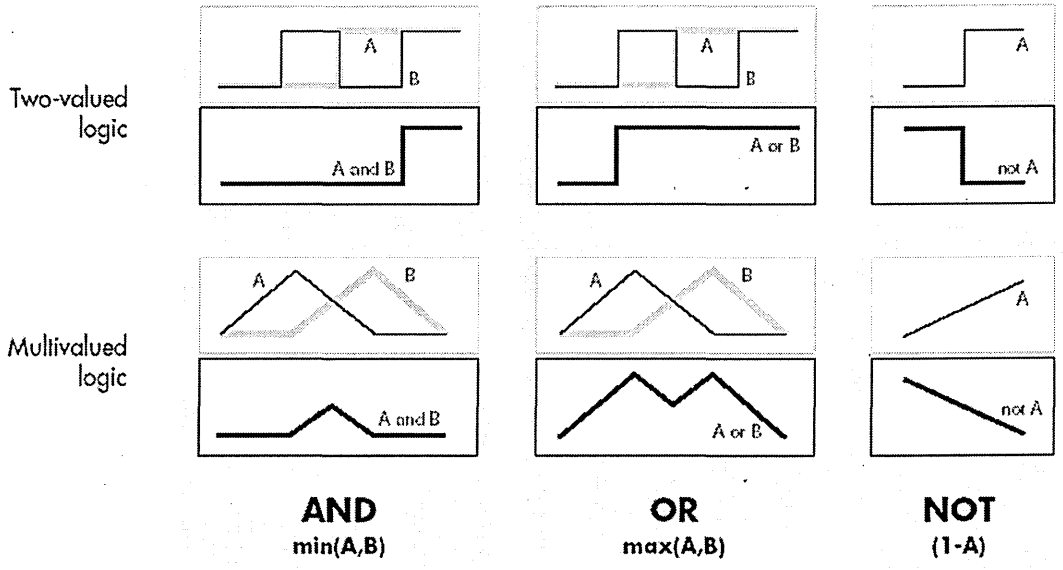


Figure 2.30: Illustrates the logical operators for crisp and fuzzy logic [MathWorks Fuzzy]

Fuzzification and Defuzzification

- Fuzzification is the conversion of a numeric input into a fuzzy input. It maps a crisp input $x \in \Omega$ into a fuzzy set A in Ω. Set A can be a fuzzy set, like the ones discussed earlier in the chapter, or a fuzzy singleton. The membership function for a fuzzy singleton is as follows:

$$\mu_A(x) = \begin{cases} 1 & \text{if } x = x_i \\ 0 & \text{otherwise} \end{cases} \qquad \text{eq}^n \text{ (2.20)}$$

  i.e. the membership is 1 only at the point x_i and zero elsewhere.
- Defuzzification is the conversion of a fuzzy quantity into a crisp quantity. Several defuzzification methods exist such as: the maximum membership, the mean of maxima, the centre of gravity methods etc.

Fuzzy Rules Processing

The two different types of fuzzy rules processing are the Mamdani-type and the Sugeno-type. The general form of the two types of rules is given below:

Mamdani: R^i: IF x_1 is A_{i1} and x_2 is A_{i2} and … and x_m is A_{im} THEN $y_i = B_i$, eq^n (2.21)

Sugeno: R^i: IF x_1 is A_{i1} and x_2 is A_{i2} and … and x_m is A_{im}

THEN $y_i = f(x_1, x_2, .., x_m)$, eq^n (2.22)

where:

- $(x_1, x_2, \ldots x_m)$ are the inputs to the system,
- y is the output

- $A_{i1}, A_{i2}, A_{im}, B_i$ are linguistic labels such as: zero (ZE), negative small (NS), positive big (PB) etc.
- $f(x_1,x_2, \ldots, x_m) = c_0 + c_1x_1 + c_2x_2 + \ldots + c_nx_n$, i.e. a linear function
- R^i denotes the i^{th} rule in the rule-base.
- $i = 1, 2, \ldots M$, where M is the number of rules.

An example of each of the above two types of rules, for a SISO system (e.g. static exercise bicycle keeping constant heart rate), would be:

Mamdani: IF heart rate is positive-big (PB) THEN make pedal torque positive-small (PS).

Sugeno: IF heart rate is positive-big (PB) THEN make pedal torque equal to 1 Nm.

It can be seen that in the Mamdani-type of fuzzy rules processing, both the antecedent (IF) and the consequent (THEN) parts are fuzzy while in Sugeno-type the consequent part is not fuzzy but a linear mathematical function. This means that in the latter there is no defuzzification operation.

A Mamdani-type controller can be seen in Figure 2.31 below:

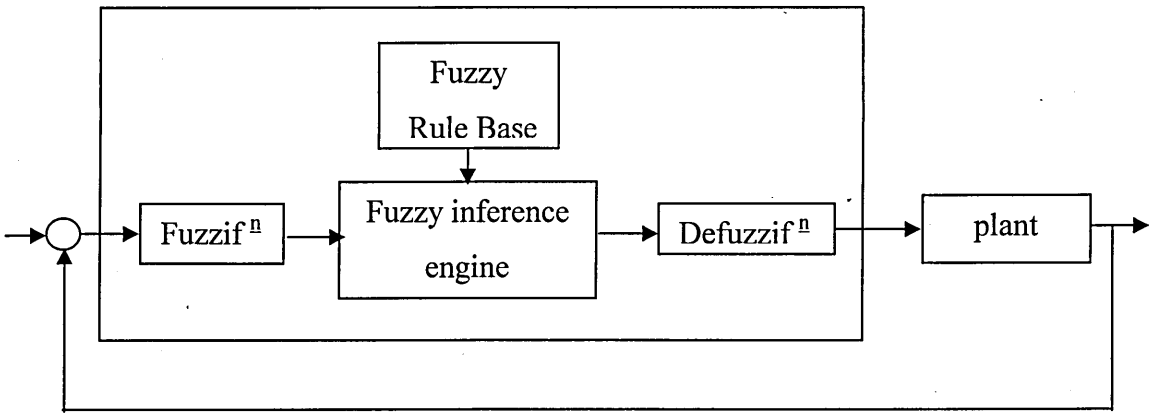


Figure 2.31: Mamdani-type fuzzy controller and plant

For a Sugeno-type controller the figure would look very similar with the difference that the defuzzification box would not be present. The "Fuzzy Rule Base" block shown in Figure 2.31 represents the knowledge about the process under investigation.

Fuzzy systems using the Sugeno-type of fuzzy rules processing are alternatively termed fuzzy TSK (Takagi-Sugeno-Kang) [Takagi 1985] systems from the names of the people that introduced it. The physical meaning of the Sugeno-type rule presented earlier is that when the input variable x is constrained to the fuzzy range characterised by the IF part of the rule, the output is a linear function of the input variables. The TSK fuzzy system

can be viewed as a piecewise linear function, where the change from one piece to the other is smooth rather that abrupt [Wang 1997]. This makes the TSK fuzzy system appropriate for piecewise linear modelling. Piecewise linear modelling is a non-linear modelling approach that uses multiple linear models. The problem is divided into partitions and a linear model is fitted in every partition. A simple fuzzy logic system example illustrating the concepts presented above is provided in Appendix B.

2.8 Summary

This chapter provided the relevant theoretical background needed to support this thesis. It presented information on Electrocardiography and the ECG signal. It discussed ECG preprocessing issues including noise removal and focused on the production of the Signal-Averaged ECG (SAECG) signal used in this study. It also discussed the process of ECG feature extraction that will be used to produce features that quantify the cardiac changes related to hypoglycaemia. ECG features that were used in a clinical study for identification of patients with the Long QT syndrome were presented. AutoRegressive modelling and the use of AR coefficients were also discussed as an alternative way, to that of feature-extraction, for ECG representation.

A review of relevant biomedical equipment was also presented. This review addressed the practical problem of measuring glucose and detecting hypoglycaemia in relation to the relevant devices. Since the ECG is involved in our approach of detecting hypoglycaemia, ECG monitoring devices were also presented. The section mainly focused on commercial devices as opposed to prototypes. Hypoglycaemia detection systems and also glucose monitoring equipment, both early and modern versions, were discussed. A few different types of commercial ECG monitors were also included. By considering this section it is realised that there is no established solution for hypoglycaemia detection, in the form of a continuous non-invasive monitoring system. The methodology for hypoglycaemia detection proposed in this thesis could, if appropriately enhanced and extended, contribute in the solution of this problem.

Moving to the classification part of this work, statistical classifiers that would be appropriate for ECG trace classification were discussed. Such classifiers were the Mahalanobis and k-Nearest Neighbour. Moreover, Artificial Neural Networks were presented including early architectures, as well as the MLP which lies among the most widely used. The use of ANN in biomedical classification problems and specifically in

the domain of ECG diagnostics was reviewed. A number of successful and promising studies employing ANN were included. Knowledge-based systems were discussed next, including both Expert Systems and Fuzzy Inference Systems. A review of such systems for ECG diagnostics was also included and their importance was emphasised.

To conclude this chapter, it is stressed that ECG feature extraction has been extensively carried out in previous studies and appears to be a wise choice to employ it in this thesis for ECG representation. Moreover Artificial Intelligence techniques, both ANN and KBS, have been successful in medical diagnostics and will be used in this thesis to achieve ECG classification into normal traces and those corresponding to hypoglycaemia. The next chapter will present the data acquisition equipment and the dataset utilised in this work.

Data Acquisition Method and Datasets

3.0 Introduction

This chapter provides information about the data used in this project and also the data acquisition equipment utilised. Data-sets consist of paired samples of ECG records and their corresponding glucose levels. Data from spontaneous hypoglycaemia were used in this project. They originated from Type 1 adult diabetic patients recruited by the Diabetic Clinic of the Royal Hallamshire Hospital in Sheffield.

3.1 Online ECG Databases

The "Physionet" [URL 13] resource for research on physiological signals contains a large number (approximately 19) of online databases [URL 14] of ECG signals accompanied by annotations by clinical experts. A few of these databases were relevant to this research, namely the European ST-T Database [Taddei 1989, 1991, 1992], the Long-Term ST Database [Jager 1996, 1998, CiC98, 2000, 2003] and the QT database [Laguna 1997]. The drawback of using such databases was that there was no glucose information to accompany the ECG records. Moreover the data of the online databases were not necessarily originating from diabetic patients. These created the need for customized ECG-glucose acquisition for this work and other related research. The data acquisition equipment and data-set are described in the following sections.

3.2 Data Acquisition Equipment

3.2.1 Portable Ambulatory System for ECG acquisition

It is advantageous to record ECG signals using an ambulatory system at the patient's own environment instead of doing so in a hospital. A portable ambulatory system (Hypoglycaemia On-line Monitoring Ensemble (HOME)) [Harris 2000] has been developed for the needs of the diabetic clinic at the Royal Hallamshire Hospital. A Hewlett Packard (HP) 200LX pocket PC attached to a single channel high gain amplifier with a serial data interface has been used, as seen in Figure 3.1.

Figure 3.1: Hypoglycaemia On-line Monitoring Ensemble (HOME)

It can record high resolution Y-lead ECG data for 1 minute every 15 minutes. It was used to record data overnight to aid the studies on nocturnal hypoglycaemia. The recording time was limited to a maximum of 10 hours by the battery life of the ECG recorder and the memory of the HP computer. Acquisition was starting at 23:00 and finishing at 7:00. The data was downloaded to a PC in the morning where off-line processing such as filtering and signal averaging could take place.

ECG data were recorded only from the Y-lead of the high resolution 3-lead orthogonal ECG. Ideally all three leads of the orthogonal ECG or all twelve leads of the standard ECG would be recorded. This was not done because of limited processing power and storage space in the HOME system. It was assumed that glucose variations are affecting the whole of the heart which means that less leads could be used. Moreover the flow of current in the heart is downwards hence most changes will happen on the Y-lead. Hence

it was assumed that changes in the cardiac function due to hypoglycaemia would be satisfactorily reflected on the Y-lead so data from just this lead would be sufficient.

3.2.2 MiniMed Continuous Glucose Monitoring System (CGMS)

The optimal way of measuring a subject's glucose is by taking a blood sample. This can be done by pricking the finger or, when repetitive sampling is necessary, by taking blood through an intra-venous (i.v.) cannula. For the data-sets used in this research, frequent sampling was necessary hence an i.v. cannula would have been used. This is only convenient for studies held in a hospital environment but is not possible for studies carried out in the patient's home. The Medtronic MiniMed CGMS system, shown in Figure 3.2, is a portable glucose meter that was used to overcome this problem.

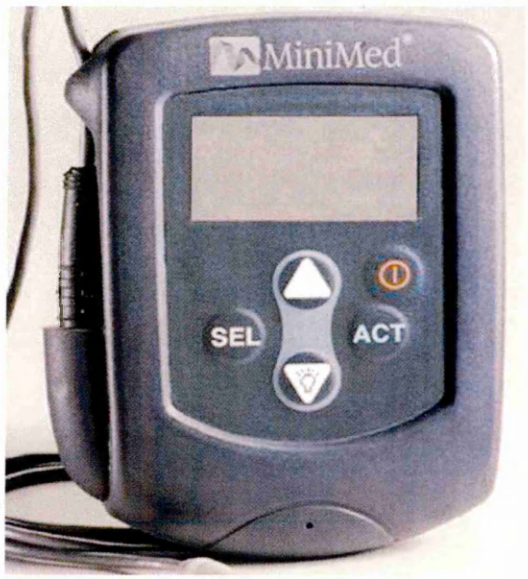


Figure 3.2: MiniMed glucose meter

A probe is inserted in the subcutaneous tissue of the tummy area and measures glucose in the tissue every 5 minutes. The probe is inserted upon a visit in the hospital and the patient can go home with the sensor and carry out his/her normal daily routine.

The difference is that MiniMed CGMS is not measuring glucose in the blood stream but in the interstitial fluid (ISF) i.e. the fluid in the connecting tissue between cells. Although ISF glucose readings closely mimic blood glucose readings, the latter is the optimal and most valid way of measuring glucose.

A limitation with measuring glucose in the subcutaneous tissue is that there exists a time-delay in glucose variations between the blood stream and the tissue which is approximately 10 minutes. It has been reported to be 9 minutes in humans [Hoss 2001], and between 5 and 12 minutes in canines [Rebrin 2000]. The delay is corrected by the software that accompanies MiniMed. Such a time-delay may be variable and could possibly depend on the rate of change of glucose. It is expected to be smaller in slow changing blood glucose but is expected to increase if the glucose is changing rapidly.

The above time-delay is another reason why ISF glucose readings are inferior to blood glucose readings, since the accuracy of the delay correction could be questioned.

Another limitation of the MiniMed glucose meter was that the minimum value it could record was 2.2 mmol/lt. There were cases where the glucose was falling below this value but these were only recorded as 2.2 mmol/lt. This is illustrated in Figures 3.5 (RHS) and 3.8 for two of the patients. For patient 202 (Figure 3.5) the glucose variable was saturated at the minimum value of 2.2 mmol/lt for records 41-58. The same happened for records 18-29 and 31-33 of patient 204. In many cases, interesting dynamics of the glucose variable may have been lost because of this since any variation below 2.2 mmol/lt was recorded as 2.2 mmol/lt. Severe hypoglycaemic events, sometimes reaching 1.5 mmol/lt or less, could not be identified because of this limitation. All hypoglycaemic events below 2.2 mmol/lt had to be treated as having the same severity.

The MiniMed CGMS system has to be calibrated using samples obtained from the blood stream by use of a finger-prick test. Three blood samples per day are needed. Calibration takes place at midnight. Calibration of the sensor should not happen during or close to an acquisition period since this could introduce disturbances on the glucose readings[5].

MiniMed CGMS has been approved by the U.S. Food and Drug Administration (FDA) but it was advised that glucose readings by CGMS were intended to supplement, not replace, blood glucose information obtained using standard home glucose monitoring devices. Moreover, the FDA panel advised that, values of glucose produced by the CGMS should not be used to make therapeutic decisions [URL 15].

More than 100 papers have been published, assessing various aspects of the CGMS system and its use in various types of studies. A number of papers[6] have been published

---

[5] Calibration at 0:00 overlapped, in our case, with the recording period (23:00-7:00). The calibration time was not customisable in the CGMS used in this study. In order to overcome this problem, the CGMS clock was shifted by 12 hours so that calibration would occur at 12 noon. The only drawback of this is that the time signatures in the spreadsheet produced, containing the captured data, had to be corrected.

[6] [Gross 2000] , [Cheyne 2002], [Gross and Ter Veer 2000], [Gross and Mastrotrotaro 2000], [Monsod 2002], [Zavalkoff 2002], [Rebrin 1998], [Sharp 2001], [Shin 2002], [Steil 2000]

that focused on the accuracy of the sensor. Gross [Gross 2000] reports good agreement of the CGMS to blood glucose meter values, under conditions of home use, in patients selected by their physicians as candidates for continuous monitoring. Unexplored areas of the sensor behaviour should obviously exist; Cheyne [Cheyne 2002] reports that "studies suggest that subcutaneous glucose levels closely mimic blood glucose levels with a lag time of only a few minutes. However, no studies have been published to show how well the sensor performs during sustained or in recovery from hypoglycaemia."

Although issues about the agreement of MiniMed readings and blood glucose readings can be raised, it must be realised that MiniMed CGMS is a niche glucose sensing system and very valuable for research purposes. It lies among the only two devices that have been approved by the FDA, the other being the GlucoWatch [URL 5] by Cygnus. The CGMS system is an invaluable tool for recording glucose profiles under circumstances, such as home self-monitoring, where no other means of measurement is available.

3.3 Dataset

The dataset used contains data on spontaneous hypoglycaemia. Forty three Type 1 adult diabetic patients were recruited for two successive nights, with one patient returning for a second acquisition which yields a total of 44 recordings. Unfortunately, not all data recorded were usable. A few of the nights recorded could not be used due to various problems such as failure of the ECG or glucose sensor, corrupted data due to noise and other artefacts and also due to human errors by the patients in handling the equipment when not accompanied by a physician or nurse. A summary of all patients constituting this dataset is included in Appendix A.

The ECG data were recorded at the patient's home using the HOME system presented earlier. One-minute worth of beat-to-beat recording was captured every 15 minutes using a sampling frequency of 125 Hz. Each one-minute recording was signal-averaged to produce a single SAECG cycle. Signal averaging was performed using the ECGLAB toolbox, based in MATLAB, which will be presented in Chapter 4. Blood glucose was recorded by the MiniMed CGMS system. The above acquisition was carried out for two successive nights, each night contributing a maximum of 33 SAECG cycles, and produced a data-set of paired ECG-glucose readings.

The profile (background info) and raw ECG data for each patient was stored in 3 binary files: one main file containing the ECG data and two complementary files containing additional information (metadata). The file extension for these three files was ".hom". The file containing the raw ECG data had a filename with the prefix "ecg" followed by the patient number (e.g. 202), followed by the ".hom" extension. The other two files had prefixes "Dsgn" & "Exp" and the remaining part of the filename was the same as before. To load the data for a given patient into the ECGLAB toolbox, all three files had to be present. (For patient 204 these would be: Dsgn204.hom, Ecg204.hom and Exp204.hom.) When the data of a patient was filtered, signal averaged and annotated, the results were stored in a MATLAB ".mat" file. Summary information for the Sheffield data-set is presented in Table 3.1 followed by presentation of the ECG and glucose profiles of sample patients.

Table 3.1: Summary information for the data-set

number of subjects`	43 + 1
number of ECG leads	1 bipolar (YY' lead)
number of records per night	≈ 33
number of nights per patient	2
ECG acquisition equipment	HOME system (125 Hz sampling freq)
glucose sensing method	MiniMed CGMS
Features extracted from (raw/SAECG)	SAECG

The ECG cycles for patient 202 are given in Figures 3.3 and 3.4 for nights 1 and 2. The ECG traces for each night are superimposed and plotted with different colours. These two figures present the ECG changes during each night. More importantly, they provide a clear presentation of the day-to-day intra-patient variability. There are extreme differences in both the P and T waves between the two nights. The T waves of the second night have lower amplitudes while the QRS complexes of the night are higher than those of night one. In addition, the ST segments of the second night are almost flat whereas the ST segments for the first night have a steep slope and are fused with the T wave upslope.

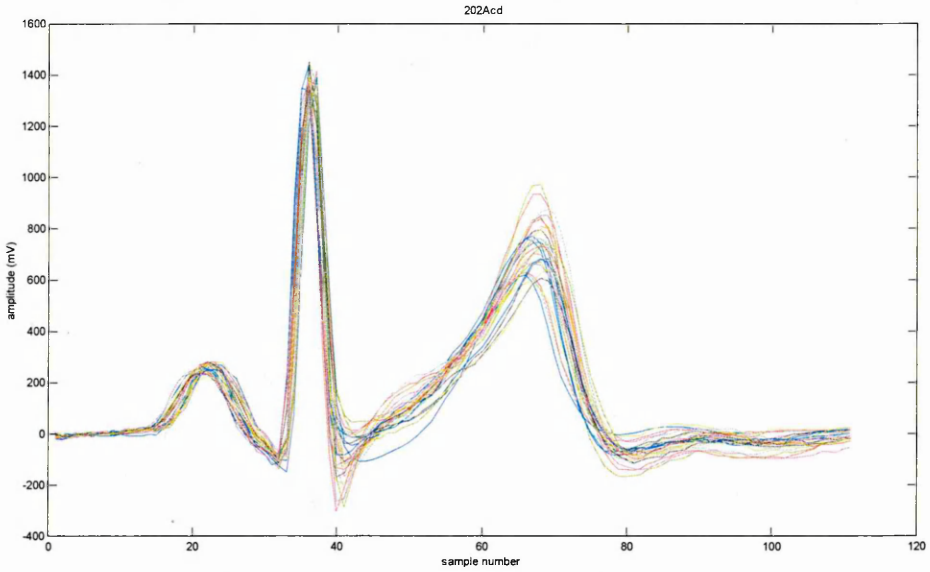


Figure 3.3: ECG traces for 202-night1 (202A)

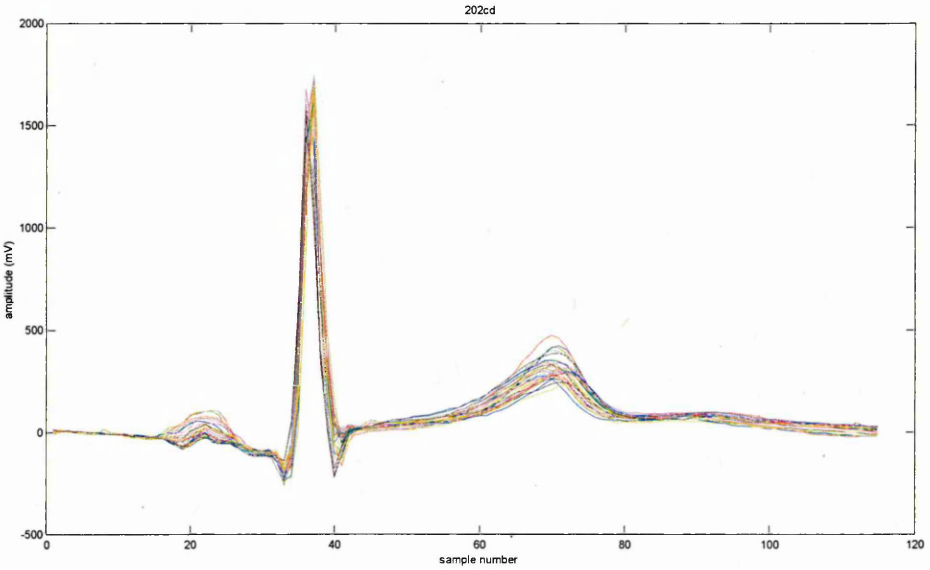


Figure 3.4: ECG traces for 202-night2 (202)

The above two figures give a clear indication of the challenges involved in classifying ECG traces, corresponding to euglycaemia and hypoglycaemia, even in the case of using a single patient (i.e. no inter-patient variability).

The glucose profiles for the two nights of patient 202 are presented in Figure 3.5.

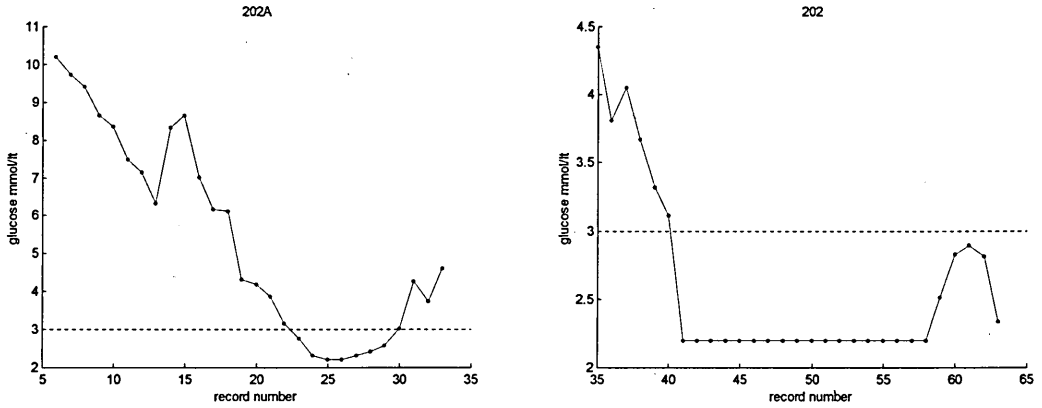


Figure 3.5: Glucose profiles for 202-night1 (LHS) and 202-nigh2 (RHS)

The horizontal dashed line marks the hypoglycaemic threshold at 3 mmol/lt. It can be seen in the two figures that the first night started as hyperglycaemic (with glucose at 10.2 mmol/lt) since it was exceeding the upper threshold of 8 mmol/lt. There was a steep descent of the glucose concentration during this night. The glucose profile of the second night was very different. The night started with glucose being at the low-end of the euglycaemic range (low-end defined as 4 mmol/lt) and went into a long period of hypoglycaemia. These two different glucose profiles of patient 202 give some reasoning for the big ECG differences between the two nights.

Based on patient 202 besides some other patients, it was observed that night-recordings starting with glucose concentrations at the low band of the euglycaemic range often exhibited flat ST segments. On the other hand, some night-recordings starting with higher glucose concentrations exhibited steep ST segments often fused with the T wave upslope. These observations were deduced by visually assessing the ST segments as opposed to using a feature extraction algorithm. Some example-cases are presented below.

Both nights of patient 223A had flat ST segments. Night-1 started at 3.6 mmol/lt with glucose increasing during the night while night-2 started at 3.48 mmol/lt and went through a long period of hypoglycaemia. Both nights of patient 229 had flat ST segments and they both went into hypoglycaemia. The first night started at 4 mmol/lt and the second started as hypoglycaemic with glucose concentrations at 2.26 mmol/lt.

Both nights of patient 228A had quite flat ST segments although the glucose concentration at the start of both nights was high, being just over the hyperglycaemic threshold of 8 mmol/lt. Night-1 went into hypoglycaemia towards the end of the

recording. Night-2 did not go into hypoglycaemia but had big glucose fluctuations. This patient was the only exception having flat ST segments while the glucose was high at the start of both night-recordings.

An example of a night-recording starting as hyperglycaemic (at 17.4 mmol/lt) that exhibited steep ST segments fused with the T upslope was 201A-night1, depicted in Figure 3.6.

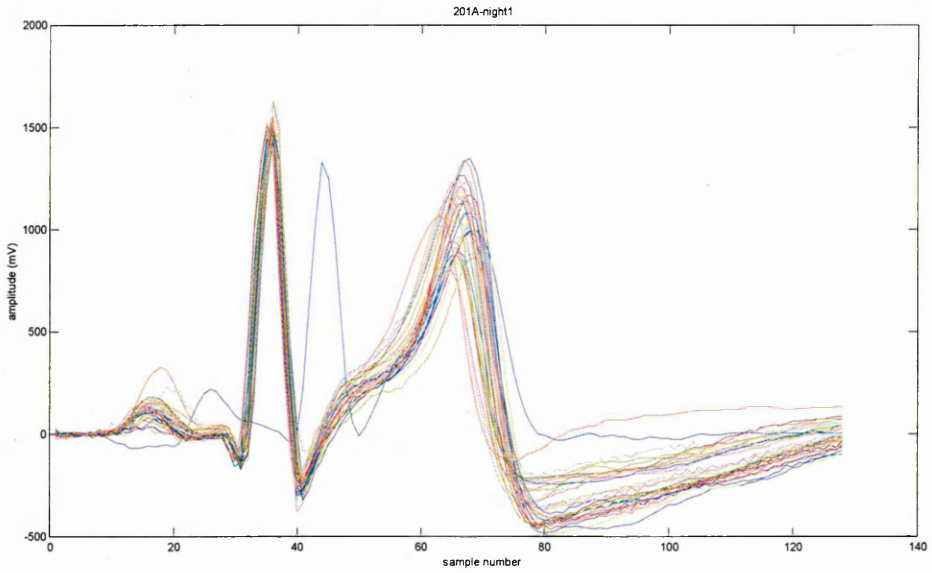


Figure 3.6: ECG traces for 201A-night1

The second night of this patient started as hyperglycaemic, at 10.12 mmol/lt, and also exhibited steep ST segments. Another example of steep ST segments in combination with high glucose concentrations at the start of the night was 205-night2, starting slightly over 8 mmol/lt.

The ECG traces of patient 204 are presented in Figure 3.7. This patient experienced very clear T wave flattening and QT prolongation in response to hypoglycaemia. Such ECG changes are the ones dictated by the research hypothesis. The glucose profile of this patient is presented in Figure 3.8.

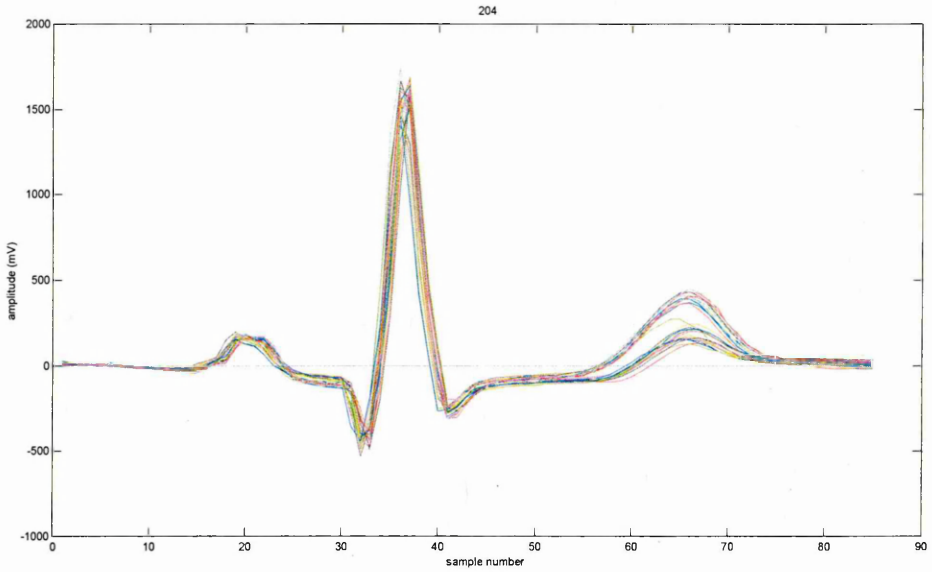


Figure 3.7: ECG traces for 204

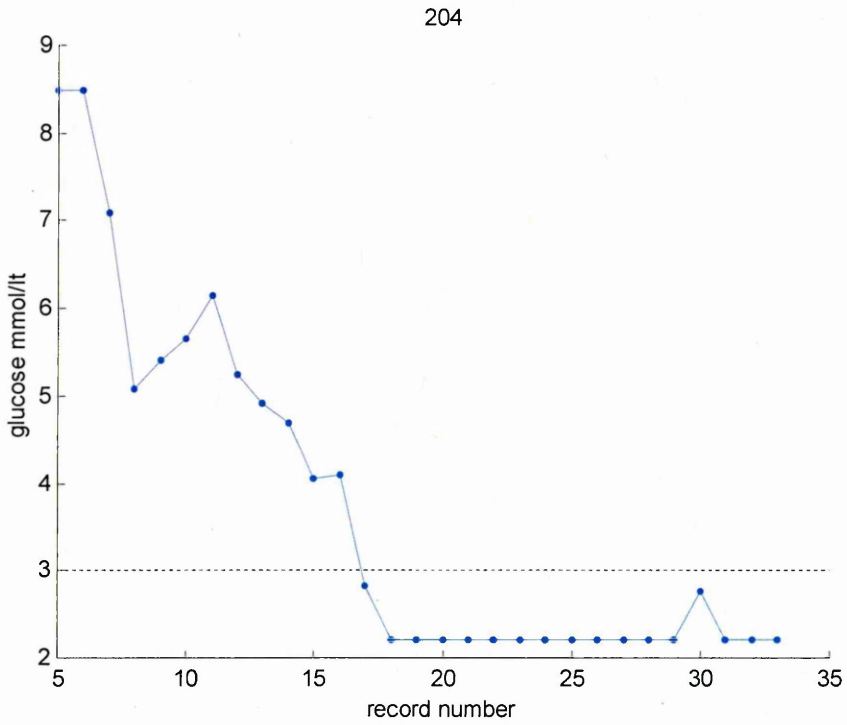


Figure 3.8: Glucose profile for 204

ECG-glucose profiles for more patients of the dataset are presented in Appendix A.

3.4 Summary

This chapter focused on the presentation of the data acquisition equipment used and the dataset utilised in this study. The equipment consisted of the CGMS glucose sensor and the custom-made ECG acquisition system (HOME). The data-set contains events on

spontaneous hypoglycaemia. Relevant details and the relevant paper publications of the studies under which the datasets were generated, were presented. Some comments and observations on the nature of the ECG and glucose profiles included in the dataset were provided. The next chapter focuses on ECG feature extraction presenting both the methodology and results.

Feature Extraction and Analysis of Signal-Averaged Electrocardiogram Signals

4.0 Introduction

This chapter presents the methodology and results related to the feature extraction process undertaken in this research. Firstly the MATLAB toolbox (ECGLAB) that was used as the main software platform for ECG processing is introduced. The issue of defining an appropriate hypoglycaemic threshold is then raised, followed by the presentation of the overall methodology for developing a hypoglycaemia detection system. The sub-processes and components of the overall system are then outlined, namely the feature extraction and classification of ECG traces. The classification of ECG traces is grouped into two different approaches: (i) the approach of using static pattern classification of ECG features with no temporal information incorporated (Chapter 5) and (ii) the approach of classifying ECG traces based on the time series of the ECG features used (Chapter 6).

ECG representation was carried out by either using direct ECG feature extraction or by AutoRegressive (AR) modelling. The former describes each ECG trace using time-interval or morphological features and the latter represents each ECG trace by means of AR coefficients. Section 4.4.1 presents the relevant algorithms that are necessary for detection of the ECG characteristic points. Once the ECG characteristic points have been defined, a comparative study of geometric methods for annotation of the T wave end is included. Then the assessment of morphology of the T wave is presented. Next, a number of ECG features are presented and some analysis of their usefulness in relation to hypoglycaemia is carried out. Following these, the autoregressive modelling of post-QRS ECG segments is discussed. This is an alternative approach to that of using individual ECG features, for ECG representation.

4.1 The ECGLAB® toolbox

ECGLAB® [Ireland 2001] is a custom-made ECG processing toolbox, running in MATLAB that has been developed by RH Ireland for the needs of the studies on spontaneous and experimental hypoglycaemia[7]. ECGLAB has a graphical user interface and it can display raw and Signal-Averaged ECGs. It allows signal averaging to be performed provided that the corrupted records are discarded manually, and allows markers to be set manually on the Signal-Averaged ECG (SAECG) records. This toolbox was also used in this research, as a tool for viewing and annotating ECGs and mainly as a platform for developing new algorithms, implemented in MATLAB. Two screenshots from ECGLAB are presented in Figures 4.1 and 4.2.

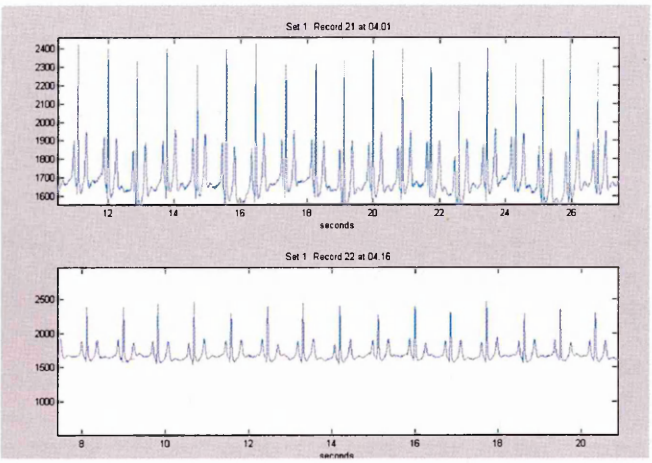


Figure 4.1: ECGLAB screenshot displaying raw (beat-to-beat) ECG

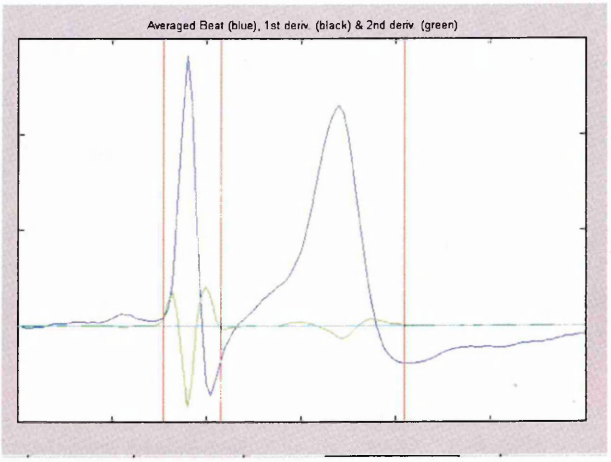


Figure 4.2: ECGLAB screenshot displaying a SAECG cycle with 2nd derivative information plotted (green) and vertical markers set.

The left figure displays raw ECG trains before the averaging process while the right figure displays a SAECG cycle with second derivative information plotted in green and with vertical markers (red vertical lines) set by the user to mark the Q, S and T end.

4.2 Choice of hypoglycaemic threshold

In order to define the medical conditions of euglycaemia (normality) and hypoglycaemia, for the purposes of this research, a hypoglycaemic threshold must be selected. A threshold lying in the interval [2.5 3.5] mmol/lt has been used in various studies. A threshold of 3.5 mmol/lt has been used by Harris et al [Harris 2000], while a threshold of 2.5 has been used by Robinson et al [Robinson 2004] in their studies on

[7] ECGLAB was developed externally and is not distributed as a commercial toolbox with the MATLAB software package (Mathworks Inc).

hypoglycaemia. In our case, thresholds of 2.5 and 3 mmol/lt were used. The choice of a threshold to define hypoglycaemia is not an easy one to make. A threshold had to be chosen bearing in mind that the data from the two classes would have to be classified by a neural or statistical classifier which is not the same situation with that of a clinical study. The threshold should be chosen so that the task of the classifier would be eased. The two classes formed should be distinguished just by using ECG features corresponding to these two classes. In some cases, data belonging to the ambiguous range of glucose values between euglycaemia and hypoglycaemia was excluded in order to show more abrupt changes of ECG features between the two classes. This transition region between euglycaemia and hypoglycaemia was normally in the interval (2.5 4) mmol/lt when a threshold of 2.5 mmol/lt was used and in the interval (3 4) mmol/lt when a threshold of 3 mmol/lt was used.

A second threshold was also necessary; this is the threshold between euglycaemia and hyperglycaemia which is the condition of abnormally high glucose levels. The choice of such a threshold is not very critical compared to the hypoglycaemic threshold. After consulting our medical collaborators this was chosen to be 8 mmol/lt.

4.3 Hypoglycaemia Detection Approach

In this research work, a methodology was proposed according to which a diagnostic system can be implemented for hypoglycaemia monitoring. The proposed approach for hypoglycaemia detection is presented in Figure 4.3.

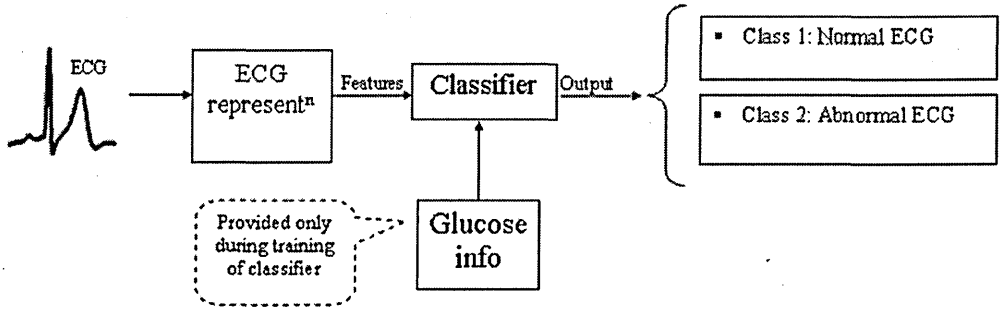


Figure 4.3: Feature Extraction and Classification System

The proposed system consists of an ECG representation stage in cascade with a classification stage. After the appropriate preprocessing and filtering, the ECG is fed to the ECG representation stage. (The system as depicted in Figure 4.3 receives a preprocessed and filtered ECG and hence an ECG preprocessing stage is not presented

as part of the system.) In the representation stage, the ECG signal is described numerically in an appropriate manner so that it can be classified in the following stage. Two approaches were used; the first one was to extract ECG features to be used for ECG representation and the second was to achieve ECG representation by means of AR coefficients. The output of the representation stage (ECG features or AR coefficients) was fed to the classifier, either a neural network (Multi-Layer Perceptron (MLP)), a statistical classifier (Linear Discriminant Analysis (LDA) or k-Nearest Neighbour (kNN)) or a knowledge based system (expert system or fuzzy inference system). During the test/monitoring phase, the classifier would classify the input vector fed as either normal or abnormal corresponding to hypoglycaemia. In the training phase, glucose information would also be available for the classifier. The glucose variable informs about the euglycaemic/hypoglycaemic state. It is used to provide the targets for the supervised classifiers (MLP, LDA, kNN) or alternatively to aid the construction of the rule base of the KBS.

4.4 ECG representation by extraction of ECG features

This section presents the feature extraction of ECG signals. Feature extraction is the process of extracting parameters (or features) from a recorded signal. In our case features are devised to describe certain physiological responses on the ECG. Feature extraction can take place in the time domain, which was the focus of this thesis, in the frequency domain, or simultaneously in both the time and frequency domains using time-frequency localisation techniques such as the Short-Time Fourier Transform (STFT). In the time domain, both time-interval and morphological features can be extracted. Time-interval features simply describe the duration of a component of the signal while morphological features can describe aspects such as the symmetry, the area under a curve, the presence or non-presence of a component.

4.4.1 ECG characteristic points

As discussed in Section 2.3.2, each ECG cycle consists of a number of ECG characteristic points (sometimes referred to as "ECG significant points"). The characteristic points most relevant to this research are: (i) the Q point, R peak, S point, which are the onset, peak and offset of the QRS complex and (ii) the T onset, T peak and T offset. These characteristic points are necessary for the process of extracting ECG features since the definition of ECG features is based on these points.

Detection of the ECG characteristic points

All algorithms discussed in this section were designed to work on SAECGs since the feature extraction process was carried out solely on SAECGs. Raw (beat-to-beat) ECGs were only used to produce the signal-averaged cycles and feature extraction of the raw cycles was not carried out. Therefore, the algorithms were not tested on raw ECG signals as this was beyond the scope of study.

Automatic detection of the R peak

This algorithm detects the temporal location of the R peak in SAECGs. Although established algorithms exist for the detection of the R peak [Balda 1977], [Moody 1982], our own code was developed. This was done firstly because most established algorithms are proprietary and were not found freely available at the time that this algorithm was developed and secondly because design of such an algorithm provided a deeper insight into ECG feature extraction. Annotation of the R peak was necessary for all ECG features using this ECG characteristic point in their definition (i.e. RT, RTapex features). The R detection algorithm developed for this project is given below in pseudocode[8] form:

```
LOAD current ECG record
Calculate 1st derivative of ECG record
FIND min & max of 1st deriv and store min_deriv_index &
max_deriv_index
Calculate difference: min_deriv_index - max_deriv_index
IF difference > 80msec
      Reduce min_deriv_index by 40msec
END
FIND new min of 1st deriv (after above reduction)
FIND max of portion of ECG lying between max_index & min_index(this
is the R peak)
```

In the pseudocode, "min_deriv_index" and "max_deriv_index" describe the temporal location of the extrema of the 1st derivative. The difference (min_deriv_index -

[8] Operations such as: calculating the length of arrays, calculating sampling intervals, printing messages in MATLAB command window etc are not included in the pseudocode since they are straightforward operations that do not contribute in the understanding of the operation of the algorithm.

max_deriv_index) is positive for upright QRSs and negative for inverted QRSs. If the difference is greater than 80 msec then the min and max indices found will not correspond to the QRS complex. The 80 msec threshold was chosen after investigation on our dataset.

The min and max indices correspond to the points of inflection (2^{nd} derivative = 0) to the left and right of the R peak. If a T wave higher than the QRS complex exists then its peak and its points of inflection may be detected instead of those of the QRS. By calculating the difference between the two points of inflection it is inferred whether they belong to the T wave or the QRS complex. The T wave being a wider wave in time, with a 1^{st} derivative not being as steep as that of the QRS, will have points of inflection further apart than the QRS will.

ECG annotation algorithms are normally tested by comparison to manual annotations produced by clinical experts. Manual R peak annotations on our dataset were not available so the R detection algorithm had to be assessed by visual inspection[9]. Visual inspection involved the assessment of the accuracy of an annotation superimposed on the current ECG trace and displayed on the visual display unit. The R detection algorithm presented here was tested visually on all ECG traces used in the studies carried out in this work and was annotating correctly the R peaks.

The objective for our R detection algorithm was to design a method that would work satisfactorily on the data used in this project as opposed to producing a robust R detection algorithm to be used for generic R annotation. Hence our algorithm was not tested on other ECG datasets.

A SAECG record (p203rec41) is presented in Figure 4.4. It depicts the detection of the R peak (marked by a black dashed line) on an ECG record that includes a T peak higher in amplitude than the R peak. The 1^{st} and 2^{nd} derivative information are also plotted in black and green respectively.

[9] carried out by C Alexakis

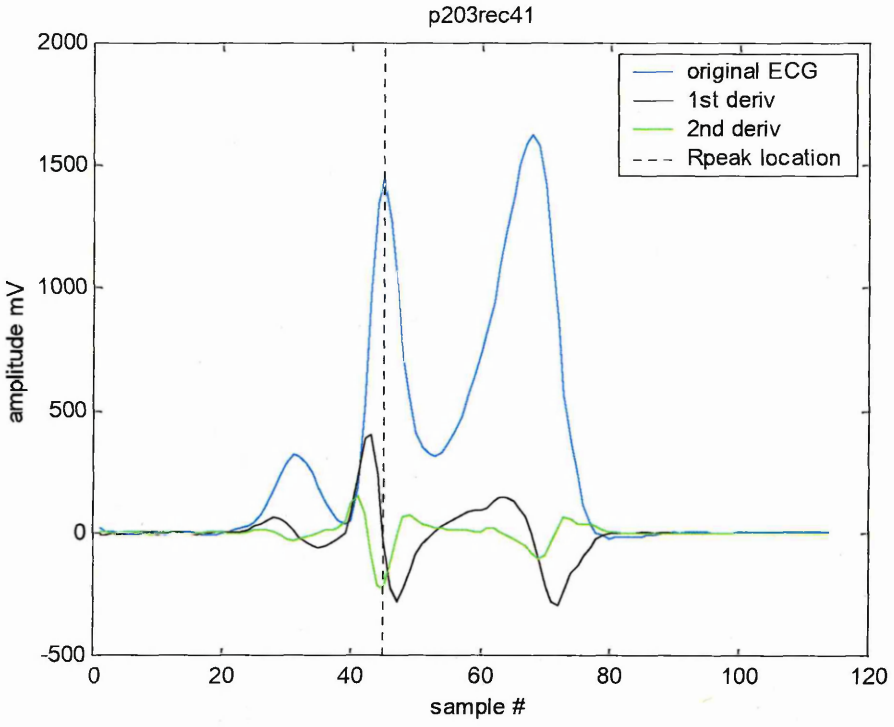


Figure 4.4: ECG trace exhibiting a T wave higher than the QRS. R peak annotated correctly by the R detection algorithm

By employing criteria based on the 1st derivative of the ECG signal, the algorithm correctly recognised the R peak, even in the presence of a T wave higher than the QRS complex. In a normal ECG cycle, the QRS is the component with highest amplitude and hence easy to detect. The algorithm was not tested on beat-to-beat ECG traces. The latter may be more difficult to annotate as there is more contamination by noise.

Algorithm for detection of the temporal location of the T peak

This algorithm detects the temporal location of the T peak in SAECGs. The T detection algorithm can detect the peak in the following cases:

1. normal T waves
2. inverted T waves
3. biphasic T waves

The algorithm was able to detect more than one T peak. In the case of biphasic T peaks the algorithm detects two peaks, one in the positive and one in the negative phase. (Inverted and biphasic T waves were discussed in Section 2.3.1.) The algorithm was tested on SAECG traces from our dataset, having a sampling frequency of 125 Hz. The pseudocode for the algorithm is given overleaf.

```
LOAD current ECG record
LOAD x coordinate of R peak (i.e. R peak temporal location,
calculated by R detection algorithm)
CALCULATE 1st and 2nd derivative of the ECG record loaded
PERFORM a forward search to find the 1st point of inflection to the
right of the R peak
CALCULATE R104 point
FIND which is the rightmost between R104 and 1st inflection point
SET as startpt (starting point for T peak search) the rightmost
point found
Set T peak search interval from starpt up to 120msec before the end
of the ECG trace

FIND all possible T peaks (both +ve & -ve peaks are considered)
STORE in descending order all possible T peaks in peaks_array

% CHECK T peak candidates found for validity
number_of_valid_peaks = 0 % initialise to 0
loop_counter = 0 % initialise to 0
WHILE number_of_valid_peaks < 2
   INCREMENT loop_counter
   GET current T peak candidate (according to loop_counter) from
   peaks_array (i.e. peaks_array(loop_counter))
   FIND points of inflection immediately to the left and right of T
   peak
   CALCULATE temporal difference between points of inflection found

   ACCEPT T peak candidate based on conditions below:
   1. Existence of points of inflection to the left and right of the
      T peak.
   2. change of sign of 1st derivative immediately to the left and
      right of T peak candidate.
   3. inflection point temporal difference ∈ [16 121] msec
   4. inflection point Voltage difference < 400 mV
   5. absolute value of 2nd derivative at T peak > 2mV
END of WHILE
STORE T peak candidate with highest absolute amplitude
```

The algorithm works by scanning the post-QRS section of the SAECG trace, i.e. detection of the R peak is a pre-requisite for this algorithm. The forward search starts at the R104 (104 msec to the right of R peak) or at the first point of inflection after the R

peak, whichever occurs later. The T peak search ends 120 msec before the end of the ECG trace.

Once a T peak candidate is found, the following acceptance criteria are applied to decide whether the identified peak is valid:

1. Existence of points of inflection to the left and right of the T peak candidate.
2. Change of sign of 1st derivative immediately to the left and right of T peak candidate.
3. Distance between the left and right points of inflection lying in the interval [16 121] msec. (The left and right points of inflection considered are the ones immediately to the left and immediately to the right of the T peak.)
4. Voltage difference (i.e. y axis difference) at the points of inflection must be smaller than 400 mV.
5. The absolute value of the 2nd derivative of the ECG trace at the x coordinate where the prospective T peak appears must be greater than 2 mV. (This criterion is necessary to distinguish between ECG components (such as small undulations) that comply with the above criteria but have a very small absolute value for the 2nd derivative at the peak detected.)

If a second T peak is detected that complies with the above criteria, it is kept only if it exceeds 60% of the amplitude of the highest peak. The highest peak is considered to be the main T peak. Figure 4.5 illustrates an example of correct T annotation in a record where the ST segment has greater absolute amplitude than the T peak.

The figure depicts the ECG trace in blue and the 2nd derivative signal in green. Dotted blue vertical markers denote the position of the points of inflection to the left and right of the T peak. The vertical green marker denotes the position of the T peak. Despite the fact that the ST has greater amplitude than the T peak, the algorithm has correctly identified the latter.

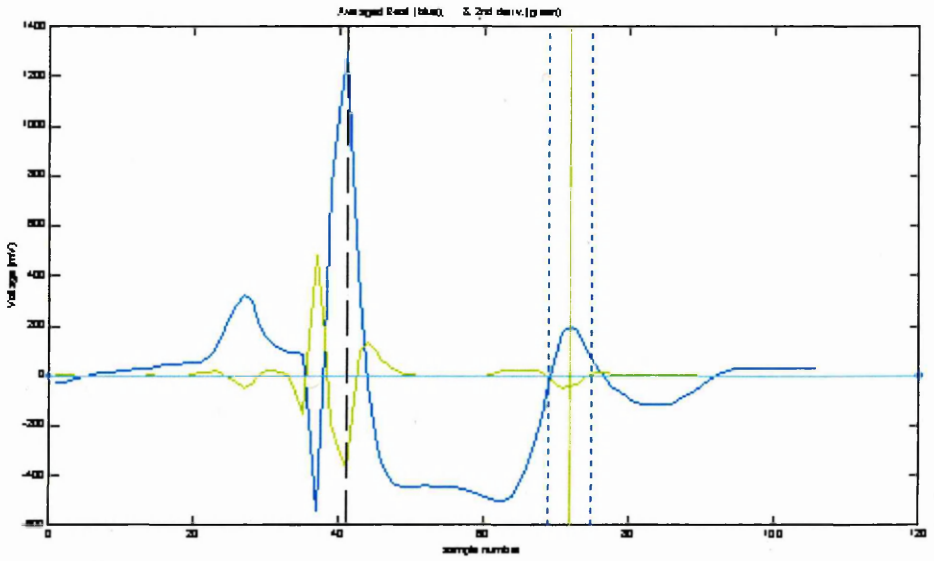


Figure 4.5: T annotation for patient 216 record 41.(BLUE: ECG trace, GREEN: 2nd derivative)

T wave End Detection

As stressed in Chapter 2, detection of the end of the T wave constitutes a major algorithmic problem in the field of ECG annotation due to the ambiguity of location of the T wave end under noisy signal conditions, disturbed post-T-wave baselines and the like. A number of existing geometric methods were adopted to perform T end annotation for the needs of this work, namely the tangent method (Maximum Slope Intercept or MSI), the Peak Slope Intercept (PSI) and the Fitting method. A brief theoretical background for these algorithms was given in Chapter 2 (Section 2.3.3). The above three algorithms were implemented in MATLAB using the ECGLAB toolbox as a platform.

Tangent Method or Maximum Slope Intercept (MSI)

This method finds the point of the T wave downslope[10] having the steepest tangent and marks the end of the T wave at the point where the steepest tangent line meets the isoelectric line[11]. The pseudocode for the tangent method is given overleaf.

[10] or upslope in the case of an inverted T wave

[11] horizontal line at 0 volts

```
LOAD current ECG record
LOAD T peak coordinates
CALCULATE 1st and 2nd derivative

IF T wave is normal (i.e. upright)
      FIND minimum of 1st derivative on T downslope
ELSEIF T wave is inverted
      FIND maximum of 1st derivative on T downslope
END

CALCULATE the point where the tangent line meets the isoelectric
line to find temporal position of T end.
STORE T end coordinates
```

After the algorithm is executed, the R coordinates and the T end coordinates are used to calculate the RT interval according to the tangent method.

Peak Slope Intercept (PSI)

This method marks the end of the T wave according to the intersection of the isoelectric line with the line defined by the T peak and the point of maximum slope (inflection point) on the T wave downslope. The pseudocode describing the PSI method for marking the T wave end is given below:

```
LOAD ECG cycle
LOAD T peak coordinates
CALCULATE 1st and 2nd derivative

IF T wave is normal (i.e. upright)
      FIND minimum of 1st derivative on T downslope
ELSEIF T wave is inverted
      FIND maximum of 1st derivative on T downslope
END

CALCULATE the point where the straight line (defined by the T peak
and the point of inflection on the T downslope) meets the
isoelectric line to find temporal position of T end.
STORE T end coordinates
```

After the algorithm is executed, the R coordinates and the T end coordinates are used to calculate the RT interval according to the PSI method.

First order fitting method (FIT)

This method marks the end of the T wave by fitting a straight line on the T downslope. The T end is defined as the point where the line fitted meets the isoelectric line. In the implementation of the method for the needs of this work the range of data on which the straight line was fitted spanned from T peak up to the point of inflection on the T downslope. The pseudocode for this algorithm is given below:

```
LOAD ECG cycle
LOAD T peak coordinates
CALCULATE 1st and 2nd derivative
FIND point of inflection on T downslope

SELECT range of fit from Tpeak to point of inflection on T downslope
Fit a 1st order polynomial in a least-squares sense (using polyfit
function)

CALCULATE the point where the fitted straight line meets the
isoelectric line to find temporal position of T end.
STORE T end coordinates
```

In a similar manner a second order polynomial can be fitted. Fitting polynomials of higher orders is not advisable since such polynomials will closely follow the shape of the T wave downslope and may mask any prolongations of the QT interval. A second order fitting method was included in the RT comparative study presented in Section 4.4.2.

The MSI method relies on a single point on the T downslope to define the T wave end while the PSI method relies on two points and the fitting method relies on a whole section of the downslope. The fact that the MSI method only relies on one point does not mean that it is inferior to the other algorithms. Reliance on a single point may appear as if the algorithm will be more sensitive to noise but in practice the algorithm proved to be robust, as it will be seen in the section where the algorithms are compared.

The following section presents a comparative study of geometric methods for T wave end detection, and is directly related to the RTc, RTapexc and QTc features.

4.4.2 Comparative study of geometric methods for marking the end of the T wave

A comparative study of three geometric methods used to mark the end of the T wave is presented in this section. By geometric methods we refer to algorithms that work on the T downslope using geometric criteria such as tangent lines or fitting of best straight lines for marking the end of T. The first 3 algorithms described in chapter 2 (msi, psi, fit) are considered to be such methods. These methods comprise relatively simple approaches, in concept, for marking the end of T. The motivation for this study was to assess these algorithms specifically on our data and compare the algorithms in the context of hypoglycaemia detection using the ECG. Moreover the tangent method has been used, in a semi-automatic way[12], by our medical collaborators in their clinical studies on the manifestation of hypoglycaemia on the ECG. Therefore, assessment of this method and comparison with other similar methods would be useful to them as well.

Five methods were considered in total. This includes the aforementioned three plus a 2nd order fitting method and a manual method for benchmarking the automatic ones. Manual marking of the end of T is the current gold standard. The manual marking of the ECG records was performed by a biomedical scientist[13] familiar with ECG annotation. The RT interval was used[14], instead of the traditionally used QT, in order to assess the various algorithms. This is because identifying automatically the Q point on the ECG can be a difficult task to perform especially in the presence of noise. The R point (the peak of the QRS complex) can be detected a lot more easily and accurately than the Q point. Moreover it is obvious that the RT interval still describes satisfactorily the duration of ventricular repolarisation so it can be used as a predictor of this arrhythmia. A few other researchers, e.g. [Porta 1994, 1998], have also considered the RT instead of the QT. The RT is defined as the time interval from the R peak to the end of the T wave.

---

[12] The T end marking was semi-automatic in the sense that a tangent line was used as a visual aid when performing manual T end annotation.

[13] Cath Davies from the Royal Hallamshire Hospital in Sheffield.

[14] R detection was automatic and T end detection was done manually by Cath Davies.

Methodology for the comparative study

All the algorithms were implemented in MATLAB using the ECGLAB toolbox as a platform. The manual marking, in order to be blinded, was carried out on a blank computer screen with no labelled axes or any sort of visual aids. The second order fitting method was included because it had not been found to have been tried in the literature. A second order polynomial would follow more closely the T wave downslope and it was suspected to partly mask the prolongations in the RT. However, the range of the T wave downslope used for fitting was small which means that the second order fitting algorithm would not closely follow the downslope. We wanted to test this algorithm and investigate the extent to which it could demonstrate prolongations in RT or QT. Moreover it was easy to include the second order method in the study since its implementation was largely based in the first order method. The way that the second order fitting method works is by fitting a second order polynomial on the downslope of the T wave. The root of the polynomial that occurs to the right of the T peak marks the end of the T wave i.e. the end is marked as the point of intersection of the polynomial with the isoelectric line to the right of the T peak. The other root, occurring to the left of the T wave is ignored. In this study the range of data on which the polynomial was fitted, ranged from the T wave peak to the point of steepest tangent on the T wave downslope. This was the case for both the first and second order fitting methods.

The above T end detection algorithms were examined and compared using a few different approaches. Firstly the graphs of RTs over time were inspected visually for each algorithm. This way it was seen how each method behaved and the level of offset (baseline level) of each one was identified. Identifying the offset indicated the degree of over/under-estimation of the RTs by each algorithm. Correlation coefficients were also used in order to identify the correlation between each of the methods. Bland-Altman plots [Bland 1986] were used in order to examine the level of agreement between the manual and each other method.

It has to be noted at this point that the RT intervals used in this study were not corrected for heart rate. This was because we were interested in comparing the performance of the algorithms. The algorithms were assessed according to how well they resembled the manual method and not according to how well they correlated with glucose or how well they could predict certain cardiac arrhythmias. So heart-rate-correction was not considered necessary. But even if correction had been applied, it would not have

introduced any change since all RTs would have been corrected using the same algorithm. RT intervals from all algorithms would have been divided by the same number (RTc = RT / √RR) according to Bazett's correction formula [Bazett 1920]. This would effect as a form of normalisation which was not necessary.

Comparative Study Results

The diabetic patients used for this study were: 201A, 203, 204, 205, 207, 208, 209, 215, 216, 222 and 227. Data exists for two nights (normally 66 records) for all patients except 204 and 222 which gave single-night records due to problems during data acquisition. This means that approximately 660 ECG cycles were used in this study. Patients having inverted or biphasic T waves were not included in the study. Figure 4.6 illustrates the RT intervals produced by the five different methods for the first night of subject 207. It can be seen that the manual method produced the longest RTs and the second order fitting method produced the shortest.

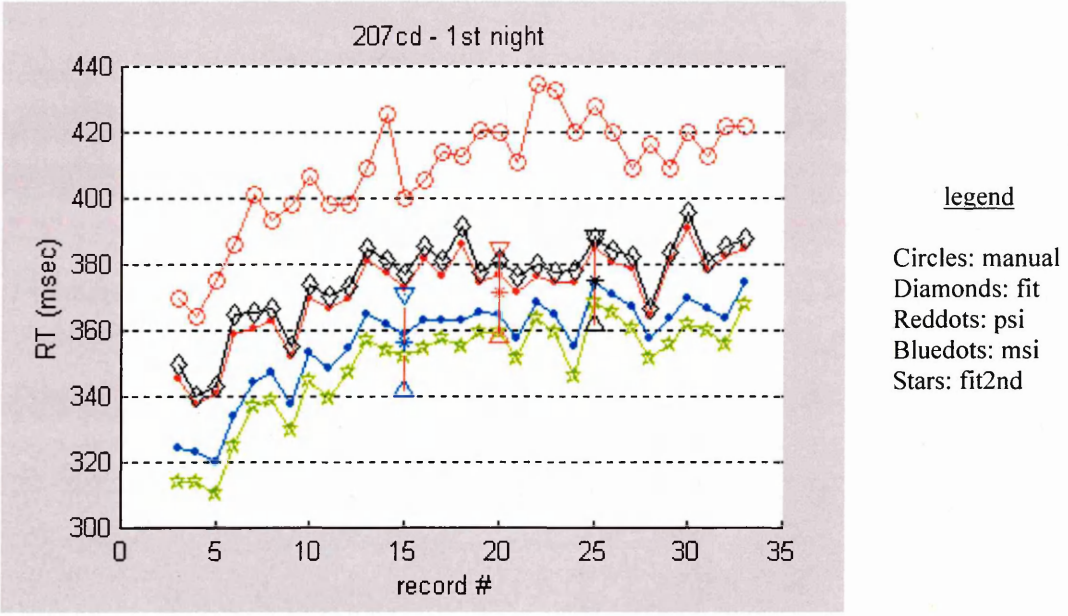


Figure 4.6: RT intervals by the 5 algorithms for 207-night1

For each RT measurement method used, the mean and standard deviation of the RT values over all patients and all nights were calculated. This gives an indication of the overestimation or underestimation of the RT interval by each method. The results are shown in the bar chart below (Figure 4.7):

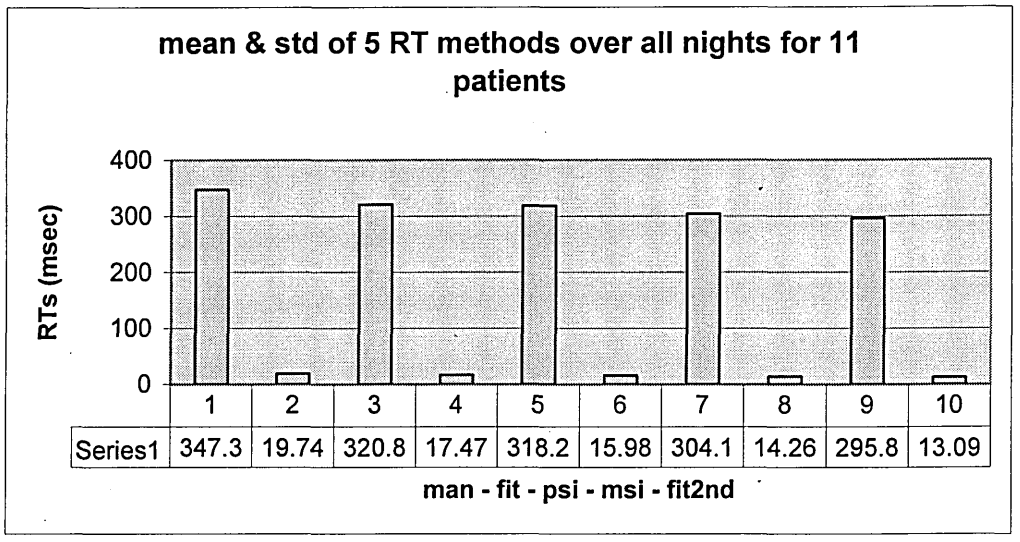


Figure 4.7: Mean and standard deviations across all patients, for the 5 Tend annotation methods

Bars labelled with odd numbers correspond to the mean values, while bars labelled with even numbers correspond to standard deviations. The table incorporated in the figure informs about the exact height of the bars.

It can be seen from Figure 4.7 that the manual method gives the longest RTs followed by the first order fitting method, then by the peak slope intercept method, the maximum slope intercept method (tangent method) and then by the second order fitting method (man>fit>psi>msi>fit2nd). The psi and fit methods gave very similar results. This is also apparent from the correlation coefficients between psi and fit.

Exceptions. The level of offset of the RT interval identified by each algorithm was different but consistent across patients (man>fit>psi>msi>fit2nd). Nine exceptions were observed, in the above result, out of the 660 records considered. They occur in the following patients, among the 11 patients used:

- Patient 205-night1: psi and fit longer than manual for records: 2, 3, 5.
- Patient 205-nght2: psi and fit longer than manual for records: 51, 54, 57, 59.
  (For record 42 fit is slightly longer and psi is shorter than manual. This record was not considered among the exceptions.)
- Patient 227-night2: psi and fit longer than manual for records: 39 and 46.
  (For record 50 fit is slightly longer and psi is slightly shorter than manual. This record was also not considered among the exceptions.)

Agreement among first derivatives of the RT methods

Although the first derivatives of all the methods agree in most of the cases, i.e. all the methods ascend, descend and change direction in the same way, there have been

observed some exceptions to this. One example is record 22 of patient 208-night1. The RT drops for the manual and second order fit method but rises for psi and fit. It almost stays the same for msi. This is highlighted in Figure 4.8 below.

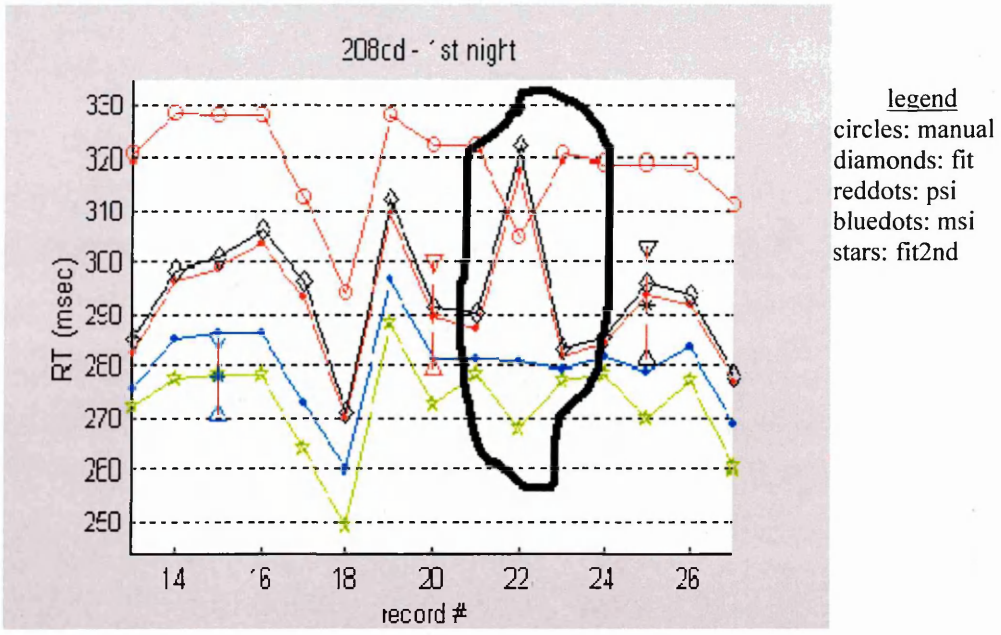


Figure 4.8: RT by psi and fit increases while it decreases for manual and fit2nd

Correlation coefficients

Correlation coefficients were calculated for all methods. Correlation matrices were produced for the 11 subjects. Each correlation matrix contains the correlation coefficient of each RT measurement method with all the other RT methods. As an example, the correlation matrix for patient 205 is given in Table 4.1 below:

Table 4.1: correlation coefficients for patient 205

p205	man	msi	psi	fit	fit2nd
man	1	0.948	0.8255	0.8125	0.6062
msi	0.948	1	0.8515	0.8326	0.557
psi	0.8255	0.8515	1	0.9986	0.5112
fit	0.8125	0.8326	0.9986	1	0.511
fit2nd	0.6062	0.557	0.5112	0.511	1

Each method has a different offset but the differences between the methods are not exclusively due to a difference in offset. This can be seen from the correlation coefficients between the different methods. If the differences between the methods were

purely due to offset then the correlation coefficients would be equal to 1. But this is not the case as seen in Table 4.1.

It must be mentioned here that the correlation coefficient ρ can only identify linear relationships between two variables. Even if the correlation coefficient does not indicate so, a non-linear relationship may exist. Looking at the correlation of the manual methods with all the other methods, it was observed that: the msi method had the highest correlation with the manual method for 7 out of 10 patients. For the other 3 patients the fit2nd method had the highest correlation with the manual method.

The following rankings of the correlations, from highest to lowest, were observed between the manual and the other methods:

- $\rho(man,msi) > \rho(man,psi) > \rho(man,fit) > \rho(man,fit2nd)$, for 3 patients
- $\rho(man,msi) > \rho(man,fit2nd) > \rho(man,psi) > \rho(man,fit)$, for 3 patients
- $\rho(man,fit2nd) > \rho(man,msi) > \rho(man,psi) > \rho(man,fit)$, for 2 patients

Comparison of T end methods using Bland-Altman plots

J. M. Bland and D. G. Altman [Bland 1986] have developed a visual method for assessing agreement between two methods of clinical measurement. This method was included in this study to aid in the comparisons of the different RT algorithms. As stressed by Bland and Altman [Bland 1986], the concept of good correlation should not be confused with the concept of good agreement. Two measurement algorithms measuring the same quantity should ideally agree in the readings they produce and not just correlate well with each other. So generally the Bland-Altman plot is a more useful tool than the correlation coefficient ρ when trying to identify agreement between two methods.

A Bland-Altman plot for comparing two measurement methods is basically a scatter diagram. Each point of the x-axis is the average value between each pair of measurements produced by the two methods studied. Each point of the y-axis describes the difference between the pair of measurements. A scatter diagram of the differences vs the average values is produced. Horizontal lines marking the mean and the ±2sd (standard deviation) limits are also included.

Bland-Altman plots give an idea of how the difference and the spread of readings between two methods change as the average RT between the two methods changes. The Bland-Altman plots that compare the manual method with all the other methods, for patient 201, are given in Figure 4.9.

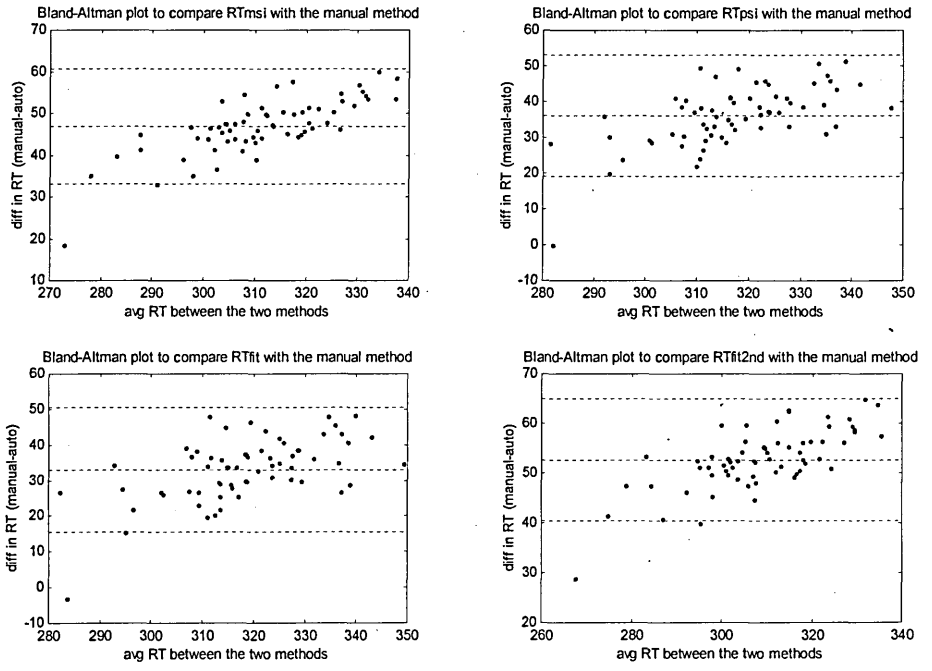


Figure 4.9: Bland-Altman plots for patient 201A

The top-left plot examines the agreement between the manual method and the msi method and the top-right examines the agreement between the manual and the psi. The bottom-left compares the manual with the fit method and the bottom-right the manual with the fit2nd method. The x-axis displays the average RT between the two methods compared while the y-axis displays the difference in RT between the manual and the corresponding automatic method. The dashed lines mark the mean and ±2sd limits. For patient 201A all the data points, but one, lie within the 2sd limits. This happens for all four plots of this patient. It can be seen that the difference between the manual and each other method increases as the average RT increases. This means that the disagreement between the manual and each automatic method in turn increases as the length of RT to be measured increases. This piece of information is not provided by the correlation coefficient. To remove this trend from the graphs a logarithmic transformation can be used, before the data is plotted.

It is apparent from the graphs that the mean difference (marked by the middle dashed line) lies well above zero in all the cases. This is because of the different levels of offset in the methods. If the agreement between the manual and each other method was very

close, then the mean-difference-line should lie very close to zero. The best agreement is found for patient 204 and for the psi and fit algorithms. For most of the patients the magnitude of the difference is quite significant compared to the magnitude of the RT intervals. Because Bland-Altman plots comprise a visual method for assessing correlation and agreement, correlation coefficients were also useful since they express the level of correlation numerically.

Final discussion on comparative study

The manual method gave the longest RTs for all patients and all nights. There were 9 cycles as exceptions to this, out of the 660 cycles studied. The tangent (msi) method was the one that correlated mostly with the manual method according to what the correlation coefficient indicates. The psi and fit methods correlate very well with each other. The correlation coefficients between the two are the highest ones observed in the study. The performance of the first order fitting method could possibly be improved by changing the portion of the T downslope used to fit the best straight line. Different ranges could be used to optimally tune the algorithm. This was not undertaken due to time constraints. Another fact observed was that the differences between the algorithms were not purely due to offset as the correlation coefficient indicated.

There was a need to choose a T end detection algorithm to be used as part of the feature extraction process. The automatic tangent method was chosen as the method to use. This method correlated the highest with the manual method which is the current gold standard. Moreover the semi-automatic version of the tangent method has been already used for manual feature extraction by medical experts [Ireland 1998, 2000] which indicates that this method is already accepted in the medical community.

4.4.3 Evaluation of the Symmetry and Morphology of the T wave

One of the main motivations in performing ECG feature extraction was to quantify the level of symmetry and the morphology of the T wave. A number of features were proposed and the relevant algorithms used to extract these features were designed. Three novel features were introduced: the Half-Areas Ratio, the T wave skewness and the T wave kurtosis feature. The latter two features were inspired from the 3rd and 4th central moments, used in statistical theory to evaluate the symmetry and peakedness of distributions. In our case the definitions of skewness and kurtosis were adopted to evaluate the morphology of the T wave.

Half-Areas Ratio (HAR) algorithm

This algorithm was designed for producing a basic symmetry ratio for the T wave. The ratio of the areas to the left and right of the T wave was used and the feature was termed HAR (Half-Areas Ratio). A similar symmetry feature had been used in the past by Benhorin et al [Benhorin 1990] to evaluate the symmetry of the T wave. Benhorin's Symmetry Ratio (SR) was different to our version (HAR) and was involving the ST segment. As mentioned in Section 2.4.2, Benhorin's symmetry ratio was defined as the ratio between the integrated area over SoTm and TmTo intervals (SR= SoTm/TmTo). In our case we wanted to avoid involving the ST segment in the quantification of T wave symmetry. A symmetry measure was introduced that involved only the T wave portion, enclosed by T onset and offset. The Half-Areas Ratio algorithm calculates the ratio of the areas to the right and left of the T peak and can give us a simple measure of the symmetry of the T wave. The HAR feature is given by the formula:

$HAR = Area_{RHS} / Area_{LHS}$

If the two areas are equal then the ratio is 1. If the area to the right of T peak ($Area_{RHS}$) is greater than $Area_{LHS}$ then the ratio is greater than 1 and vice versa[15]. Figure 4.10 helps illustrate the concept behind the algorithm. Assuming an upright T wave, the two areas involved are calculated as follows: the left-hand-side area ($Area_{LHS}$) is left-delimited by the vertical line going through Tonset, right-delimited by the vertical line going through

[15] Benhorin's symmetry used a ratio of the LHS upon the RHS area. We chose the inverse (RHS/LHS) so that the HAR value would increase for T waves skewed to the right and decrease for T waves skewed to the left. This is in agreement with how the skewness of a distribution works and allows easier inspection of feature values by a human observer.

T peak, bottom-delimited by the isoelectric line and top-delimited by the ECG portion corresponding to the T wave upslope. Replacing the Tonset by Toffset and the T wave upslope by the downslope produces the definition of the right-hand-side area ($Area_{RHS}$).

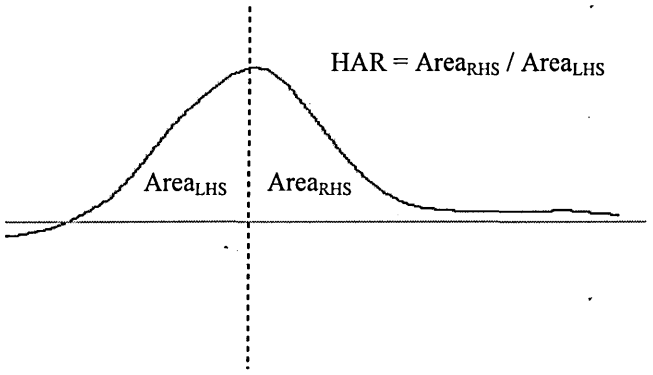


Figure 4.10: Truncated T wave to demonstrate Half-Area Ratio

Figures 4.11 and 4.12 illustrate two T waves from patient 202-night1 (202A) that possess the extreme values of HAR for the given patient. The onset and offset of the T wave were calculated by the tangent method (msi).

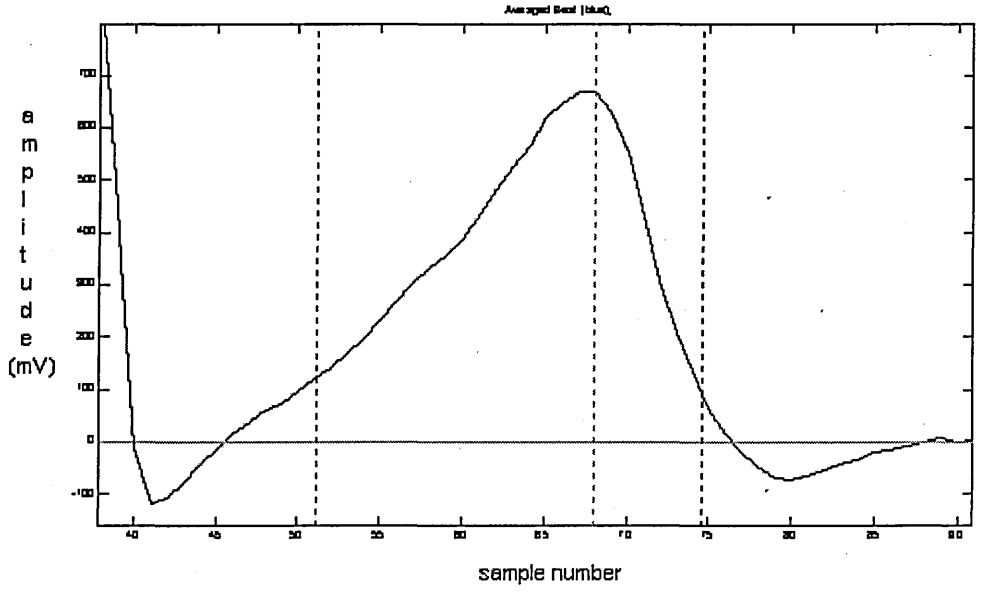


Figure 4.11: T wave of 202A record 14 having a HAR value of 0.547

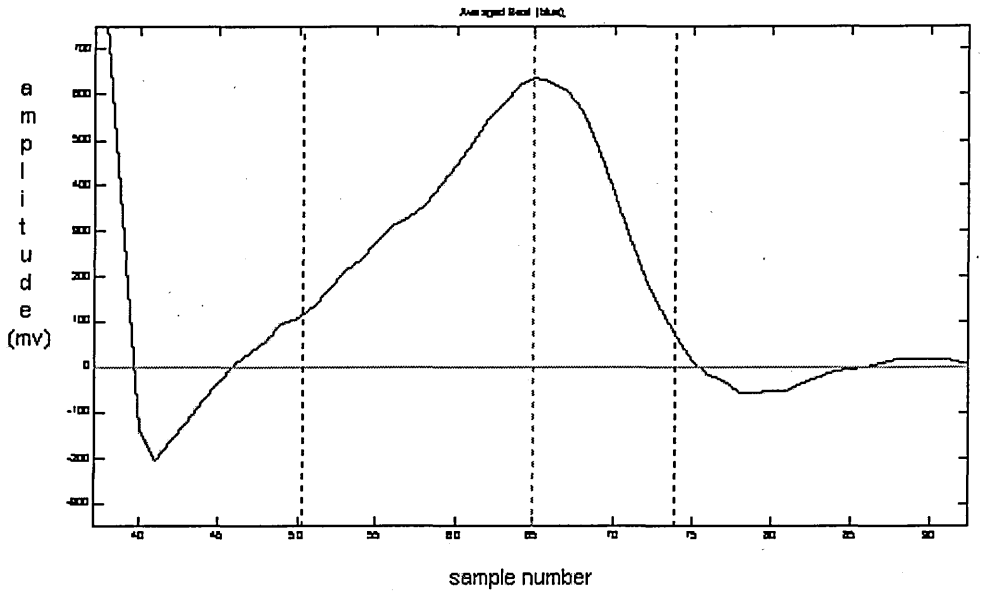


Figure 4.12: T wave of 202A record 16 having a HAR value of 0.6605

The HAR variation for this night shows only subtle changes. All T waves in the night possessed a LHS area, greater than the RHS area. The change in T wave symmetry according to HAR was small although the above night contained hypoglycaemic records.

The HAR feature is dependent on the onset and offset annotations of the T wave. Using different algorithms to mark the onset and offset of the T wave will yield different values of the feature. For instance, if we consider two identical T waves, with the second T wave having its end located a few msec to the right compared to the first T wave, then the HAR value will be different for the two waves although they will have the same symmetry. In order to investigate the sensitivity of HAR to the annotation algorithms, the feature was calculated, based on 3 different annotation algorithms.

Figure 4.13 illustrates the HAR feature values from patient 209 when the T wave offset is calculated in turn by the three algorithms (msi, psi, fit) described in Section 4.4.1.

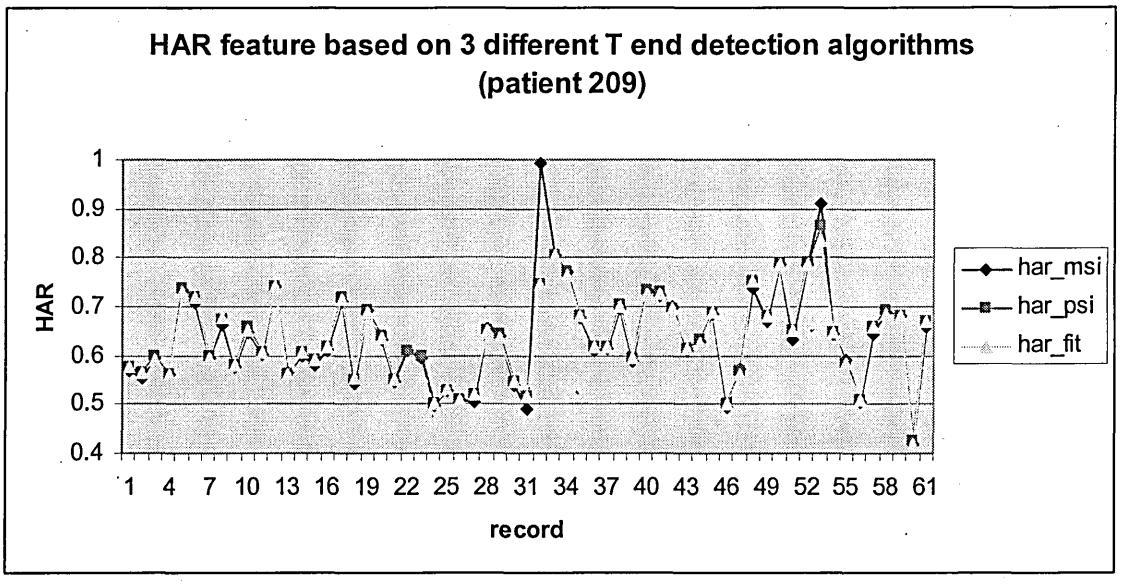


Figure 4.13: HAR feature based on 3 different T end detection algorithms (patient 209)

The T onset is calculated by the msi algorithm (tangent method) to avoid introducing algorithmic variations on the annotation of the onset. This allows investigation of the effect of the use of different offset algorithms. It can be seen on the graph that the HAR profile depends on the annotation algorithm used. Even if the same algorithm is used for annotation of onset and offset, the existence of some variation will be possible. For instance, the behaviour of the tangent method will be different on the T upslope and downslope because they have different slopes. The following section discusses approaches that were followed in order to reduce the dependence of the HAR algorithm on the T onset/offset annotations.

Reducing the dependence of HAR on the T onset/offset annotations

A number of steps were taken to make the HAR feature (and also the skewness and kurtosis features, presented later) as independent as possible to the T onset/offset annotations. One approach was to make the feature dependent on only one annotation, either the onset or offset, instead of two. This could by achieved by defining the onset and offset of the T wave in such a way that they would lie at the same amplitude (in mV). This would define a horizontal threshold level and only the portion of the T wave above this threshold would be used for symmetry/morphology calculations. This would contribute in overcoming variations of the HAR feature due to the onset/offset annotations being at different amplitudes. The point (either onset of offset) being at the highest amplitude was extrapolated (i.e. projected) at the other side of the T wave. For instance, if the T offset was at a higher amplitude than the T onset, then the T offset

would be extrapolated on the T upslope and would define a new onset for the wave (T-onset^{extrap}).

The opposite was done when the onset was at a higher amplitude than the offset; in that case the onset was extrapolated on the other side. What is achieved by the extrapolation process is that the value of each of the three symmetry/morphology features depends only on one annotation (either onset of offset) rather than two. The new onset/offset defined by projection from the other side of the wave, would only be used for the purposes of calculating the HAR feature and the other two features assessing symmetry (skewness, kurtosis). The extrapolated onset/offset would not be used for calculating other features e.g. time interval ones (T-duration etc).

The above process is illustrated in Figure 4.14.

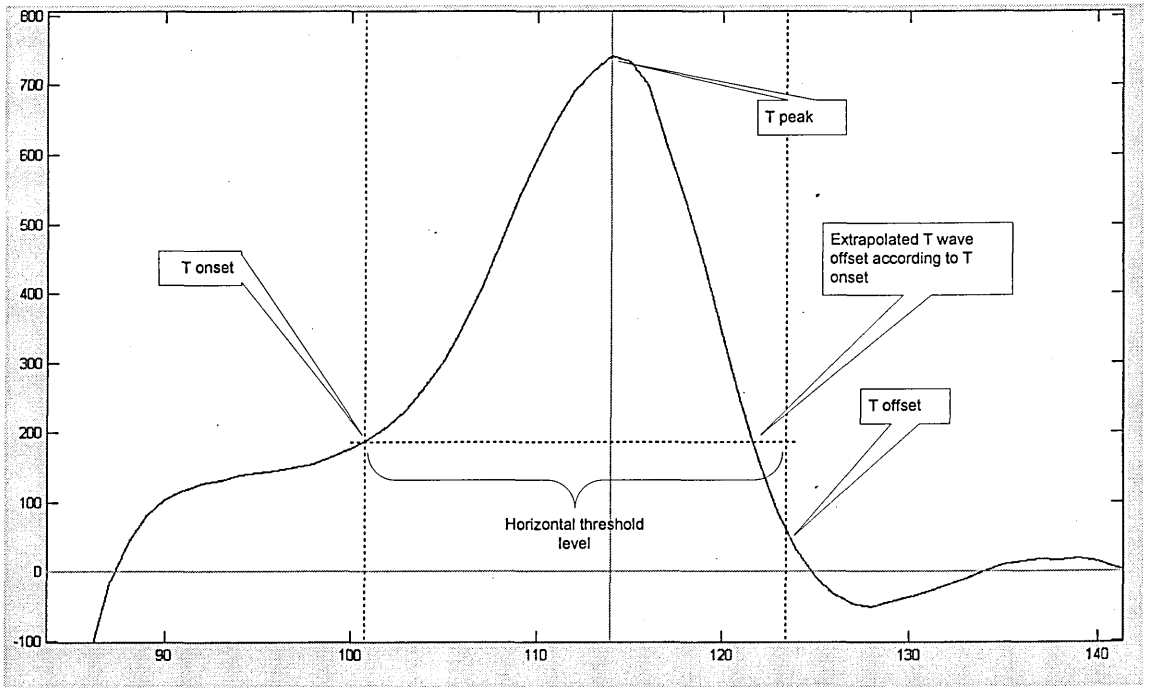


Figure 4.14: An illustration of extrapolating the T onset on the T downslope (p209 record 6)

The figure shows the projection of T onset on the T downslope. The T onset annotation lies at a higher amplitude than the T offset so it is the one to be projected on the other side to define a horizontal threshold level. The extrapolated-HAR (HARx) feature for this T wave will be calculated based on T-onset and T-offset^{extrap}. When the T wave offset lies at a higher amplitude it is extrapolated on the T upslope to define a new onset.

A second approach used to evaluate the symmetry of the T wave based on the HAR feature, with no dependence on the onset and offset annotations, was to use the point of inflection on the T downslope to define the horizontal level. Only the portion of the T wave above this horizontal level was used to calculate the symmetry of the wave.

In the approaches using a horizontal level to set a lower threshold for selecting the T wave portion for morphology calculations, we were faced by problems due to low sampling rates of the ECG signal. The closest ECG samples to the horizontal threshold chosen, had a significant difference in amplitude between them, as illustrated in Figure 4.15 and would almost never lie at the same horizontal level. This partly defeats the purpose of using a horizontal level as a step to improve the T wave morphology calculations. Although the difference in amplitude between the onset and offset annotations is reduced, it is not eliminated. Even for high sampling rates, the amplitudes of the offset and extrapolated onset, or vice versa, will be similar but not identical. In the case of our dataset the sampling rate was low (125 Hz) and the above problem was more prominent. One solution to this would be to use interpolation for upsampling the data to a higher sampling rate. Due to the low sampling rate in our dataset, the amount of interpolation needed was high and this was, in some cases, causing some distortion to the signal. Although the distortion was not severe, the approach of interpolating was not followed. T wave symmetry changes can be subtle and even the slightest distortion due to interpolation could contaminate the features.

Boxplots are presented in Figure 4.16 to allow comparisons of the three versions of the HAR feature. The boxplot [Tukey 1977] is a very useful tool for data inspection and will be used in the remaining of the chapter as part of the ECG feature analysis. A box represents the inter-quartile range (mid-50% of the data) while the median is also marked with a line across the smaller dimension of the box. The whiskers display the extent of the remaining of the data. Outliers are represented by "+" as seen on the graph. The notches in the box are a graphic confidence interval about the median of a sample. Box plots do not have notches by default. The use of notches is a feature of the boxplots produced by the Statistics toolbox [MathWorks Statistics] in MATLAB.

Figure 4.15 compares the three versions of the HAR feature for the conditions of normality[16] and hypoglycaemia. The LHS graph presents the raw feature, the middle one presents the HAR feature after projection of the T onset or offset on the other side (HARx) and the RHS figure presents the HAR feature when the T wave portion used is selected by the point of inflection of the downslope (HARxIP). It can be seen that all three features have similar behaviour. In all three cases there are big differences between euglycaemia and hypoglycaemia and the feature changes appear to be statistically significant, according to the notches plotted. There is no overlap of the notches for the boxplots corresponding to normality and hypoglycaemia. Hypoglycaemia also appears to have significantly larger ranges of values compared to normality. Comparing the normal records for all three features, it is seen that all three boxes are quite symmetric with the raw HAR feature having the largest range. Comparing the hypoglycaemic records for all three features, it is observed that the HARx and HARxIP features are significantly skewed to the left (i.e. bottom in this plot) compared to the raw HAR feature which is slightly skewed to the right (i.e. top in this plot).

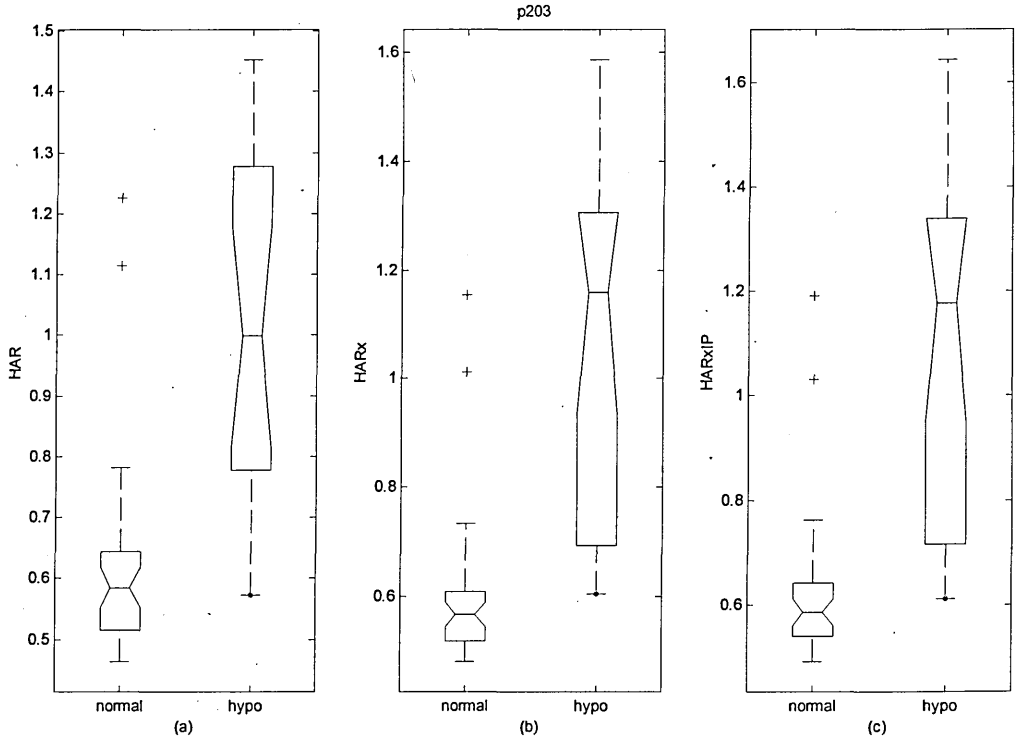


Figure 4.15: Boxplots for the 3 versions of the HAR feature for patient 203 (both nights): (a) HAR, (b) HARx, (c) HARxIP

[16] Including hyperglycaemic data if present.

Although the above boxplot was informative about the behaviour of the HAR feature, it does not provide temporal information. In order to investigate the dynamic changes with respect to time and in relation to the changing glucose, Figure 4.16 is provided. The top graph presents the three versions of HAR vs time. The bottom graph presents the glucose levels. Two successive nights for this patient are given. The vertical dashed line in black splits the two nights. The horizontal dashed line in black, at the bottom graph marks the hypoglycaemic threshold of 3 mmol/lt.

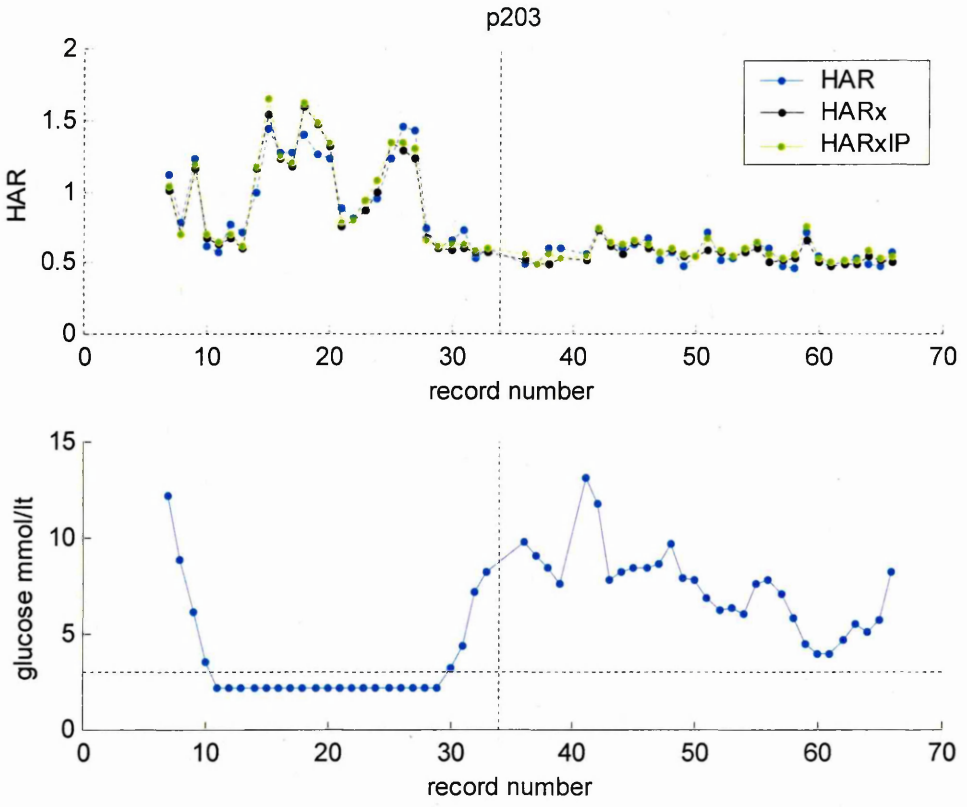


Figure 4.16: HAR features and glucose level versus time for 2 successive nights of patient 203

The 3 versions of the HAR feature have similar performance with the HARx and HARxIP being very close together. The graph indicates that this feature, in all its versions, is a very informative one in relation to hypoglycaemia detection. The first night was hypoglycaemic while the second was normal (including some hyperglycaemic records). The HAR feature (all three versions) had very small variation for the second night (RHS of the vertical dashed line) while it had great variation during the first night which was hypoglycaemic as can be seen from the glucose profile in the bottom graph.

It is emphasised that the performance of the various ECG features, as predictors of hypoglycaemia, is expected to vary from patient to patient. The HAR feature for patient 227 (both nights) is presented in Figure 4.17. It can be seen that there are clear changes

between the two clinical conditions, but the changes are not as prominent as those of patient 203 presented in Figure 4.16. Comparing the three versions of the HAR feature for patient 227, it is observed that the HARx was the worst version of the feature judging by the overlap that the notches of the boxplot have.

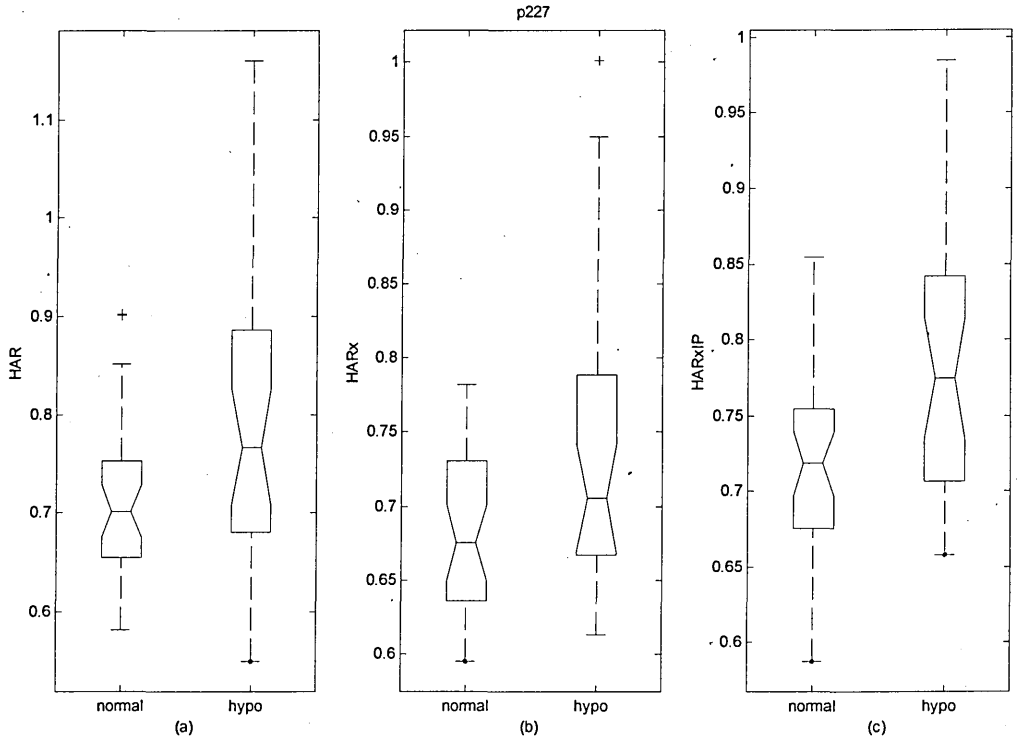


Figure 4.17: Boxplots for the 3 versions of the HAR feature for patient 227 (both nights): (a) HAR, (b) HARx, (c) HARxIP

Assessing T wave symmetry using the concept of skewness

As part of the efforts to devise algorithms that will assess the T wave morphology, an algorithm was designed that uses a normalised form of the 3rd central moment (skewness) to assess the symmetry of the T wave. An analogy was formulated that was considering and assessing the symmetry of T waves in a similar way to how the symmetry of a statistical distribution is described by the normalised 3rd central moment. Skewness is calculated by dividing the 3rd central moment, about the mean, by the cube of the standard deviation as described by the formula: $\gamma_1 = \frac{E(x-\bar{x})^3}{s^3}$ eq^n (4.1)

where E denotes the "expected value" operator.

This is the skewness of a random variable X with sample mean $\bar{x}$ and sample standard deviation s. It describes the symmetry of the distribution of X. Skewness is negative when the distribution is skewed to the left and positive when it is skewed to the right. A

skewness of zero corresponds to a perfectly symmetric (i.e. Gaussian) distribution. Skewness in this study is not used for quantifying the symmetry of the distribution of a given ECG feature but rather to evaluate the symmetry of the T wave shape. The T wave shape resembles the shape of a bell-shaped distribution which justifies the choice of the concept of skewness to assess the symmetry of the T wave.

In order to evaluate its symmetry, the T wave component is truncated from the rest of the ECG trace and is treated as a functional form. Its shape is treated as if it is the shape of the frequency curve of a, highly-sampled, discrete distribution. The skewness of this functional form is calculated and, effectively, we obtain a measure of symmetry of the T wave. (We utilise the term *frequency curve* instead of *probability curve* because the T wave curve originates from a sampled signal and, more importantly, it is not normalised to have unity area underneath[17].)

Let us consider a random variable X. $f(x)$ is the frequency of occurrence of X at value x and $\int_{-\infty}^{+\infty} x \cdot f(x)\,dx$ eq^{n} (4.2)
describes the area under the frequency curve. When the curve is normalised to have unity area underneath ($\int_{-\infty}^{+\infty} x \cdot f(x)\,dx = 1$) then $f(x)$ is referred to, as the probability density function (p.d.f.) of x.

In our case, and in order to calculate the symmetry of the T wave, X describes the time (x-axis variable) and $f(x)$ the corresponding voltage of a given sample of the ECG trace (y-axis variable). Using this analogy and for an analogue ECG signal, equation 4.2 becomes: $\int_{Tonset}^{Toffset} t \cdot ECG(t)\,dt$ eq^{n} (4.3)
where t describes the time at any time instant and ECG describes the corresponding amplitude of the ECG trace (only for the T wave portion of the ECG) for this time instant. In the discrete domain this can be written: $\sum_{i=Tonset}^{i=Toffset} t_i \cdot ECG(t_i)$ eq^{n} (4.4)

Using this analogy, we are calculating the normalised 3rd central moment of the T wave sample points about the T wave peak and this leads to a new ECG feature for T wave symmetry evaluation. The analogy used is illustrated in Table 4.2:

[17] The term *probability curve* refers to continuous distributions with unity area underneath.

Table: 4.2: Analogy introduced to allow calculation of T wave skewness and kurtosis

Analogy	
statistical theory	**assessment of T wave morphology**
random variable X	discrete time t_i (sampling instants at which T wave is sampled)
frequency of occurrence of X, (f(X))	T wave portion of ECG signal, $ECG(t_i)$ (mV)
mean of X	temporal position of T wave peak
variance of X	2nd moment of ECG sample points around the T wave peak
standard deviation of X	square root of the 2nd moment of ECG sample points around the T wave peak
skewness of the distribution of X	3rd moment of ECG sample points around the T wave peak normalised by the cube of the standard deviation
kurtosis of the distribution of X	normalised 4th moment of ECG sample points around the T wave peak

According to the analogy, the mean value of the distribution corresponds to the T peak. The standard deviation formula expresses the square root of the 2nd moment of the ECG sample points around the T peak (σ_{Tpeak}), and the skewness expresses the normalised 3rd central moment around the T peak. t_{Tpeak} denotes the temporal position of the T wave peak. The algorithm assessing the kurtosis of the T wave is based on the same analogy and will be discussed in Section 4.4.3.

According to the above analogy, the skewness formula becomes:

$$\text{T skewness} = \frac{\sum_{i=1}^{N}(t_i - t_{Tpeak})^3 \cdot ECG(t_i)}{\sum_{i=1}^{N} ECG(t_i)} / \sigma_{Tpeak}^{\,3} \qquad \text{eq}^{\text{n}}\ (4.5)$$

$$\text{where } \sigma_{Tpeak} = \sqrt{\frac{\sum_{i=1}^{N}(t_i - t_{Tpeak})^2 \cdot ECG(t_i)}{\sum_{i=1}^{N} ECG(t_i)}} \qquad \text{eq}^{\text{n}}\ (4.6)$$

The analogy was based on equations from population statistics (using true mean and standard deviation) as opposed to the equations from sample statistics. When using equations from sample statistics, and for a sample of N points, the denominator for the

above equations would be $\sum_{1}^{N} ECG(t_i) - 1$ instead of $\sum_{1}^{N} ECG(t_i)$, i.e. one degree of freedom (DOF) would be subtracted. Our situation although being only an analogy to a statistical problem, resembles a problem where the whole population is considered instead of a sample of the population. The T wave is abstracted as the shape of a true distribution and not as a sample distribution producing an estimate of the true distribution of the population. Each T wave shape assessed, purely describes itself and is not an estimate of any other T wave. In other words, we know the temporal position of the T wave peak instead of calculating it from the T wave data, which is analogous to knowing the true mean of a distribution, in a statistical problem.

The behaviour of the skewness algorithm was tested experimentally. The choice of the exact form of the skewness equation being the one originating from population statistics, was also tested. A perfectly symmetric T wave was constructed to investigate the above. This was an artificially synthesized T wave but was based on real ECG data. The upslope of a nearly symmetric T wave (p203 record7) was chosen and its mirror image, along the vertical, was produced. The two shapes were joined together to produce a perfectly symmetric T wave. This is shown in Figure 4.18.

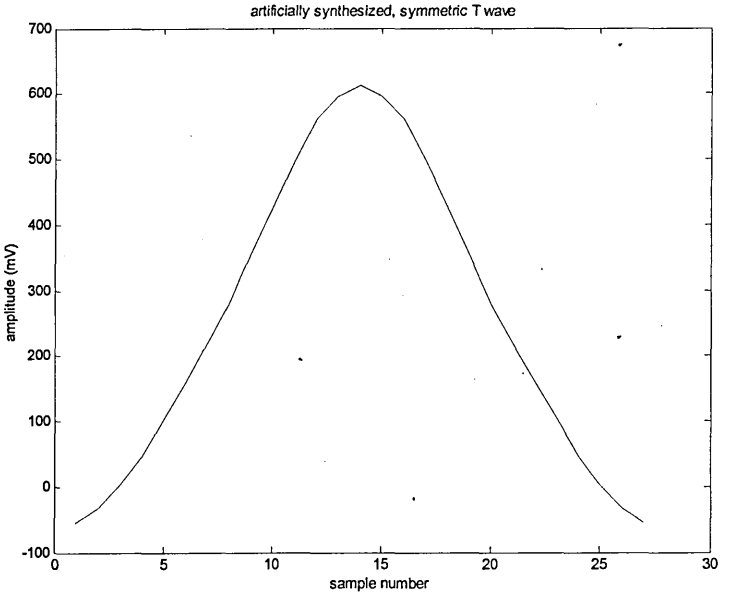


Figure 4.18: artificially synthesized T wave, exhibiting perfect symmetry

The HAR feature for this T wave was 1 as expected. The skewness feature (SKEW) for this wave for N-1 DOFs (i.e. sample statistics) was $-6.9831 \cdot 10^{-17}$ while it was $-6.9821 \cdot 10^{-17}$ for N DOFs (i.e. true population statistics). It can be seen that the

difference between the two values of skewness is negligible. Values of skewness in the order of 10^{-17} are practically zero, which demonstrates that the skewness algorithm assessed correctly the perfectly symmetric T wave corresponding to a theoretical skewness value of zero. It also demonstrated that the choice of the number of DOFs did not affect the result. This is due to the fact that the T wave shape assessed was not normalised to have unity area. Normalising to unity area leads to a value of skewness in the order of 10^{-15} when using N-1 DOFs, while the skewness values using N DOFs remains in the order of 10^{-17}.

The skewness algorithm needs the T onset and T offset annotations in order to function. The onset and offset specify the subsection of the ECG trace to be used. The skewness algorithm for calculating the symmetry of the T wave is illustrated in Figure 4.19 that displays two truncated T waves of calculated skewness -0.049 and 0.352 respectively. The data originates from patient 204 and the traces illustrated in the figure are records 9 and 26 respectively which exhibited the extrema of skewness for this patient. The T wave in record 9 is slightly skewed to the left according to the skewness algorithm while the T wave in record 26 is skewed to the right. The green line marks the T peak while the black dashed lines to the left and right of the peak mark the onset and offset of the wave respectively. Taking in account the T onset and offset it is apparent from the figures that the LHS graph looks quite symmetric while the RHS one is skewed to the right. Obviously the calculation of skewness depends heavily on the annotation of the T wave onset and offset. For instance if the T wave offset on the RHS graph above was marked a few msec to the left of its current position, the value of skewness calculated would appear to be less positive.

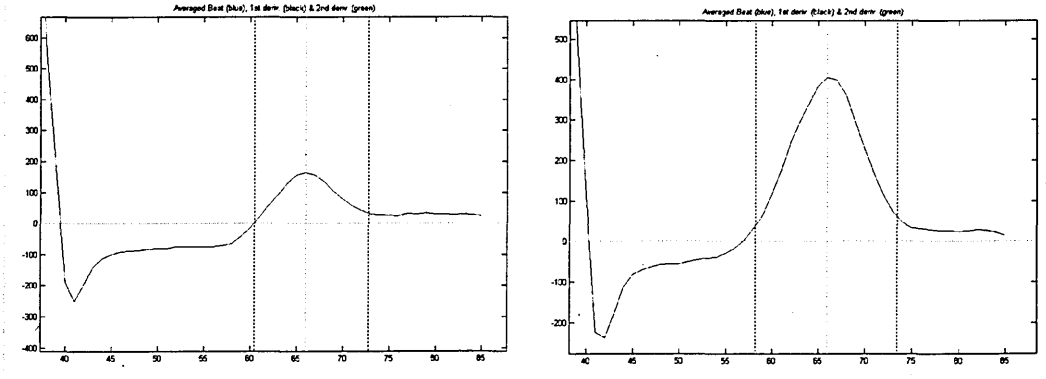


Figure 4.19: (LHS) p204rec9 (skew=-0.065); (RHS) p204rec26 (skew= 0.291)

In order to analyse this phenomenon, the skewness was calculated for each record of patient 204 using 3 different T end annotations by 3 different algorithms, namely MSI

(tangent method), PSI and FIT. This was a similar analysis to the one assessing the effect of the T onset and offset annotations on the HAR feature. In order to allow comparisons of the effect of T end, the T onset was calculated by the MSI algorithm in all three cases. We are focusing on the effect of T end annotation on the calculation of skewness rather than the effect of T onset annotation because the former is the one that constitutes a more difficult algorithmic problem. Annotating the T onset is an easier task since the waveform is likely to be less affected and distorted by dropping glucose. Hence it is crucial to choose a T end detection algorithm that releases the full potential of the skewness feature in evaluating the symmetry for the wave. For instance an algorithm that overestimates the QT interval, i.e. it marks the end of T always to be to the right of where it should be, will tend to present the T wave being skewed to the right which will mask the real symmetry of the wave.

The skewness calculated using the three T end algorithms is tabulated in Table 4.3. The table includes the record numbers and the glucose levels as well.

Table 4.3: Skewness based on three different algorithms of T end detection

rec	gl	skew_msi	skew_psi	skew_fit
5	8.48	0.183	0.632	0.632
6	8.48	0.302	0.415	0.415
7	7.09	0.233	0.451	0.451
8	5.08	0.017	0.177	0.177
9	5.41	-0.049	0.282	0.282
10	5.66	-0.031	0.022	0.022
11	6.14	0.014	0.091	0.091
12	5.23	0.273	0.388	0.388
13	4.91	0.040	0.300	0.400
14	4.69	0.023	0.202	0.202
15	4.06	0.047	0.112	0.112
16	4.09	0.052	0.138	0.138
17	2.82	0.225	0.475	0.475
18	2.20	0.259	0.503	0.503
19	2.20	0.228	0.559	0.648
20	2.20	0.179	0.357	0.357
21	2.20	0.220	0.436	0.544
22	2.20	0.301	0.631	0.631
23	2.20	0.272	0.364	0.364
24	2.20	0.040	0.133	0.218
25	2.20	-0.023	0.128	0.128
26	2.20	0.352	0.470	0.586
27	2.20	0.306	0.419	0.419
28	2.20	0.181	0.289	0.289
29	2.20	0.274	0.386	0.386
30	2.76	0.222	0.556	0.658
31	2.20	0.035	0.122	0.122
32	2.20	-0.009	0.138	0.138
33	2.20	0.040	0.187	0.187

The contents of the table are also presented as a graph in Figure 4.20.

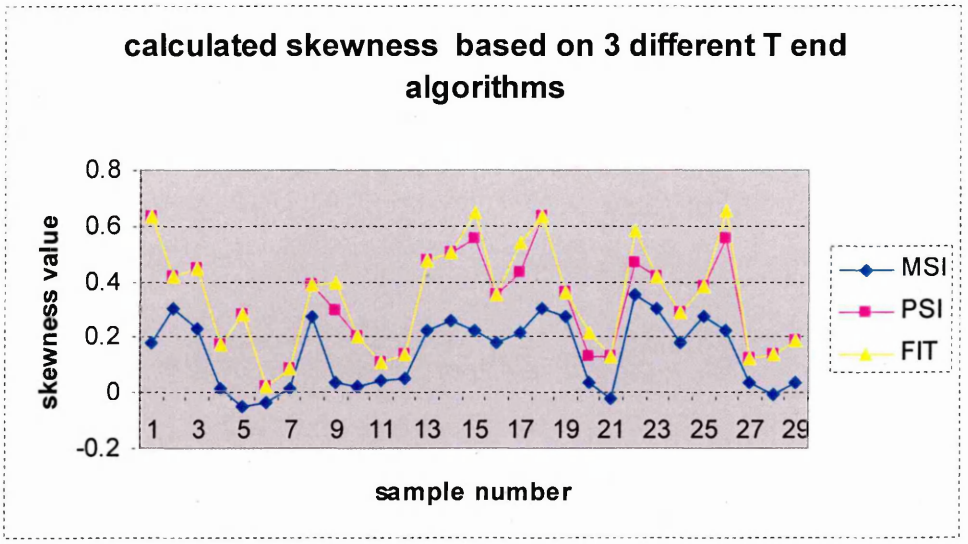


Figure 4.20: Skewness according to three different algorithms of T end detection

It can be seen that the skewness calculated by the peak-slope-intercept (PSI) algorithm matches very closely the skewness calculated by the fitting algorithm (FIT). The dc component of the skewness calculated by the tangent method (MSI) is significantly less. According to this algorithm the T wave appears to be negatively skewed (i.e. to the left) in more samples than with the other two algorithms. The other two algorithms overestimate the RT interval and hence they cause most T waves to be assessed as skewed to the right.

In order to aid the investigation of the effect of the T end annotation on the calculation of skewness, the correlation coefficient between glucose and the values of symmetry by the three different T end annotation algorithms (MSI, PSI and FIT) is calculated. The correlations are tabulated in Table 4.4.

Table 4.4. Correlation coefficients between glucose and skewness, based on three different algorithms.

	correl(gl,msi)	correl(gl,psi)	correl(gl,fit)
glucose	-0.1311	-0.0041	-0.05814

The correlation coefficients calculated are very low for all algorithms but they suggest that the skewness calculated by the MSI algorithm correlates less badly with the glucose variable.

Advantages of the skewness feature (SKEW) over the HAR feature

Both these features are used to assess the symmetry of the T wave. A question could easily be raised as to why use the skewness feature and what are the advantages of this feature over the simpler HAR feature. The limitation of the HAR feature is that it quantifies symmetry simply by using the areas of the T wave halves to the left and right of the T peak. This feature will not detect the asymmetry of a wave whose halves to the left and right of the T peak have the same area. On the other hand, the skewness feature does not depend on the areas for assessing the symmetry and will distinguish the asymmetry in the above example. This can be easily understood by considering the simple illustration in Figure 4.21. The illustration depicts a piecewise linear representation of the T wave. The LHS of the T wave is defined by a right-angle triangle and the RHS by a trapezium joined with a right-angle triangle that simulates the case when the T wave is skewed to the right. The points corresponding to the T onset, offset and peak are marked on the figure. By considering the lengths marked on the figure (arbitrary units) the HAR value can be calculated:

$Area_{LHS} = (9*10)/2 = 45$

$Area_{RHS} = (10 + 2)*6/2 + (9*2)/2 = 45$

$HAR = Area_{RHS} / Area_{LHS} = 1$

Calculating the skewness of the piecewise linear T wave[18] we get 0.7804, which is significantly different to a skewness value of zero that would correspond to a perfectly symmetric T wave. The above value for skewness proposes a significantly skewed to the right, T wave. It can be easily concluded that the HAR feature will fail to demonstrate the asymmetry of any T wave that possesses equal areas to the left and right of its peak. This introduces a need for a feature that will be able to identify such an asymmetry. The proposed skewness feature demonstrates that it can identify such asymmetries.

The skewness feature could prove a useful one in detecting T wave changes under hypoglycaemia but it should not replace the HAR feature. The two features measure different quantities and could both be used in order to obtain a more complete measure of the morphology of the T wave.

[18] The shape in the figure was sampled at 0.0001 intervals for calculation of the skewness value.

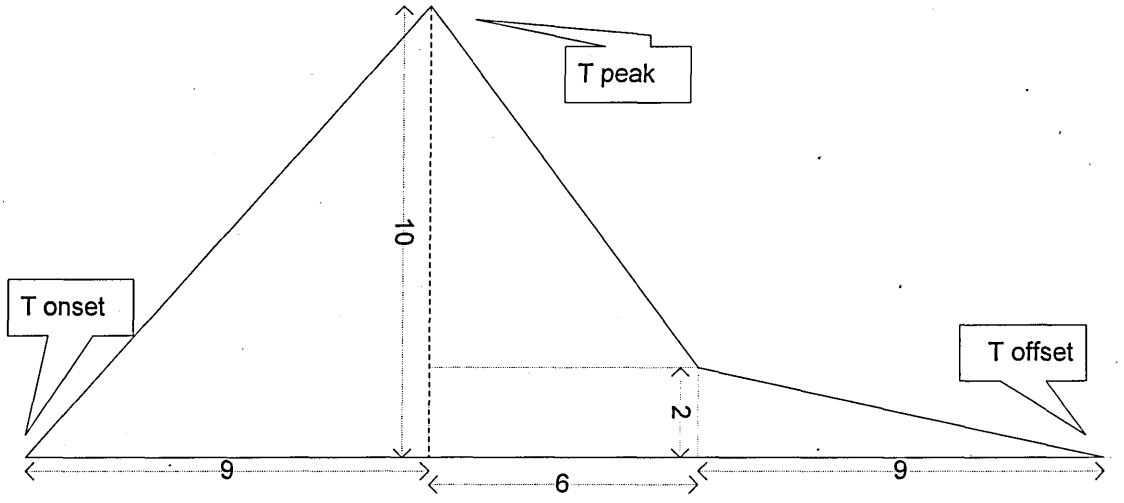


Figure 4.21: Piecewise linear representation of the T wave for comparison of HAR and SKEW

Reducing the effect of T onset/offset annotations on the SKEW feature

The approach that was used for reducing the dependency of the HAR feature on the onset and offset of the T wave was also applied in the case of the SKEW feature. Three versions of the feature were produced: the raw feature (SKEW), a version where the onset or offset was extrapolated on the other side of the wave (SKEWx) and a version of the feature that was using the point of inflection on the T downslope to define the portion of the T wave to be used (SKEWxIP). Boxplots are presented in Figure 4.22 to allow comparisons of the three flavours of the feature.

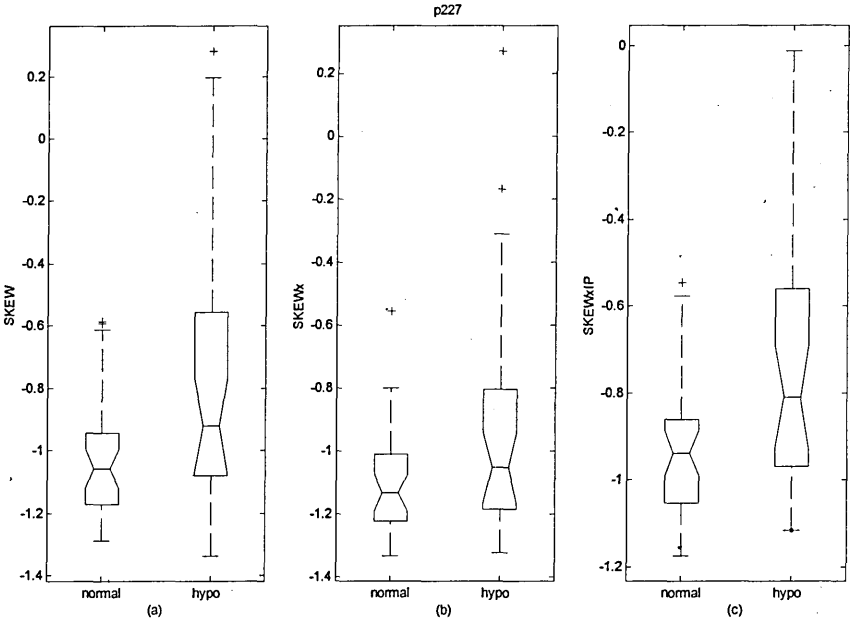


Figure 4.22: Boxplots for the 3 versions of the SKEW feature for patient 227 (both nights) (a) SKEW, (b) SKEWx, (c) SKEWxIP

It can be seen from the graph that all versions of the feature have quite similar behaviour. The range and inter-quartile range increases in all three cases, under hypoglycaemia. For this patient, choice of one of the three features to be used will not be crucial to the performance of a classification system.

Assessing T wave morphology using the concept of kurtosis

The kurtosis of a statistical distribution is calculated by the formula:

$$\beta_2 = \frac{E(x-\bar{x})^4}{s^4} \qquad \text{eq}^{\text{n}}\ (4.7)$$

This is the kurtosis of a random variable X with sample mean $\bar{x}$ and sample standard deviation s. It describes the degree of peakedness of a distribution, defined as a normalized form of the fourth central moment. A distribution with a high peak is called leptokurtic ($\gamma_2>0$), a flat-topped curve is called platykurtic ($\gamma_2 < 0$), and the Normal distribution is called mesokurtic ($\gamma_2=0$). γ_2 is the kurtosis excess [Kenney 1951] defined as kurtosis minus three ($\gamma_2 = \beta_2 - 3$). Since the kurtosis of the Normal distribution is 3, using the kurtosis excess is more convenient since it is zero-valued for the Normal distribution. This justifies the subtraction of the value of 3 and the convenience of using the kurtosis excess quantity instead of the kurtosis.

Similarly to the skewness feature, the concept of kurtosis was not used for assessing the shape of a distribution but instead for assessing the T wave shape through the analogy introduced earlier. Again the T wave component is truncated from the rest of the ECG trace and its shape is assessed in the same way that the probability curve of a probability distribution is assessed using the kurtosis formula from statistical theory. The calculated kurtosis of a given T wave gives a measure of the morphology of the T wave.

According to the analogy between the shape of a distribution and the T wave shape, presented in Table 4.2, the kurtosis formula becomes:

$$\text{T kurtosis} = \frac{\sum_{1}^{N}(t_i - t_{Tpeak})^4 \cdot ECG(t_i)}{\sum_{1}^{N} ECG(t_i)} \Big/ \sigma_{Tpeak}{}^4 \qquad \text{eq}^{\text{n}}\ (4.8)$$

Similarly to the equations used for calculation of skewness, this is the equation corresponding to true population statistics (when the true mean μ and true standard deviation σ are used).

Figure 4.23 illustrates the T waves corresponding to the minimum (LHS) and maximum (RHS) kurtosis for patient 204.

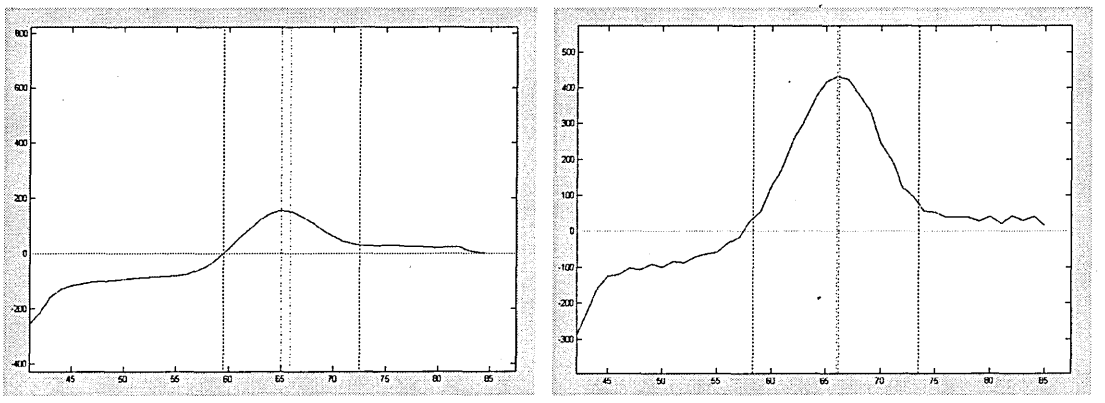


Figure 4.23: (LHS) p204rec22 (kurt=2.224) and (RHS) p204rec24 (kurt= 2.464)

It is observed that the range of the kurtosis feature is very small (0.239) compared to its mean value (2.385) but even differences in skewness of around 0.2 correspond to very differing T wave morphologies as seen in the above figure. If we calculate the kurtosis excess, it will be -0.615 for record 22 and -0.536 for record 24 which informs us that both T wave shapes are classed as platykurtic.

The approach used for reducing the effect of onset and offset annotations on the symmetry features were also applied to the kurtosis excess feature. The three versions of the kurtosis excess feature are compared in the boxplot presented in Figure 4.24.

The figure suggests significant changes for all three versions of the feature in response to abnormal glucose changes. The limits of the inter-quartile ranges change significantly between normality and hypoglycaemia. The figure also suggests that T waves become more platykurtic under hypoglycaemia which is in line with the research hypothesis.

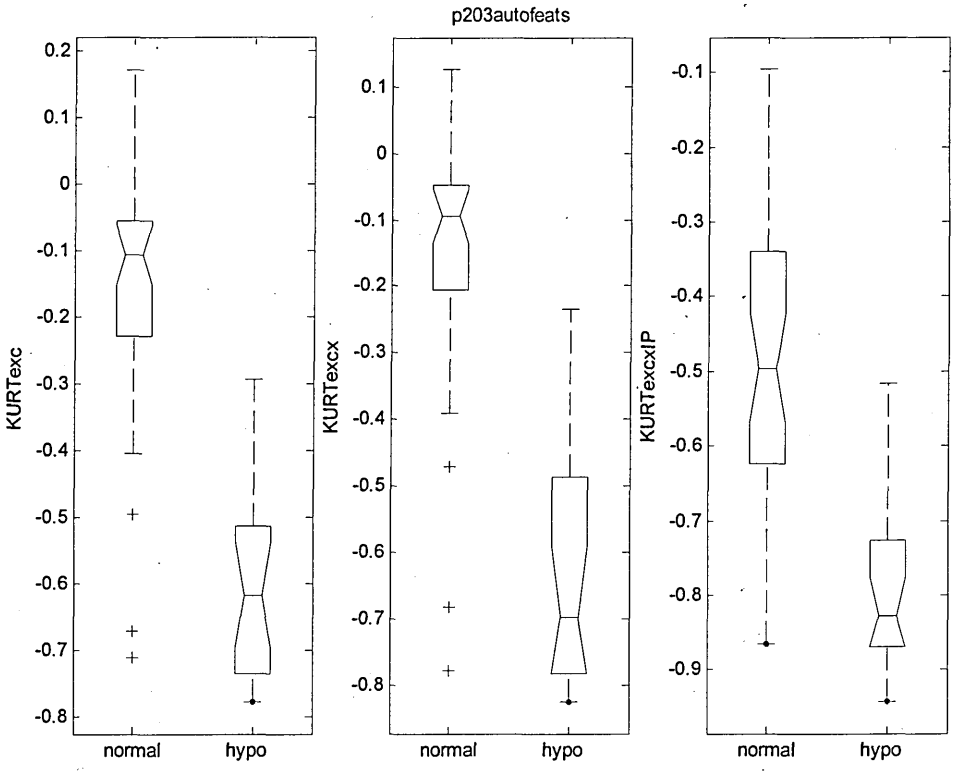


Figure 4.24: Boxplots for the 3 versions of the kurtosis excess feature for patient 203 (both nights)

4.5 ECG features

Time-domain ECG features were extracted and used in order to represent ECG traces for classification to be carried out. From a theoretical point of view, such ECG features were discussed in Section 2.4. Some of the features listed below have been presented already in this chapter in the sections where the relevant algorithms for extracting them were presented. The ECG features that we considered, assessed and used were:

1. RR: this is an instantaneous measure of Heart Rate. Heart Rate variations are possible under hypoglycaemia and hence the RR was considered as a feature.
2. RT: time interval from R peak to T wave end. The corrected version of this interval (RTc) was mainly used.
3. RTc: Heart-Rate-corrected RT interval. This was chosen instead of QTc in order to avoid possible variation due to the Q point detection process. The R peak is easier to detect than the Q point and hence the annotation of the former is expected to be more robust.
4. RTapex : time interval from R peak to T wave peak. The corrected version of this interval (RTapexc) was mainly used.

5. RTapexc: Heart-Rate-corrected RT interval. In certain studies, the RTapexc was used instead of the RTc in order to avoid possible variation due to the T end detection process. Detection of the T wave end can suffer from noise and artefacts and becomes difficult under T wave flattening and presence of U waves. In order to avoid such problems, the RTapexc was used to investigate whether it can still highlight effectively the delayed VR process.
6. T-duration (Tdur): time duration from T onset to T offset.
7. T-durationc (Tdurc): Tduration corrected for heart rate.
8. T amplitude (Tampl): amplitude of the T wave from the isoelectric (0 Volts) level. This is a significant feature since changes in T amplitude are related to changes in plasma potassium that occur under hypoglycaemia. The T wave amplitude is expected to reduce under hypoglycaemia. Alternative ways of calculating the T amplitude would be to use the ST segment as a reference instead of the isoelectric level. Other approaches would be to use the ratio of Tampl/Rampl in order to achieve some form of normalisation of the Tampl values.
9. T-area: Area under the T wave. The area is enclosed by the ECG trace, the isoelectric line and the vertical lines defined by the T onset and offset markers.
10. T wave Symmetry (Half-Area Ratio (HAR)). This is a measure of symmetry of the T wave. It was used because T wave symmetry changes are observed under hypoglycaemia. This feature was discussed in detail in Section 4.4.3.
11. T wave skewness. This is an alternative measure of T wave symmetry, as mentioned earlier. It was inspired by the 3^{rd} central moment from statistical theory used to calculate the skewness of distributions. In this study it is not used as a statistical measure but using an analogy it is used to assess the symmetry of the T wave.
12. T wave kurtosis. This feature was used to assess the peakedness of the T wave and was also inspired from statistical theory and more specifically the 4^{th} central moment. Again it is not used as a statistical measure but, using the analogy, it is used to assess the flatness of the T wave peak.

When extracting the above features, the R and T peaks were detected using the algorithms presented in this chapter. The T onset and offset were detected using the tangent method (msi). An example of the set of features extracted in most studies is given in Table 4.5.

Table 4.5: Extracted ECG features for patient 204

Record	gl	HR	RTmsi	RTcmsi	Tdur	Tdurc	Tampl	Tarea	HAR	SKEW	KURTexc	RTapex	RTapexc
5	8.48	84.55	291.38	345.89	100.88	119.75	206.62	1536.83	1.49	1.03	-0.33	232.00	275.41
6	8.48	83.30	294.13	346.57	107.02	126.10	248.57	1937.30	1.48	1.16	-0.15	232.00	273.36
7	7.09	86.97	297.79	358.54	105.79	127.37	226.08	1737.40	1.18	0.77	-0.44	240.00	288.96
8	5.08	85.81	297.92	356.28	122.20	146.13	368.19	3304.47	0.86	-0.27	-0.60	240.00	287.01
9	5.41	86.70	291.59	350.52	122.20	146.89	404.66	3636.03	0.98	-0.11	-0.60	232.00	278.88
10	5.66	95.56	283.54	357.83	118.78	149.90	362.94	3133.20	0.69	-0.74	-0.55	232.00	292.79
11	6.14	86.46	292.87	351.57	123.78	148.59	409.12	3766.22	0.72	-0.60	-0.51	240.00	288.10
12	5.23	86.77	293.65	353.14	106.16	127.66	201.02	1606.02	0.95	0.43	-0.53	240.00	288.62
13	4.91	87.77	296.20	358.25	121.81	147.33	391.82	3514.68	1.26	0.51	-0.53	232.00	280.60
14	4.69	84.48	297.87	353.44	125.63	149.07	435.32	3962.74	0.83	-0.32	-0.60	240.00	284.78
15	4.06	89.98	285.42	349.52	119.86	146.78	391.05	3453.42	0.76	-0.47	-0.56	232.00	284.11
16	4.09	92.30	277.29	343.92	104.57	129.70	271.56	2161.09	0.78	-0.40	-0.61	232.00	287.74
17	2.82	88.12	291.84	353.69	92.68	112.32	132.84	912.46	1.17	0.74	-0.45	240.00	290.86
18	2.20	90.40	295.01	362.11	98.98	121.49	154.70	1153.99	1.26	0.87	-0.33	240.00	294.59
19	2.20	86.86	290.00	348.94	100.46	120.88	168.75	1275.51	1.40	1.02	-0.32	232.00	279.15
20	2.20	89.10	282.57	344.34	97.13	118.37	171.53	1229.85	1.07	0.49	-0.61	232.00	282.72
21	2.20	92.09	290.14	359.45	98.57	122.11	155.88	1159.65	1.58	1.16	-0.25	232.00	287.42
22	2.20	88.53	291.86	354.53	104.14	126.50	157.15	1185.12	1.36	1.09	-0.36	232.00	281.82
23	2.20	92.31	294.11	364.81	97.73	121.22	127.94	962.25	1.22	0.88	-0.37	240.00	297.69
24	2.20	84.16	300.01	355.33	121.76	144.21	432.17	3982.23	1.05	0.17	-0.53	240.00	284.25
25	2.20	86.82	297.47	357.82	123.14	148.12	446.38	4002.90	0.92	-0.19	-0.60	240.00	288.69
26	2.20	84.54	294.02	349.00	98.69	117.15	162.38	1193.52	1.26	1.03	-0.30	240.00	284.88
27	2.20	91.03	294.74	363.06	100.72	124.06	179.40	1262.76	1.16	0.82	-0.32	240.00	295.62
28	2.20	86.74	288.66	347.07	96.63	116.19	162.56	1184.09	0.94	0.22	-0.62	240.00	288.57
29	2.20	88.92	292.33	355.87	104.35	127.04	215.67	1766.22	0.87	0.25	-0.57	240.00	292.17
30	2.76	87.17	290.94	350.69	105.43	127.08	208.13	1608.22	1.34	0.96	-0.40	232.00	279.64
31	2.20	86.15	302.17	362.07	123.68	148.20	404.52	3757.61	0.72	-0.56	-0.55	248.00	297.16
32	2.20	90.59	289.88	356.19	123.52	151.77	390.61	3531.06	0.91	-0.19	-0.60	232.00	285.07
33	2.20	88.88	294.15	358.01	124.84	151.94	413.89	3782.15	1.04	0.14	-0.56	232.00	282.37

Boxplots presenting the ECG features for patient 227 (both nights) are presented in Figure 4.25. The versions of the features where no heart-rate-correction was applied were excluded from the graph. In each subplot the LHS box corresponds to normality while the RHS corresponds to hypoglycaemia. The y-axis label informs about the ECG feature plotted. This subject shows moderate feature changes under hypoglycaemia and it can be seen that the biggest changes occur for the morphological features (HAR, SKEW, KURTexc).

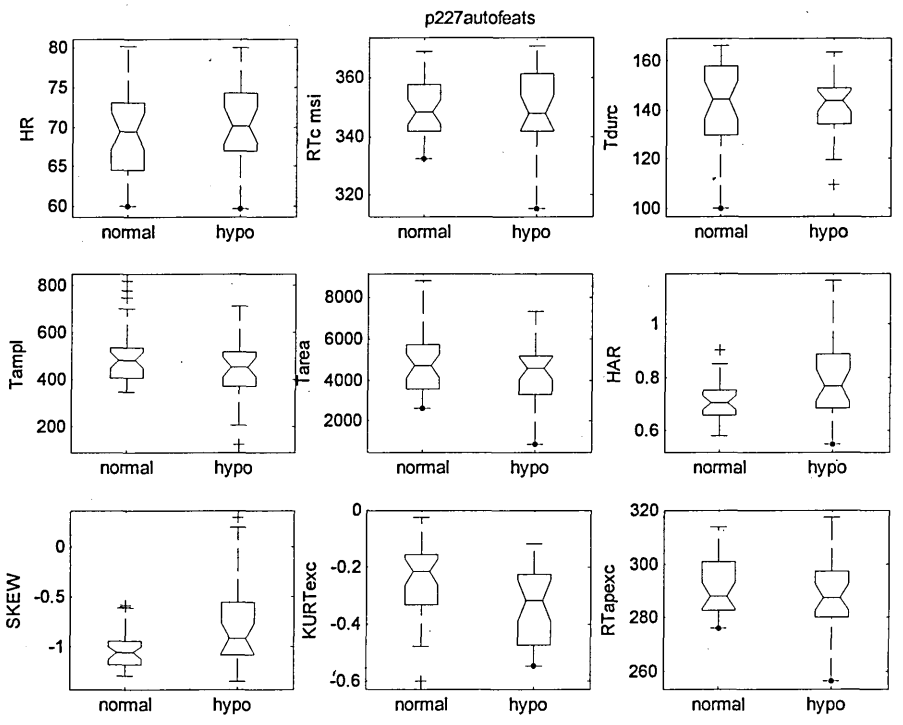


Figure 4.25: Boxplots for 9 ECG features, for patient 227 (both nights)

Box plots of the RTc, T amplitude and HAR features for all patients are presented in Figures 4.26, 4.27 and 4.28. These plots allow for inspection of the variation in feature ranges across patients. Inter-patient and intra-patient variability is apparent in the figures.

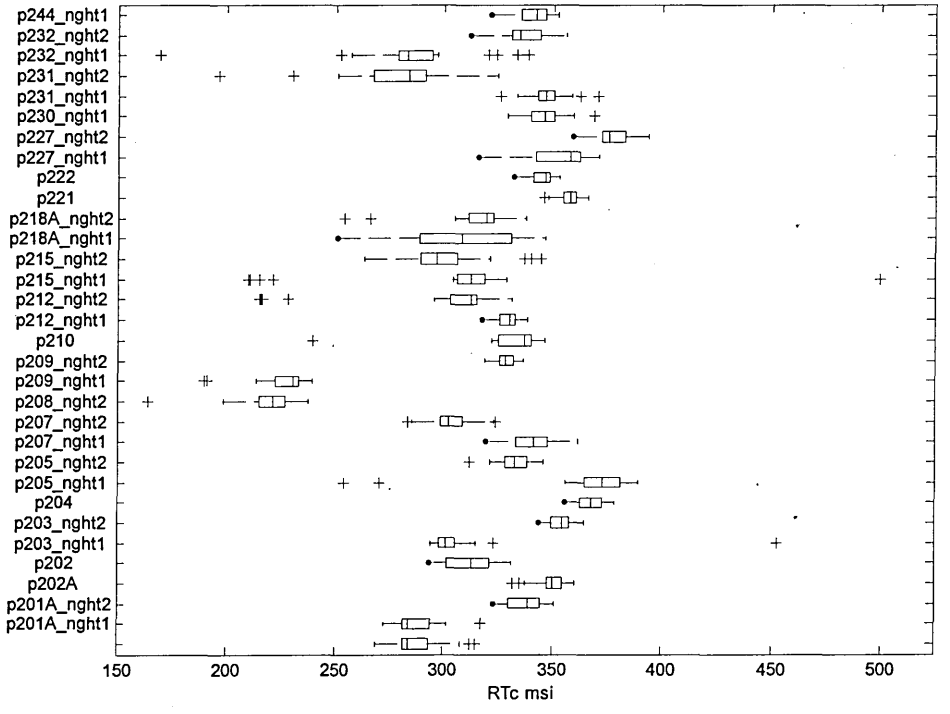


Figure 4.26: Boxplots for the RTc feature across all patients

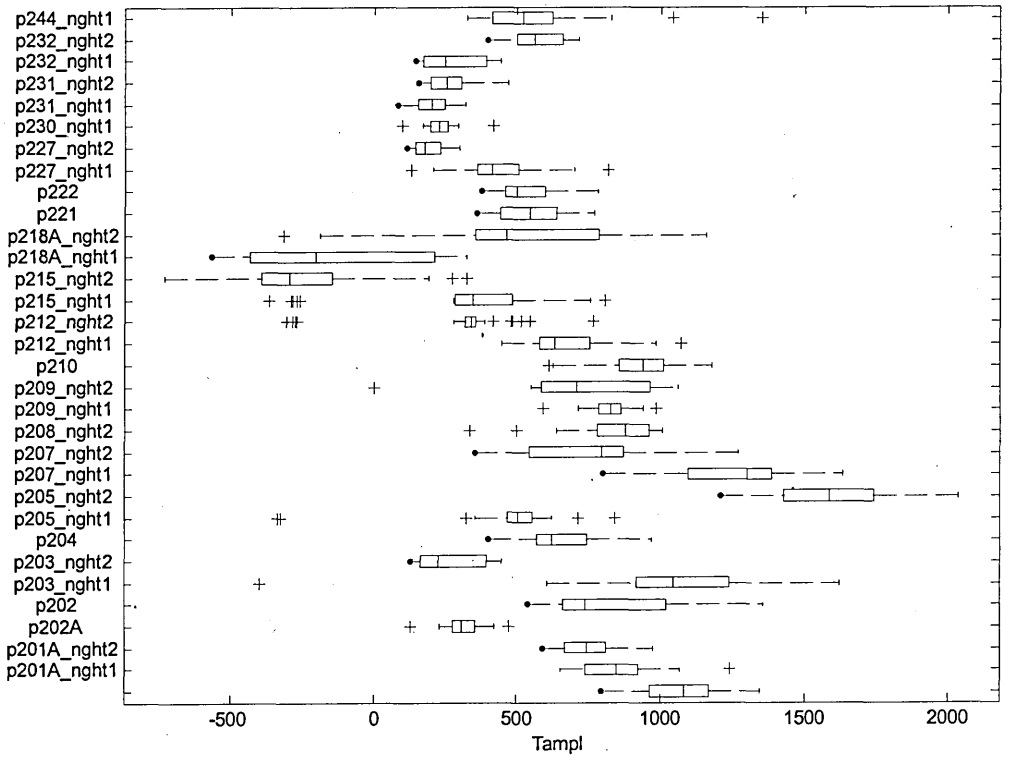


Figure 4.27: Boxplots for the Tampl feature across all patients

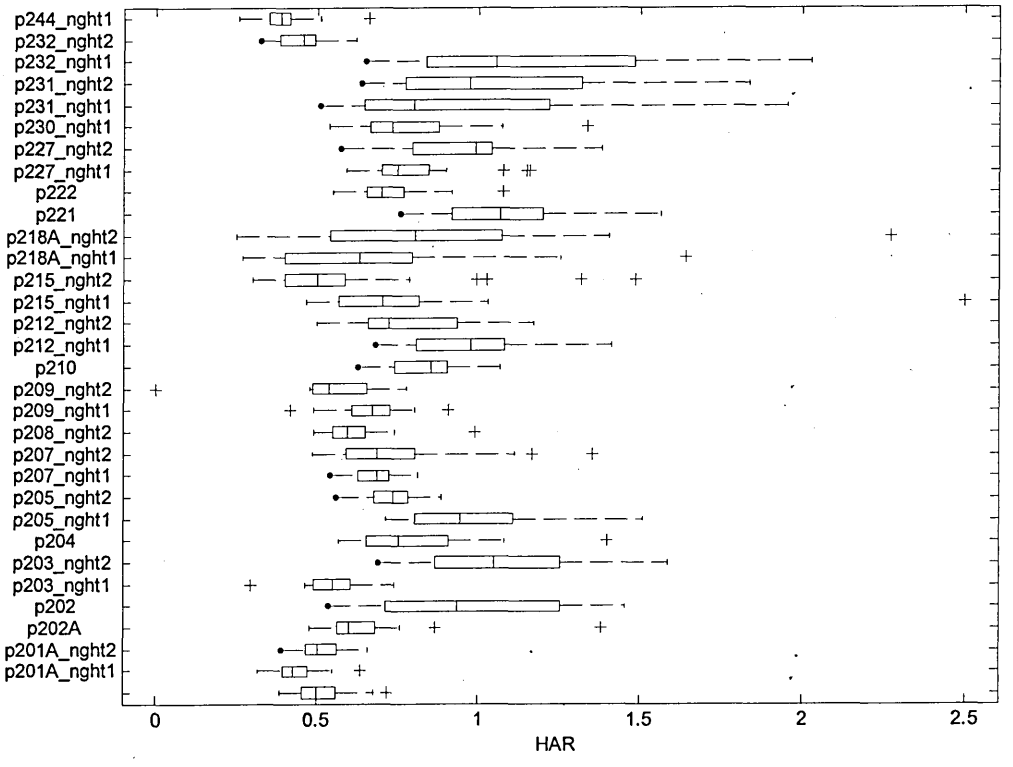


Figure 4.28: Boxplots for the HAR feature across all patients

4.5.1 Tests of significance on the ECG features change between the conditions of euglycaemia and hypoglycaemia

In order to check for statistical significance in the feature changes due to hypoglycaemia, hypothesis tests were carried out. Two-sample t-tests at the 5% level of significance were employed for each ECG feature considered. For each feature, the null hypothesis H_0 stated that the changes in the mean value between hypoglycaemia and normality were not statistically significant. The alternative hypothesis H_1 stated that the changes in the mean value were significant i.e. the mean was different between the conditions of normality and hypoglycaemia. Two-tailed tests were carried out, which means that the alternative hypothesis was describing changes in mean in both directions (greater or smaller).

The tests were carried out separately for each patient. Only patients that contributed data representative of both clinical conditions (euglycaemia and hypoglycaemia) were included in the tests. Outliers were removed using the 3 SD criterion prior to the tests. A hypoglycaemic threshold of 3 mmol/lt was used to distinguish between hypoglycaemia and euglycaemia. Before proceeding any further in the discussion it is important to bear in mind that the t-test assumes a Normal distribution in the data. The t-test results can be doubted if the features do not follow the Normal distribution.

Table 4.6 contains the t-test results for all ECG features considered. A value of 1 denotes that the outcome of the t-test indicated rejection of the null hypothesis at the 5% level of significance i.e. the mean values corresponding to normality and hypoglycaemia were different. A value of 0 denotes acceptance of the null hypothesis. The column labelled "sum" in the table contains the sum of each row. For each ECG feature included, the t-tests gave different results across patients. Calculating the sum of each row gives an indication of the number of patients for whom the changes in a given feature were statistically significant. The best features according to this metric were, RTc and SKEWx. For these features, the t-test indicated that statistically significant changes occurred in 4 out of 7 patients.

Table 4.6: t-test results at the 5% level of significance for 7 patients

patient	p202	p203	p204	p209	p212	p227	p244	sum
HR	1	0	0	0	0	0	1	2
RTmsi	1	1	0	0	0	0	1	3
RTcmsi	1	1	0	0	1	0	1	4
Tdur	1	0	1	0	0	0	1	3
Tdurc	0	0	0	0	1	0	1	2
Tampl	1	1	1	0	0	0	0	3
Tarea	1	1	0	0	0	0	1	3
HAR	0	1	0	0	0	0	1	2
SKEW	0	1	0	0	0	0	1	2
KURTexc	0	1	0	0	0	1	0	2
RTapex	1	1	0	0	0	0	1	3
RTapexc	1	1	0	0	1	0	0	3
HARx	0	1	1	0	0	0	1	3
HARxIP	0	1	0	0	0	1	1	3
SKEWx	1	1	1	0	0	0	1	4
SKEWxIP	0	1	0	0	0	1	0	2
KURTexcx	1	1	0	0	0	0	0	2
KURTexcxIP	0	1	0	0	0	1	1	3

The t-test results give indications that a single feature combination cannot be used to detect hypoglycaemia in all patients. Different features may have to be used to detect the onset of hypoglycaemia in different patients. Different components of the ECG will be affected by hypoglycaemia in different patients. A typical example of the above is the presence of U waves under hypo. Presence of a U wave could be an indication of hypoglycaemia but does not happen in all patients. It can be seen on the table that for patient 203, fifteen features exhibited statistically significant changes in response to hypoglycaemia. On the other hand, for patient 209 the changes in all features studied were statistically insignificant. Such a result is possible since some patients can be asymptomatic during hypoglycaemia. Specifically for the case of patient 209, the result has to be treated with caution since the sample size for the hypo group was extremely small (n=4 for hypo and n=50 for euglycaemia) and the t-test is probably not accurate.

4.6 AutoRegressive Modelling (AR) of post-R ECG traces

This section presents the approach of representing the ECG traces using Autoregressive (AR) models. This type of ECG trace representation was used as an alternative to the approaches that used ECG features. The autoregressive coefficients computed were used to describe the modelled ECG traces. This approach has been used in other studies for detection of certain cardiac arrhythmias [Srinivasan 2002]. In this study it was used for the detection of the delayed ventricular repolarisation often exhibited during hypoglycaemia. Under hypoglycaemia, and due to the T wave flattening, the AR model

parameters are expected to be different because of the smaller undulations existing in the data (post-R segment) to be modelled.

An n[th] order AutoRegressive (AR) model is a linear recursive model described by the difference equation:

$$y(k+1) = -\sum_{i=1}^{i=n} a_i \cdot y(k-i+1) + e(k+1) + \beta, \quad k,n \in \aleph \qquad \text{eq}^{n}\ (4.9)$$

where *e(k)* denotes a noise parameter and β denotes an offset parameter.
For an n^{th} order model, the output y at sample number $k+1$ is modelled as the sum of n scaled versions of previous values of the output plus a noise component and an offset component. The estimates of the optimal model parameters (a_i and β) that achieve minimal error need to be calculated. The Least Squares (LS) algorithm can be used to find the estimates of the optimal model parameters. It works by minimising the sum of the squares of the model errors.

In matrix form we can write: $\mathbf{Y} = \Phi \times \hat{\mathbf{B}}$ eq[n] (4.10)

Each row of vector **Y** contains the LHS of eq[n] (4.9) for a different value of *k*:

$$\begin{bmatrix} y(2) \\ y(3) \\ \ldots\ldots \\ \ldots\ldots \\ y(k+1) \end{bmatrix}$$ The vector $\hat{\mathbf{B}}$ contains the model parameters: $$\begin{bmatrix} a_1 \\ a_2 \\ \ldots\ldots \\ a_n \\ \beta \end{bmatrix}$$

If the number of elements in **Y** is denoted by N then, the N by $(n+1)$ matrix Φ has the form:

$$\begin{bmatrix} -y(1) & 0 & \ldots\ldots & 0 & 1 \\ -y(2) & -y(1) & \ldots\ldots & 0 & 1 \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ -y(k) & -y(k-1) & \ldots\ldots & -y(k-n) & 1 \end{bmatrix}$$

By multiplying the N by $(n+1)$ matrix Φ with the $(n+1)$ by 1 vector $\hat{\mathbf{B}}$, we get an N by 1 column vector where each row is the RHS of the AR model presented in eq[n] (4.9).

The least squares estimate is obtained by minimising the sum of the squares of the model errors. The mean squared error (MSE) is calculated as follows:

$$MSE = \frac{1}{N}\sum_{k=1}^{N}(y_k - y_k^m)^2 \qquad \text{eq}^n\ (4.11)$$

where y_k is the actual output at sample instant k and y_k^m is the modelled version of the output at the same sample instant, calculated by multiplying the row of matrix Φ corresponding to k with the column vector $\hat{\mathbf{B}}$ containing the model parameters.
In matrix form, eq^n (4.11) is written as:

$$\mathbf{MSE} = \frac{1}{N}(Y - \Phi\hat{\mathrm{B}})^T(Y - \Phi\hat{\mathrm{B}}) \qquad \text{eq}^n\ (4.12)$$

Solving with respect to $\hat{\mathrm{B}}$ we get:

$$\hat{\mathrm{B}} = (\Phi^T\Phi)^{-1}\Phi^T Y \qquad \text{eq}^n\ (4.13)$$

The vector $\hat{\mathrm{B}}$ contains the estimates of the optimal model parameters. AR modelling was used in this thesis as an alternative to the feature extraction process presented earlier, for the representation of ECG traces.

When modelling the ECG signal, the general form of an nth order AR model takes the form:

$$ECG(k+1) = -\sum_{i=1}^{i=n} a_i \cdot ECG(k-i+1) + e(k+1) + \beta, \quad k,n \in \aleph \qquad \text{eq}^n\ (4.14)$$

ECG(k+1) is the ECG amplitude, in mV, at sample *k+1*. The ECG section to the right of the *R* peak and until the end of the trace was used because this is the section affected by hypoglycaemia. The Least Squares (LS) algorithm was used to find the estimates of the optimal model parameters. A 3rd order AR model was employed which yields four model parameters (a_1, a_2, a_3 and β). Figure 4.29 illustrates a SAECG trace together with its modelled version using a 2nd order AR model.

The whole ECG cycle is plotted (solid blue) but only the post-R peak section is modelled (dotted black). It must be noted that the model order was set to 2 only for the purposes of generating the Figure (4.29) illustrating the modelling process. Using a 3rd order model or higher would cause the solid and dotted lines to almost overlap so the differences would not be visible.

Apart from the illustration in Figure 4.29, all the ECG modelling was carried out using a 3rd order AR model. The correlation coefficient between actual and modelled ECG trace was for all but one patient greater than 91% with an average of 95% across patients. The model order can be increased so that each ECG trace is more closely modelled but this will produce extra model parameters that the classifiers have to handle and classify. Emphasis was placed on making the model simple and hence keeping the classification task simple.

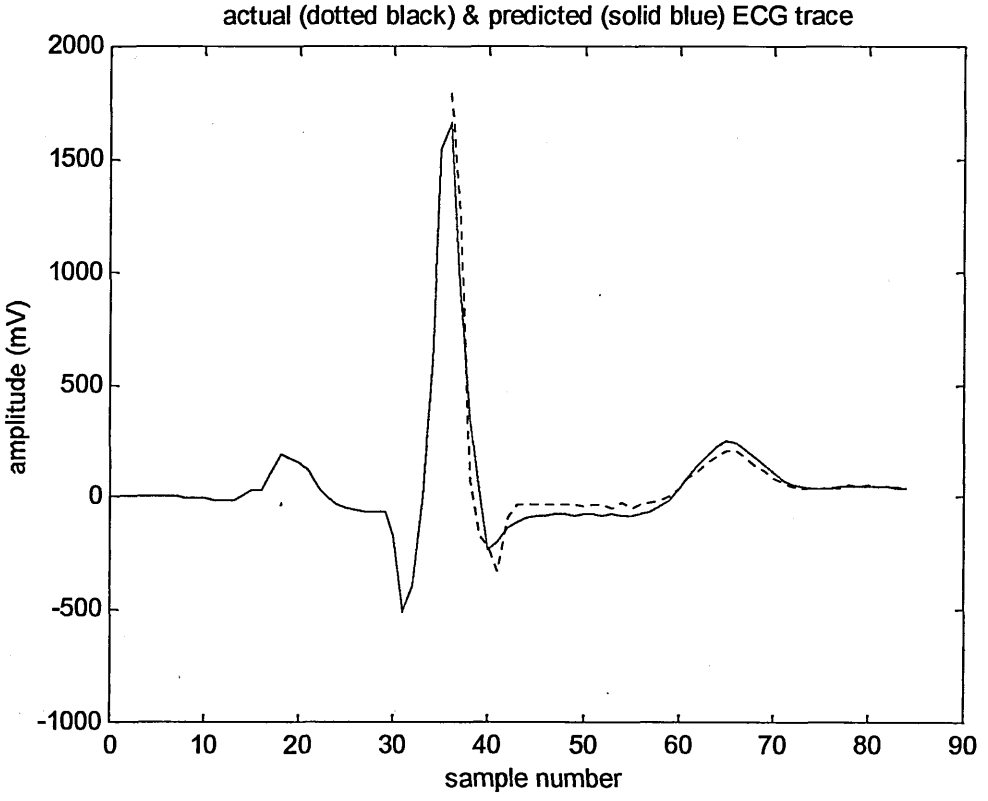


Figure 4.29: Actual (solid) vs modelled

The classification results in the approach of modelling the ECG by means of AR coefficients are presented in Section 5.2.5. Linear Discriminant Analysis was used for classifying the AR coefficients into two groups corresponding to euglycaemia and hypoglycaemia.

4.7 Conclusion

This chapter focused on the representation of ECG traces by means of ECG features. The feature extraction algorithms and the features introduced and used were presented. The feature extraction process was put in the context of the overall methodology proposed. This process precedes the classification process in the hypoglycaemia detection methodology proposed. The ECG features extracted are fed to a classifier for detection of the abnormal clinical condition. Annotation algorithms for detecting the R

and T peaks and also the T onset and offset were implemented in MATLAB. A comparison of T end annotations was carried out. Three novel ECG features were introduced for assessment of the T wave morphology. The visual inspection and statistical analysis carried out on all the features considered, indicated the existence of inter-patient and intra-patient variability. Moreover there were indications that a fixed set of features may not be able to effectively represent the ECG cycles of all patients. Certain features are useful for certain patients depending on the dynamics of each patient's ECG signature. The next chapter focuses on the classification of ECG traces using neural and statistical classifiers.

Classification of Signal Averaged Electrocardiogram Signals using Artificial Neural Networks and Statistical Classifiers

5.0 Introduction

This chapter contains Methods and Results for the classification of SAECG signals using artificial neural networks (ANN) and statistical classifiers. Feed-forward neural networks (multi-layer perceptrons (MLP)), Linear Discriminant Analysis (LDA) and k–Nearest Neighbour (kNN) classifiers were used for classification. A number of different approaches to classification are presented and discussed. The chapter is structured as follows: initially the data-set used is addressed followed by the preprocessing methodology for the data used. Following that, the neural network architecture is discussed detailing all the necessary aspects of using MLPs for ECG trace classification. After that, the statistical classifiers (LDA and kNN) are presented and the approach used for evaluation of performance of all classifiers is given.

Moving into the Results section, the classification studies carried out during the initial phase of the research are first presented together with an assessment of the ECG features inspired by Benhorin's work [Benhorin 1990] and extracted semi-automatically in ECGLAB. The use of various feature combinations is presented. Then the focus moves to the individual-patient-oriented classification studies where emphasis is given on customising the system for individual patients in order to tackle inter-patient variability problems. Statistical classifiers are also included in these studies. Next the use of ECG representation by Autoregressive Modelling coefficients is presented and LDA classification results are given. Final results in the chapter involve the use of improved preprocessing combined with utilisation of a reduced set of ECG features.

5.1 Methodology

5.1.1 Data

Data from the dataset presented in Section 3.3.1, were used in the classification studies. The approach according to which the data were partitioned into training and test sets is presented below. Issues related to inter-patient variability are also discussed.

5.1.2 Preprocessing

A number of preprocessing steps were followed to aid the classification process. Firstly, data points of each ECG feature were removed as outliers if they did not lie within 3 standard deviations from the mean. This was carried out separately for each ECG feature used[19] since the various ECG features had different ranges of values (Section 4.5). Outliers can occur because of noise or other artefacts, or due to bad performance of the feature extraction algorithms upon an ambiguous ECG trace. It is essential to remove outliers, since they can degrade the performance of the ANNs. If outliers are present they can contaminate the range of the feature values and once normalisation is applied, the useful information will be squashed to a significantly smaller range, leading to loss of useful information.

A second pre-processing step was to normalise the data in the interval [-1 1]. This was also done separately for each feature, because of the differences in range among features. The various features were describing different quantities and hence their range of values was different. For instance, some features were measuring amplitudes in mV and others were describing time intervals in msec. For each feature, the samples available were mapped linearly from the original space to [-1 1].

The next pre-processing step was to remove the baseline value (i.e. the dc component) from each ECG feature. This was not applied in all studies carried out. For the studies aiming to produce a global classifier by forming datasets that included data from all patients, baseline removal was essential in reducing inter-patient variability. The first recorded sample of every night was considered the baseline and it was subtracted from

[19] i.e. the mean and standard deviation used in the outlier removal criterion were calculated separately for each feature.

all the samples. Alternatively the mean value of all samples can be used as the baseline. Using the mean has some disadvantages and was not preferred in our case. First of all using the mean is not convenient for online classification. Using a mean value as the baseline, requires some portion of the data fed to the system to be used for calculation of the mean before classification can start. For the offline classification studies included in this thesis, the first value baseline is normally a good sample and is always euglycaemic, since all accepted nights were starting with the patient having normal glucose levels. On the other hand, the mean value can be contaminated by artefacts and will vary depending on whether hypoglycaemic samples are contained in the night.

Figure 5.1 illustrates the preprocessing stages that were included in the ECG-trace-classification process. The baseline-removal stage succeeds the outlier-removal stage but is not depicted since it was not included in all studies.

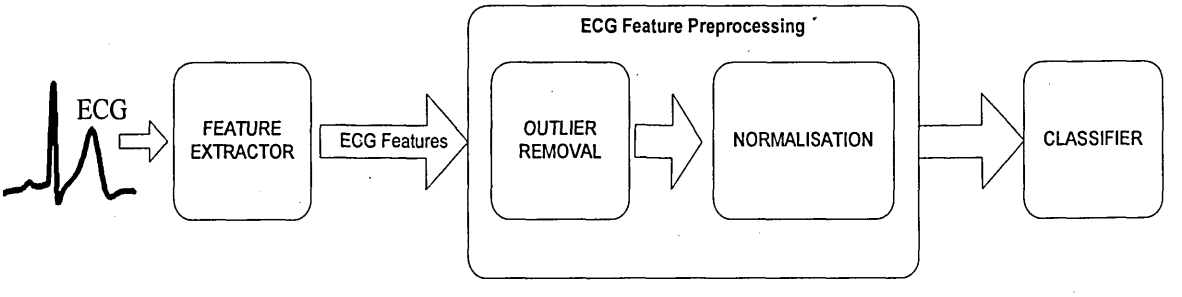


Figure 5.1: Preprocessing of ECG features before they are fed to the classifier

5.1.3 Formation of data-sets (training and test files)

Two approaches were followed in the formation of data-sets. According to the first one, a number of patients were mixed together to form the datasets. This approach suffers from problems related to inter-patient variability. The behaviour of the ECG signal and hence the baseline, mean, standard deviation, range etc, of ECG features varies from patient to patient. This was highlighted in Figures 4.28 and 4.29 for two of the ECG features considered. This variation undermines the generalisation ability of the classifiers. Variability among patients may be due to variation in many parameters such as: age, gender, duration of diabetes, level of glycaemic control, fitness level and so on.

Among the above parameters, the only one that was taken in account when forming datasets was the gender of the patient. The gender is a binary parameter (either male or female) so it was a lot easier to take it in account when forming datasets. And more importantly, there are indications that the gender is a factor that may affect the baseline value of some ECG features. Healthy females are found to have longer QTc intervals

compared to healthy males [Molnar 1996]. Datasets were formed by using data from all patients and alternatively, using only male and only female patients in order to avoid variations due to gender and also attempt to investigate the effect of the gender. The variation in the remaining parameters (e.g. age, duration of diabetes etc) were not taken into account since this was beyond the scope of this research.

To overcome inter-patient variability problems a second approach was also followed according to which a different neural-network/statistical-classifier was used for each patient or in some cases for each night. This overcomes the problem of inter-patient variability by allowing a classifier to be customised to a specific patient. The problem with the second approach is that the data available are significantly reduced. As mentioned in Section 3.3.1 presenting the dataset, a maximum of two nights is available per patient which gave a maximum of 66 SAECG records. 66 patterns were not the ideal amount for training and querying an ANN but the advantage was that the only problem to be overcome was intra-patient variability since inter-patient variability was eliminated. Considering the trade-off between having longer datasets and eliminating inter-patient variability, it was clear that the latter was more important. Such a choice was a significant step towards improving classification performance. In order to make maximum use of the data available, 5-fold cross-validation was applied. The partitioning of data in training and test sets, under cross-validation was repeated each time a neural network was re-trained from different random initial conditions. This is advantageous over the approach of generating the cross-validation groups only once and then training the networks from different initial conditions. Since the cross-validation groups were formed many times after shuffling of the feature vectors, the formation of training and test files was unbiased.

According to cross-validation the data was partitioned in groups. 5-fold means that 5 groups of approximately equal size were used i.e. each cross-validation group consisted of 20% of the total data available. Care was taken so that each group contained equal number of euglycaemic and hypoglycaemic feature vectors. Four groups were merged together (i.e. 80% of the data) to form the training file for the classifier and the remaining 20% was reserved for testing. The process was repeated five times so each group would be left out once, and the classification results were averaged over 5. 10-fold cross-validation is also a very common alternative but only leaves 10% of the data out for testing each time. In a datafile of 66 pattern vectors this means approximately 6-

7 vectors. In our case and in order to maximise the size of the test data, 5-fold cross-validation was chosen. This process will be further discussed in the Results section (5.2.2).

5.1.4 Neural Network Classifiers

The Neural Network Architecture chosen for this research was the Multi-Layer Perceptron (MLP) presented in Section 2.7.1. We needed a supervised neural classifier and the MLP was an obvious choice since it is an established and very widely used architecture. Moreover the MATLAB neural network toolbox was available which contains extensive tools (mainly in the form of various pre-processing and training algorithms) for the MLP. Networks were trained by the batch gradient-descent back-propagation algorithm. The neural network toolbox (versions 3.0.1 (R11) and 4.0.2 (R13)) of MATLAB (MathWorks Inc) was used in all the neural networks that were set up.

Neural network size

Neural networks consisting of one hidden layer were trained. The number of neurons in the hidden layer were variable in the interval [2 5] or in some cases in [2 10][20] depending on the study carried out. The network size was fixed during training. Each neural network was trained 4 or 9 times (depending on the range of allowed hidden neurons in each study) with a different number of neurons in the hidden layer each time.

Neural network initial conditions

The initial weights and biases of ANNs were initialised to random numbers. Training a network more than once from different random initial conditions can yield different results as ANNs are very sensitive to initial conditions. In most cases, a hundred networks or more were trained for the same configuration and the best ones chosen. This was done to overcome the sensitivity of the network performance to random initial conditions.

Output of the neural networks

Two approaches were followed regarding the type of the neural network output. In the first case the neural network was presented with the continuous glucose data as a target

[20] hidden layer sizes in [2 10] were only used in the studies producing global classifiers, where the datasets were longer.

for training (Section 5.2.1) while in the second one the glucose was coded into two classes, a normal and a hypoglycaemic class[21]. The output of the ANNs was not simply interpreted as normal or hypoglycaemic. A third class was considered. This is the class of undetermined classifications. The output of the neural network is a real number which varies from –1 to 1 (since the activation function used was a hyperbolic tangent, as discussed later). In order to map the output of the neural network to the "normal" class or to the "hypoglycaemic" class, a threshold had to be considered. The most straightforward value to choose for this threshold was zero. The simplest approach of mapping would be to assign anything greater than zero to the "hypoglycaemic" class and anything below to the "normal" class i.e. the equivalent of using a hard-limiter activation function. This mapping though will not show us the cases where the network was unable to choose between the two classes (i.e. output stuck around zero). A better approach is to use the dual-threshold method and map output values greater than a chosen number (0.5 or 0.8 for instance) to "hypoglycaemic" and those smaller than the chosen number (e.g. – 0.5 or – 0.8) to "normal". This way a third group, that of undetermined classifications ("don't-know" outcomes), is created. The number of times that the ANN output lies in the region of undetermined classifications will give us an idea of how clearly it can distinguish the two classes. If the network maps many input patterns to the undetermined region, then there is an indication that it cannot separate the two classes easily. The partitioning of the above three classes is illustrated in Figure 5.2.

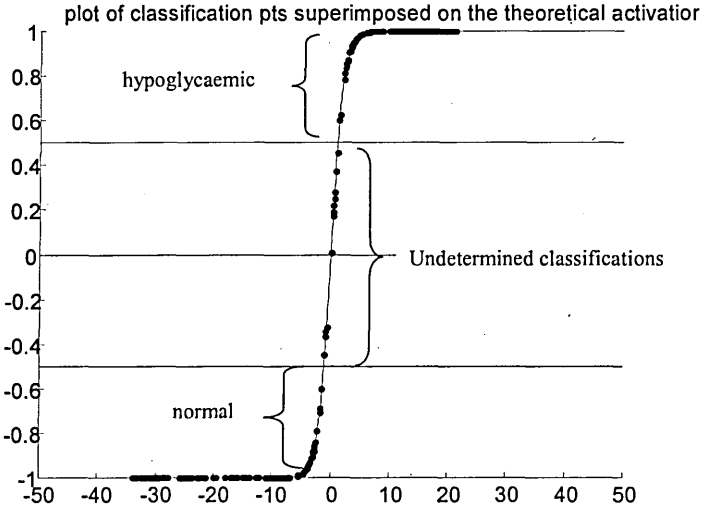


Figure 5.2: Output activation function with mapping to the 3 classes (normal, hypoglycaemic, undetermined) shown. Dots denote data points lying on each class

[21] in some studies the normal (euglycaemic) class also contained hyperglycaemic data while in other studies the inclusion of strictly euglycaemic data was tried.

Activation functions

Hyperbolic tangent activation functions were used both for the hidden and the output layers. The hyperbolic tangent (tansig) we used is smoother than the built-in one contained in the MATLAB neural network toolbox. It has the same shape as the built-in logistic sigmoid (logsig) function contained in the toolbox and is given by the formula: $y = 2 / (1 + exp(-n)) - 1$ instead of $y = 2/(1+exp(-2*n))-1$ which is the default function. The largest the value of the gain inside the exponential term, the steeper the transition will be between the two extremes. The modified tansig, being smoother in slope, gives more candidates in the undetermined region. This means that the two classes must be more clearly separated by the network when using the modified tansig compared to the built-in one, so that they will not fall into the undetermined region.

Training parameters

The training algorithm used was batch gradient descent with or without momentum and with or without a variable learning rate (functions "traingd" and "traingdx" in the MATLAB neural network toolbox). When momentum was used it was chosen among values in the interval [0.1 0.7]. When a variable learning rate was used its initial value was set to 2, its increment was calculated by multiplying with 1:02 and its decrement by multiplying with 0.7.

5.1.5 Statistical classifiers

Statistical classifiers were also considered in order to allow comparisons with ANNs. The main statistical classifiers used were Linear Discriminant Analysis (LDA) and the k-Nearest Neighbour (kNN). LDA[22] works by minimising the Mahalanobis distance as discussed in Section 2.6. The k-Nearest Neighbour (kNN) was using a Euclidean distance metric (Section 2.6). The same ECG features that were fed into the ANN were used in LDA and kNN. Classification was binary into normal and arrhythmic records. Cross-validation was applied to achieve better use of the data available. Partitioning of the data into training and testing was exactly the same as for the ANN.

5.1.6 Measurement and Analysis of Performance

The accuracy of classification on its own is not sufficient to describe the performance of the classifiers. More aspects of performance are considered and the performance metrics

[22] *The LDA classifier was implemented using the "classify" command of the statistics toolbox in MATLAB*

used are: accuracy, hitrate (sensitivity), false-alarm-rate, specificity (true-negatives-ratio) and missed-hypos (false-negatives ratio).

They are defined as:

- Accuracy or concordance = tp + tn / (tp+tn+fp+fn) (eq^n 5.1)
- Sensitivity or Hitrate = tp / (tp +fn) (eq^n 5.2)
- False-alarm-rate = fp / (fp +tn) (eq^n 5.3)
- Specificity (TNratio) = tn / (tn + fp) (eq^n 5.4)
- Missed-hypos = fn / (fn + tp) (eq^n 5.5)

where *tp*, *tn*, *fp* and *fn* stand for: true positives, true negatives, false positives and false negatives respectively. Positive refers to hypoglycaemia while negative refers to euglycaemia. True positives are those classifications of the ANN where the real class is positive and the ANN also classified the ECG as positive. Similarly for the other three quantities:

True Negatives (TN): answer = NO, network said NO

False Positives (FP): answer = NO, network said YES

False Negatives (FN): answer = YES, network said NO

The above four quantities are summarised in Table 5.1 known as the confusion matrix:

Table 5.1: Confusion matrix

	Classifier YES	Classifier NO
Actual YES	TP	FN
Actual NO	FP	TN

Accurate classifiers produce as few FP and FN as possible. In other words a good classifier will give a confusion matrix with large numbers on the main diagonal and very small numbers on the second diagonal.

Sensitivity (hitrate) describes the number of arrhythmic traces classified correctly while false-alarm-rate describes the number of normal traces that were classified as arrhythmic (i.e. false alarms). Specificity (TNratio) describes the number of normal traces classified correctly while missed-hypos describes the number of arrhythmic traces classified as normal, i.e. the number of hypoglycaemic events that were missed[23]. By combining equations (5.1) and (5.5) the relationship between sensitivity and missed-

[23] The terms *sensitivity* and *hitrate* will be used interchangeably throughout the text. Similarly for the terms *specificity* and *TNratio*.

hypos is revealed: *sensitivity = 1 – missed-hypos*. Combining equations (5.3) and (5.4) it is shown that specificity and false-alarm-rate also sum to 1: *specificity = 1 – false-alarm-rate.*

5.2 Results Section

Classification results from a number of approaches and classification "recipes" are presented in this section. The various studies are presented in chronological sequence to show the development process of the ECG classification. Classifiers trained on global datasets are compared against classifiers tailored to the dynamics of specific patients. A number of different ECG feature combinations are compared while both neural and statistical classifiers are used.

5.2.1 Glucose level inference from ECG features

As has already been mentioned, most classification studies carried out in this research were binary, for identifying ECG traces into normal and arrhythmic. Besides these studies, the approach of inferring the approximate glucose level of the patient through the ECG was attempted. This was a very optimistic attempt and was carried out at the early stages of the research. The possibility of achieving a mapping between ECG and glucose was examined. The neural networks used were required to produce output values corresponding to the continuous level of glucose of the patient. By consulting the ECG, the MLP should infer the level of glucose of the patient at the time. Of course, this goes beyond our research hypothesis. The hypothesis assumes an existing relationship between the ECG and abnormally low glucose while this study examined whether the glucose could be inferred from the ECG regardless of whether it was abnormally low, normal or abnormally high.

In terms of the data needed, the advantage of such an approach where modelling of the ECG-glucose relationship is attempted, is that maximal use of the data available is possible. We do not necessarily need many hypoglycaemic nights. The study could be carried out on purely normal nights. This is because we are trying to infer the absolute glucose levels from the ECG. If this modelling was possible, having a good range of glucose levels, even if they did not lie in the hypoglycaemic band, could be enough to train a neural network.

This approach did not produce useful results as the objective set was extremely optimistic. Although there are indications towards a relationship between abnormally low glucose and hypoglycaemia, it is highly unlikely that the ECG strongly correlates with the glucose variable outside hypoglycaemia. Numeric results from this approach are not presented and this study is reported only in the form of a short discussion. The neural network simply failed to learn on the training data so numeric results cannot be presented.

5.2.2 Classification of ECG traces by MLPs trained on multiple patients, using ECG features extracted semi-automatically

In this study multi-layer perceptrons were used to classify ECG features extracted semi-automatically. The ECG features were extracted from ECG traces corresponding to euglycaemia and hypoglycaemia and hence distinguishing the two classes of feature vectors successfully would lead to detection of the two clinical conditions.

The feature extraction is classed as semi-automatic because the T end was marked manually but using a tangent line as a visual aid (i.e. use of the tangent method) and based on the above annotation of the T end, a number of features based on Benhorin's paper [Benhorin 1990] (presented in Section 2.4.2) were extracted. The Q and S markers were set manually (with no visual aids) by the human expert performing the annotations. The human expert was Cath Davies from the Royal Hallamshire Hospital (RHH) in Sheffield.

ECG Features used

Fifteen features were extracted from ECG traces using ECGLAB, based on the semi-automatic annotation by the human expert. The 15 features extracted along with their brief definition are given below:

1. RR – instantaneous Heart Rate
2. QT – time interval from Q point to T wave end
3. QTc – Heart-Rate-corrected QT
4. QRS – QRS duration, from Q point to S point
5. QRSc – QRS corrected for heart rate
6. SoTm – time interval from S point to T wave maximal amplitude (i.e. T peak)
7. SoTmc – SoTm corrected for heart rate
8. TmTo – time interval from T peak to T wave end

9. Time Area (25%) to Area (75%) - the time to accumulate the mid-50% of total absolute repolarisation area from its 25% to its 75% value
10. Total Area - Total absolute repolarisation area from S point to T end
11. Symmetry - T wave area symmetry ratio (SR); the ratio between the integrated area over SoTm and TmTo intervals. (SR= SoTm/TmTo)
12. % of Total Area to To – Total Area accumulated at To
13. (Area of ToUo)/Area of TmaxUo) – area under ECG from Tend until Uend upon area from T peak to Uend
14. %Tmax/Tbaselinemax – ratio of current T amplitude upon T amplitude at the start of the night
15. %Tmax/Rmax - ratio of T amplitude upon R amplitude

Out of the total number of 15 features produced, the following were used: RR, QT, QTc (corrected by Bazett's formula), Benhorin's T wave symmetry (SR) and %Tmax/Tbaselinemax. The features were selected because they were believed to be the most significant clinically. RR was used because variations in heart rate are expected in a hypoglycaemic event. QTc is a classical predictor of delayed VR among the clinical community [Harris 2000], [Ireland 1998, 2000]. The uncorrected version of QTc (QT) was also used to investigate its usefulness as a feature and also to provide clues towards the quality of heart-rate correction[24]. The symmetry of the T wave can be seen visually to be affected under hypoglycaemia hence a feature describing the symmetry was included. Finally the T amplitude (described by the %Tmax/Tbaselinemax feature) is expected to drop under hypoglycaemia due to the changes in potassium.

The remaining features extracted semi-automatically in ECGLAB were not used. The features excluded were: QRS, QRSc (corrected for heart rate), SoTm, SoTmc (corrected for heart rate), TmTo, Time Area (25%) to Area (75%), Total Area, % of Total Area to To, (Area of ToUo)/Area of TmaxUo), %Tmax/Rmax. After performing visual inspection, these features were identified as either not informative or not robust i.e. the extraction algorithm had poor performance and was not reliable. Some of the features were informative but were highly correlated with other features that were used. For instance the %Tmax/Rmax feature had almost identical behaviour to the

[24] The RR and QT were fed instead of the QTc in some cases, in order to test whether the neural network could detect QT changes that were uncorrelated to hear rate.

%Tmax/Tbaselinemax feature so only the latter was used in this study. An example of this can be seen in Figure 5.3 that depicts the two features (%Tmax/Tbaselinemax and %Tmax/Rmax) for patient 202A.

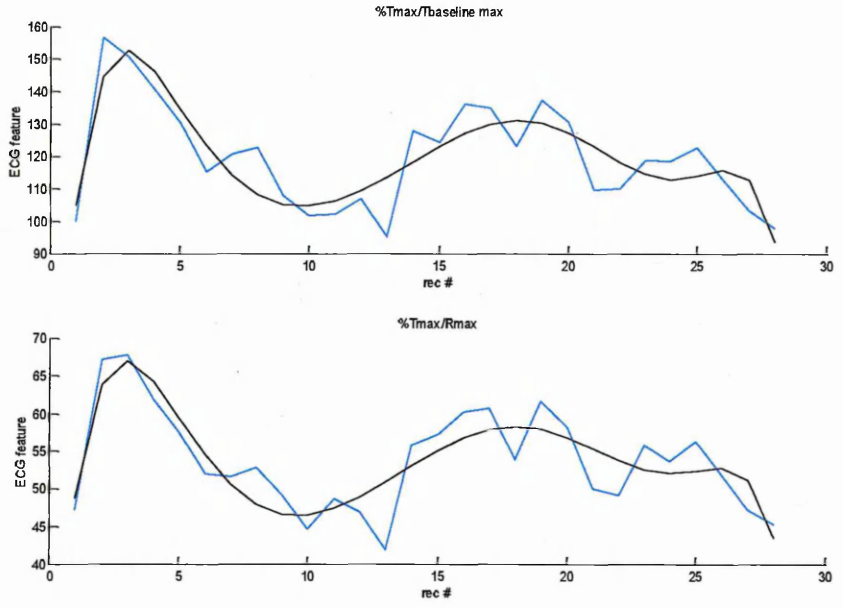


Figure 5.3: BLUE: Tmax/Tbaselinemax (top) and Tmax/Rmax (bottom)
BLACK: trend line (polynomial fitted)

The top graph presents the %Tmax/Tbaselinemax feature (in blue) during the night and the bottom graph the %Tmax/Rmax (in blue). Polynomials are fitted (black lines) in both cases in order to mark the trend of the feature. Trendlines, although not necessary in this graph since the features have very similar behaviour, were used during the process of visual inspection and this is the reason why they are included in this figure. It can be concluded from the figure that the two features have similar behaviour. The main difference is that the two features have a different baseline value and different range which can be seen by inspecting the marks of the two y-axes. Some subtle differences also exist. For instance the two features have slightly different behaviour around record 10. The top feature has a flat segment while there is a negative peak in the bottom one. A few more such subtle differences can be observed on the figure. Despite these differences, the two features have very similar behaviour.

Regarding the QRS (and QRSc) feature, it describes the ventricular depolarization of the heart and is not very useful for the detection of hypoglycaemia; hypoglycaemia affects the ventricular repolarisation process. Moreover, the QRS feature extracted by ECGLAB was not very robust. Figure 5.4 (top graph) depicts the QRS feature during

the night, for patient 201A-night1. The bottom graph presents the glucose profile for that night. It can be seen that the feature values oscillate between two levels which do not seem to be related to glucose.

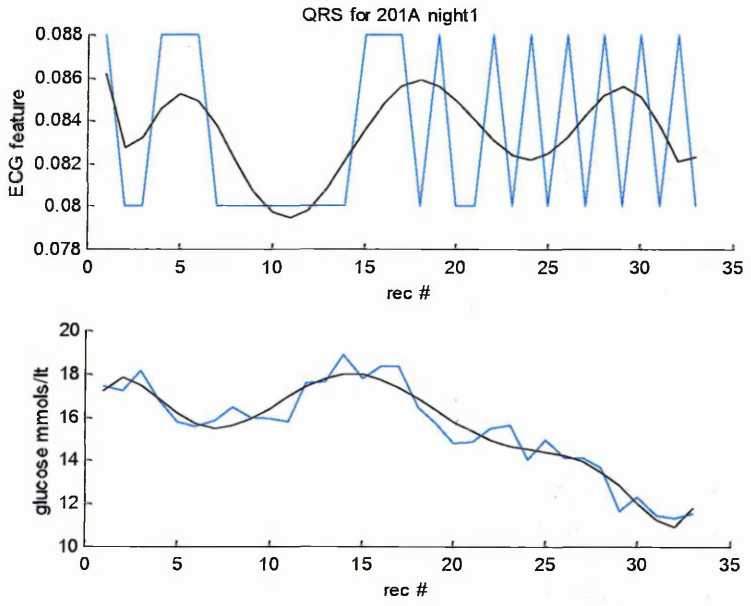


Figure 5.4: BLUE: QRS feature (top) and glucose profile (bottom). BLACK: trend line (polynomial fitted)

The SoTmc feature (along with its uncorrected version) were not used because they were found not to convey any useful information regarding hypoglycaemia. Figure 5.5 depicts this feature along with the glucose profile, for patient 204. This patient experienced a hypoglycaemic episode around the middle of the night and until the end of the acquisition.

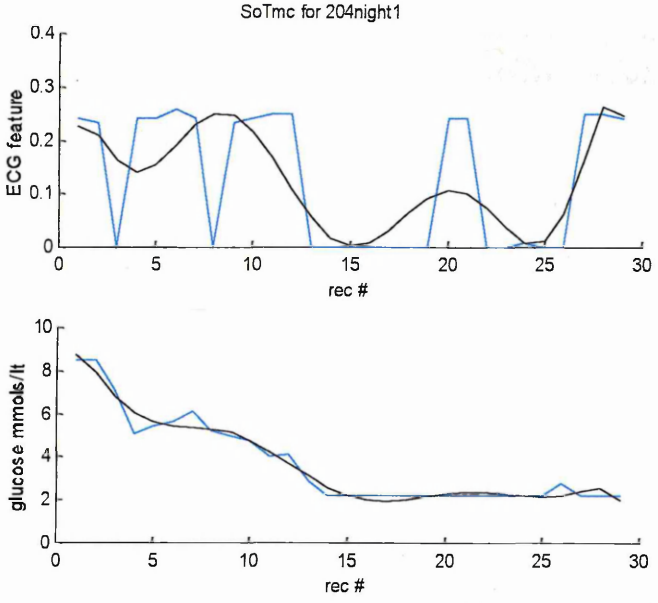


Figure 5.5: BLUE: SoTmc feature (top) and glucose profile (bottom.)BLACK: trend line (polynomial fitted)

Figure 5.6 depicts the TmTo feature (top graph) for the 2nd night of patient 203. The glucose profile is depicted in the bottom graph. It can be seen on the graph that the feature does not appear to be correlated to glucose.

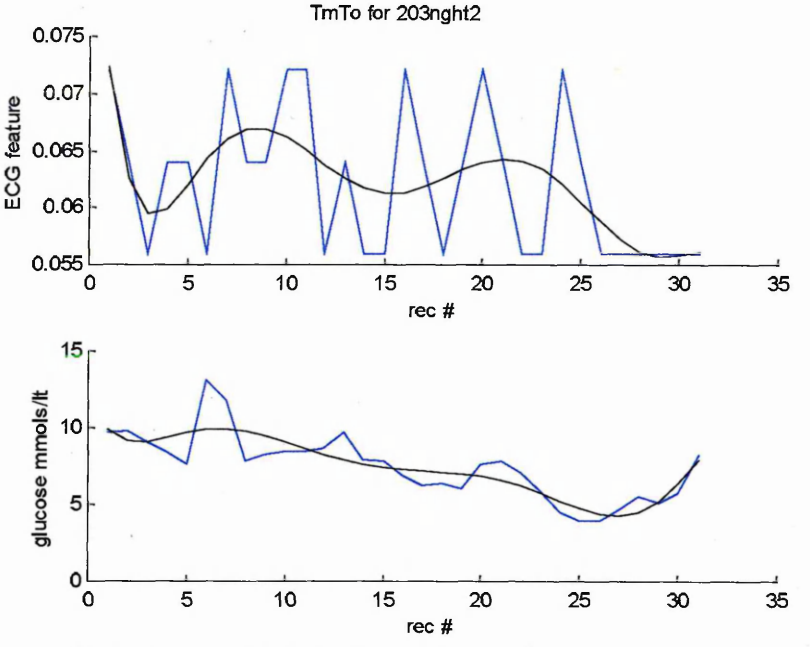


Figure 5.6: BLUE: TmTo feature (top) and glucose profile (bottom)
BLACK: trend line (polynomial fitted)

Figure 5.7 depicts the (Area of ToUo)/(Area of TmaxUo) feature (top graph) for the first night of patient 209. The glucose profile is depicted in the bottom graph. It can be observed on the graph that the feature does not appear to be correlated to glucose.

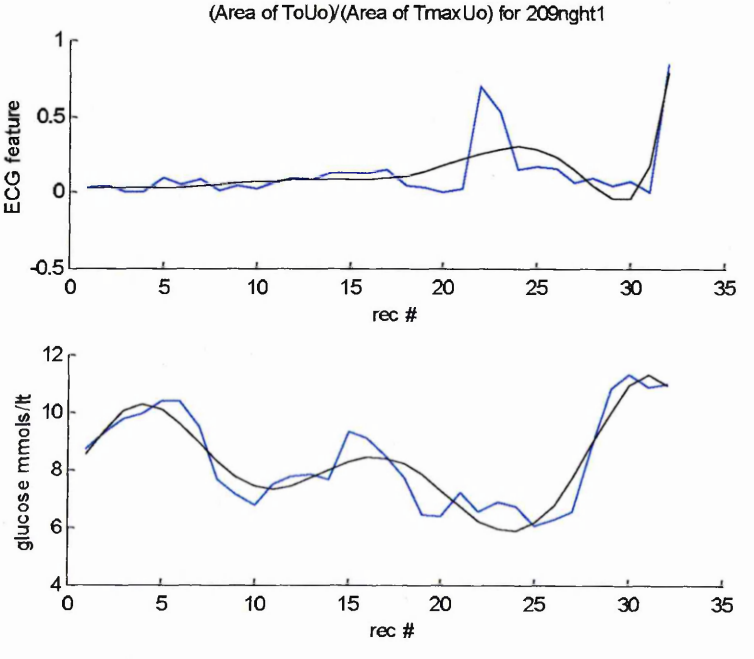


Figure 5.7: BLUE: (Area of ToUo)/(Area of TmaxUo) feature (top) and glucose profile (bottom)
BLACK: trend line (polynomial fitted)

It is noted that some of the features excluded may be useful predictors of hypoglycaemia. They were excluded from this study because they were less important and useful than the features used. Poor performance of a feature not used may be due to weaknesses of the feature extraction algorithm or its implementation in MATLAB and not due to the concept behind the feature being weak. Higher sampling rates may also reveal useful variations in some time-interval features that were excluded.

Correlation coefficients

The correlation coefficient (ρ) was used to calculate correlations between ECG features and glucose in order to gain an insight on the interrelationships between the cardiac function and the glucose variable. Correlation matrices containing the correlations between glucose and all features were calculated. Examples of these correlations are given in Table 5.2 for three of the patients of the dataset:

Table 5.2: Correlation Coefficients for 207-night1 (a), 209-night1 (b) and 201A-night1 (c)

(a)

207nght1	glucose
glucose	1
(Area of ToUo)/Area of TmaxUo)	0.822
QT	0.8
HR	0.773
Total Area	0.754
Time Area (25%) to Area (75%)	0.744
QRS	0.62
QRSc	0.425
SoTmc	0.372
QTc	0.348
TmTo	0.271
%of Total Area to To	0.197
%(Tmax)/(T baseline max)	0.167
Symmetry	0.133
SoTm	0.032
%Tmax/Rmax	0

(b)

209nght1	glucose
glucose	1
QTc	0.315
%of Total Area to To	0.267
QRSc	0.25
QT	0.238
TmTo	0.216
%Tmax/Rmax	0.193
Heart Rate	0.183
SoTmc	0.178
Symmetry	0.133
SoTm	0.109
QRS	0.097
Total Area	0.054
%(Tmax)/(T baseline max)	0.049
(Area of ToUo)/Area of TmaxUo)	0.026
Time Area (25%) to Area (75%)	0.018

(c)

201Anght1	glucose
glucose	1
Heart Rate	0.318
Time Area (25%) to Area (75%)	0.292
QT	0.277
Total Area	0.276
QTc	0.266
QRSc	0.245
(Area of ToUo)/Area of TmaxUo)	0.233
Symmetry	0.212
SoTm	0.209
%of Total Area to To	0.169
%Tmax/Rmax	0.134
SoTmc	0.114
%(Tmax)/(T baseline max)	0.062
TmTo	0.024
QRS	0.012

The features are sorted in descending order according to the magnitude of the correlation coefficient. The remaining columns of the correlation matrices are not included; only the relationships with glucose are presented. By referring to the correlation matrices, the relationships among all the features can be examined. It must be stressed that the correlation coefficient is very useful but not the ultimate method for identifying correlated data, since it can only identify linear relationships. It is useful

when ρ identifies a relationship but if it fails to show a relationship, it does not necessarily mean that one does not exist.

It was found that the rank of features was not very consistent across patients. This was observed by considering all the nights studied and is also apparent in the above tables. For 207-night1 the most correlated feature to glucose was "(Area of ToUo)/(Area of TmaxUo)" while it was the QTc for 209-night1 and the Heart Rate for 201A-night1. The Heart Rate was the only feature consistently ranked highly (top-ranked for 6 out of 13 nights studied).

Figure 5.8 presents a 2-dimensinal scatter diagram of %Tmax/Rmax vs QTc. On the graph, dots denote euglycaemic (normal) ECG feature vectors (pairs of QTc - %Tmax/Rmax values) and circles denote hypoglycaemic feature vectors. Two classes can be formed on the graph: a euglycaemic (Class A) and a hypoglycaemic (Class B). However, the two clinical conditions (euglycaemic, hypoglycaemia) cannot be distinguished by a linear or non-linear decision boundary. The situation is more complex than that and non-linearly enclosed areas (clusters) are needed. An example of an ambiguous point is encircled in the graph. It would be extremely difficult for a classifier using only the above two features to distinguish between the two clinical conditions for the encircled point. More features would be needed as inputs to the classifier to help classify such ambiguous cases. Alternatively, an MLP with two hidden layers would, in theory, be able to classify the cases where a class of one type (e.g. euglycaemia) exists within a class of another (e.g. hypoglycaemia); the class within a class problem.

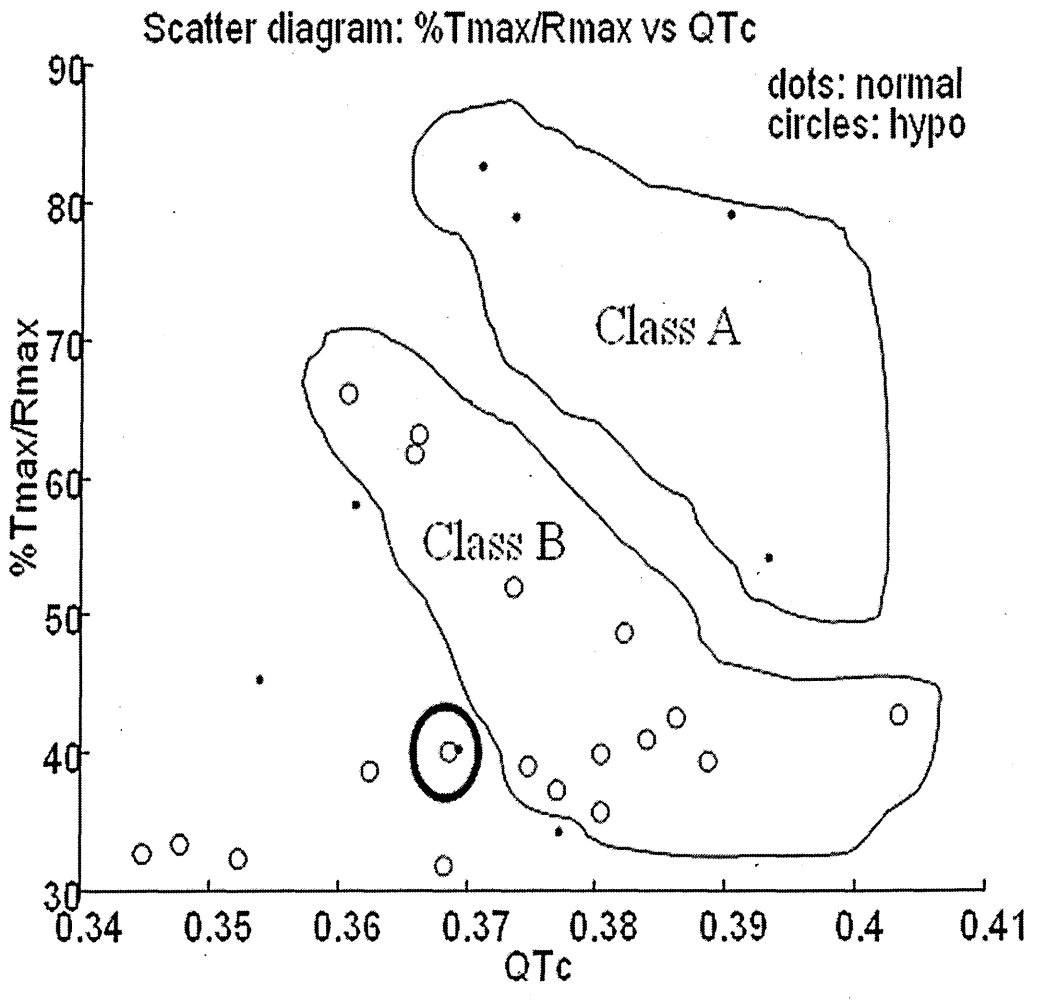


Figure 5.8: Scatter diagram - %Tmax/Rmax vs QTc

Scatter diagrams of higher dimensions (greater than 3) cannot be visualised easily and hence visual inspection is not a sufficient technique for feature selection.

ECG Feature Combinations

A number of ECG features were used in different feature combinations. Selection of the features was carried out by taking in account the clinical significance of features besides using information from visual inspection. The latter approach was used for exclusion of a number of features that were not useful. The number of features fed was 3 in most cases. The 1st derivatives of the features were also included, in most cases, giving a total of 6 features. The features combinations used, consisting of 3 features and their first derivatives were:

- RR, QT, QTc
- RR, QTc, Symmetry
- RR, QTc, %Tmax/Tbaselinemax

- RR, Symmetry, %Tmax/Tbaselinemax
- QTc, Symmetry, %Tmax/Tbaselinemax

Another combination used, consisting of four features, not including 1st derivatives was:

- RR, QTc, Symmetry, %Tmax/Tbaselinemax.

The first derivatives were excluded in order to reduce the number of inputs to the classifier from 8 to 4. The reason behind using the 1st derivatives of the features is illustrated in Figure 5.9.

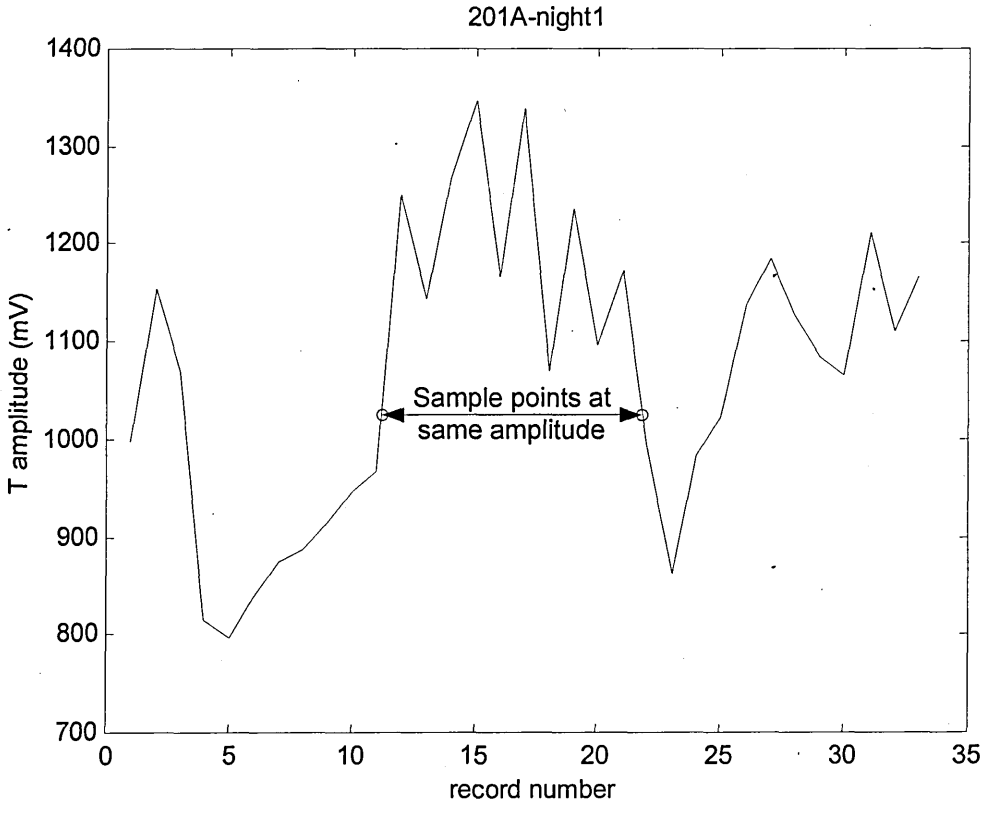


Figure 5.9: Illustration of two sample points considered under a static and dynamic context

The derivative value supplied with the feature value informs whether the feature was increasing or decreasing. The two sample points in Figure 5.9 are the same in a static context but not in a dynamic context. At the LHS sample point the feature value increases whereas it decreases for the RHS sample point. This information could be useful in our case. It had been stressed in the biomedical background discussion (Chapter 1) that T wave flattening and QTc prolongation are observed under hypoglycaemia. The opposite changes most likely do not convey useful information. Therefore feeding the derivative signal to the classifiers may help in ignoring feature changes, to the opposite direction, that are unrelated to hypoglycaemia.

Results from various feature combinations

Twenty one Type 1 adult diabetic patients from the dataset were included in the study each one contributing a maximum of two nights. Each night can contribute a maximum of 33 SAECG cycles. 34 nights were acceptable out of a total of 42. 275 SAECG traces were available out of a total of 1122 after preprocessing and data selection. The classification details are tabulated in Table 5.3 below:

Table 5.3: Parameters for classification based on semi-automatic features

Classification parameters	
Patients:	Adult type1 diabetics
ECG leads used:	YY' from 3-lead orthogonal ECG
Feature extraction:	Semi-Automatic
T wave onset method:	not used
T wave offset method:	tangent method (semi-automatic)
ECG features used:	RR, QT, QTc, Symmetry of T wave, %Tmax/Tbaselinemax in various combinations
Baseline removal method:	1st value baseline only in %Tmax feature
Number of output classes:	2 (euglycaemic – hypoglycaemic)
Hypoglycaemic threshold:	2.5 mmol/lt or 3.5 mmol/lt
Euglycaemic range:	[4 8] mmol/lt
Hypoglycaemic range:	[2.2 2.5] mmol/lt (according to 2.5 mmol/lt threshold) [2.2 3.5] mmol/lt (according to 3.5 mmol/lt threshold)
Ranges excluded:	(2.5 4) & (8 +∞) (according to 2.5 mmol/lt threshold) (3.5 4) & (8 +∞) (according to 3.5 mmol/lt threshold)
cross-validation method:	5-fold
outlier removal:	mean ± 3SD
Number of hidden layers:	1
Number of neurons in hidden layer:	Variable in [2 10]
Glucose sensing method:	MiniMed CGMS

Results Using Features: RR, QTc, Symmetry of T wave, %Tmax/Tbaselinemax

The classification results of the global classifier produced are presented in Table 5.4. The table presents training and test metrics (accuracy, sensitivity, specificity, ratio of undetermined classifications). The assessment of undetermined classifications is rated separately for the hypoglycaemic (undetermined-positives labelled as "und p" on the table) and euglycaemic (undetermined-negatives labelled as "und n" on the table) cases. Although the training accuracy is high, the test accuracy is unsatisfactory being less than a random classifier (i.e. 50% accuracy). The main reason for the poor performance is probably due to inter-patient variability. Further discussions on the issue of inter-

patient variability will be presented later in the chapter when various studies will be combined.

Table 5.4: Classification results for all male patients combined

malemlp_9_mc=0.1#1 hypo@2.5, RR, QTc, Symmetry of T wave, %Tmax/Tbaselinemax (no derivatives)

	Xval	tp	tn	fp	fn	und p	und n	normal	hypos	accuracy	hitrate	tnratio	undet ratio
train	1	125	122	2	2	8	14	138	135	90.48%	98.43%	88.41%	8.06%
	2	99	121	5	8	29	11	137	136	80.59%	92.52%	88.32%	14.65%
	3	125	126	3	2	10	7	136	137	91.94%	98.43%	92.65%	6.23%
	4	101	118	2	5	30	17	137	136	80.22%	95.28%	86.13%	17.22%
	5	123	127	4	2	10	6	137	135	91.91%	98.40%	92.70%	5.88%
	mean:									**87.03%**	**96.61%**	**89.64%**	10.41%
	std:									6.08%	2.66%	2.91%	5.19%
test	1	17	69	142	15	1	14	225	33	33.33%	53.13%	30.67%	5.81%
	2	10	149	66	23	1	8	223	34	61.87%	30.30%	66.82%	3.50%
	3	29	79	124	5	0	12	215	34	43.37%	85.29%	36.74%	4.82%
	4	14	70	134	11	7	21	225	32	32.68%	56.00%	31.11%	10.89%
	5	7	144	59	24	3	22	225	34	58.30%	22.58%	64.00%	9.65%
	mean:									**45.91%**	**49.46%**	**45.87%**	6.94%
	std:									13.67%	24.65%	18.03%	3.19%

Figure 5.10 contains graphs of the 4 input features fed to the classifier (top 4 graphs) and the glucose variable (bottom graph), for the whole of the dataset used. The first half of the dataset (LHS) corresponds to hypoglycaemia (the glucose variable fluctuates around 2.2 mmol/lt) and the second half (RHS) corresponds to euglycaemia.

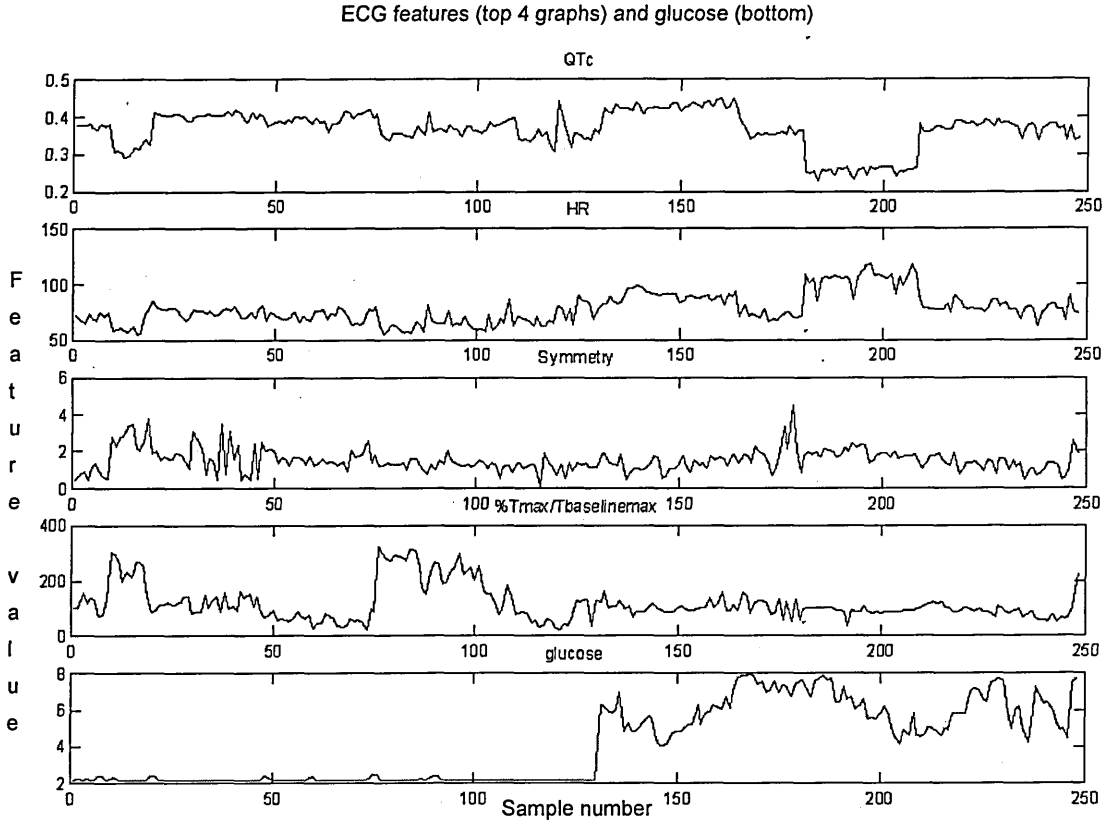


Figure 5.10: Four input features fed to the classifier (top 4 graphs) and glucose (bottom graph)

The classification results for the RR, QT, QTc feature combination are presented in Table 5.5.

Table 5.5: Results for RR, QT, QTc features (including 1st derivatives)

	malemlp_9_mc=0.1#1~1											
	Xval	tp	tn	fp	fn	und p	und n	normal	hypos	accuracy	hitrate	tnratio	undet ratio
train	1	85	131	3	24	32	4	138	141	77.42%	77.98%	94.93%	12.90%
	2	97	117	7	24	20	15	139	141	76.43%	80.17%	84.17%	12.50%
	3	70	106	4	55	15	24	134	140	64.23%	56.00%	79.10%	14.23%
	4	48	69	53	72	19	8	130	139	43.49%	40.00%	53.08%	10.04%
	5	89	80	49	34	17	9	138	140	60.79%	72.36%	57.97%	9.35%
	mean:									**64.47%**	**65.30%**	**73.85%**	11.81%
	std:									13.82%	17.02%	17.76%	2.04%
test	1	0	196	31	31	1	7	234	32	73.68%	0.00%	83.76%	3.01%
	2	0	195	36	34	0	3	234	34	72.76%	0.00%	83.33%	1.12%
	3	1	162	57	28	5	7	226	34	62.69%	3.45%	71.68%	4.62%
	4	21	132	81	11	2	16	229	34	58.17%	65.63%	57.64%	6.84%
	5	12	128	93	17	6	14	235	35	51.85%	41.38%	54.47%	7.41%
	mean:									**63.83%**	**22.09%**	**70.18%**	4.60%
	std:									9.40%	29.96%	13.82%	2.63%

Comparison of feature combinations

Table 5.6 presents the classification train and test accuracy for 5 feature combinations[25]. The optimal number of neurons in each case is recorded. Classification results were produced for both hypoglycaemic thresholds. Therefore the table allows comparisons of both the feature combinations and the hypoglycaemic thresholds.

Table 5.6: Comparison of feature combinations and hypo thresholds

		HrQTQTc	HrQTcSymm	HrQTcTmaxTbas	QTcSymmTmaxTbas	HrSymmTmaxTbas
hypo@2.5	# of neuron	9	7	7	6	8
	train acc	0.645	0.601	0.643	0.611	0.711
	test acc	0.638	0.560	0.523	0.550	0.506
	diff: tr-tst	0.006	0.041	0.120	0.061	0.206

		HrQTQTc	HrQTcSymm	HrQTcTmaxTbas	QTcSymmTmaxTbas
hypo@3.5	# of neuron	7	9	8	10
	train acc	0.545	0.596	0.797	0.783
	test acc	0.398	0.465	0.478	0.464
	diff: tr-tst	0.147	0.131	0.319	0.319

Comparing the feature combinations, consisting of groups of 3 ECG features and their corresponding derivatives, it was found that the combination which gave the highest test

[25] The RR, Symmetry, %Tmax/Tbaselinemax combination is not included in the table for the 3.5 thresh since it was not simulated.

accuracy was the one consisting of the RR, QT and QTc features. This combination yielded a training accuracy of 64.5% and a test accuracy of 63.8%. Although there were combinations with higher training accuracies, the above one gave the highest test accuracy. However, this result has to be treated with caution. By inspecting the table more carefully it is realised that the sensitivity was 0% for two of the cross-validation groups and just 3.45% for a third one. This yields a very low average sensitivity (22.09%). This is compensated by a high value of specificity (70.18%) which raises the accuracy to 63.83%. Although this combination contributed the highest accuracy, detection of positives (hypo events) was poor.

In general, the feature combinations gave very similar results for each hypoglycaemic threshold which makes it difficult to select the best one. The optimal feature combination cannot be concluded easily. Since the performance is not satisfactory, it will not be too crucial and necessary to select the best feature combination. For the 2.5 mmol/lt threshold it was RR, QT, QTc according to the test accuracy as already mentioned. Giving emphasis on the training accuracy, the best combination was RR, Symmetry, %Tmax/Tbaselinemax. However, this combination gave the lowest test accuracy and this indicates that the neural network may have overfitted on the training data. It also had the 2nd largest number of hidden neurons. Comparing the two hypoglycaemic thresholds, it is concluded that the one at 2.5 mmol/lt was a better choice according to the performance of the classifiers. The test accuracies for the 3.5 mmol/lt were very low, being worse than a random classifier. The disadvantage of the 2.5 mmol/lt threshold is that it contributes less hypoglycaemic patterns than the 3.5 threshold[26].

Looking at the size of the hidden layer, it varied slightly among the various configurations. In all the above feature combinations, the best results were achieved by the networks having a large number of hidden neurons; i.e. number of neurons lying in the upper half of the [2 10] interval.

---

[26] This effectively leads to reduction of the data that can be used. In most cases the hypo class was the one contributing less data than the euglycaemic class. Since equal numbers of training patterns from both clinical conditions were used, the hypo class was the one determining how much training data could be used.

5.2.3 Per-patient classification based on automatically extracted ECG features

The approach of producing a classifier tailored for each patient was followed since previous studies (including the study presented in Section 5.2.2) indicated that the performance of global classifiers is unsatisfactory. In this per-patient study 6 patients were used. This was because the rest of them were not suitable for a per-patient study. For such a study each patient must contribute data representative of both clinical conditions (euglycaemia and hypoglycaemia). This is necessary for training the classifiers since they have to be trained on sufficient data from both conditions in order to be able to distinguish them in the future. Only 6 patients from the dataset possessed the above characteristic.

ECG features

ECG features were extracted in a fully automatic way. The features used were RR, RTc (a subsection of QTc), Total Area under the T wave, and HAR (a symmetry measure of the T wave). These features were extracted automatically, in ECGLAB, from the Signal-Averaged ECG cycles[27]. RTc is the corrected version of RT produced by the formula: $RTc = RT/\sqrt{RR}$ (Bazett's formula) [Bazett 1920]. This correction is applied in order to decorrelate the RT interval from the RR (instantaneous heart rate). The Q point detection was manual in the previous study. In this study the R point was used instead of the Q point and the annotation was carried out using an automatic algorithm as opposed to using manual annotations. Q detection is more error-prone to R detection hence the choice of the latter. The RT interval still describes the ventricular repolarisation process and hence it is a valid alternative to the QT. It has been used in the past by Porta et al [Porta 1994, 1998].

The detection of the end of the T wave was carried out using a fully automatic version of the tangent method (msi). Testing of the automatic tangent method was carried out using visual inspection by the researcher[28]. The onset of the T wave was marked using the intersection of the ECG trace with the isoelectric line. Unlike the downslope of the T wave, the T upslope almost always crosses the isoelectric line. This means that using the

[27] The features used in this study were different to the features extracted semi-automatically by the human expert (Cath Davies) working in RHH.

[28] Charilaos Alexakis.

intersection of the T upslope with the isoelectric line was an acceptable algorithm for T onset detection. Such an algorithm though, could not be used for the T end detection because in the event of no crossing of the downslope with the isoelectric line, the algorithm would not provide an annotation for the T end.

Neural Network Classifiers

In this study a different neural network was used for each patient in order to overcome inter-patient variability problems. Global classifiers (i.e. classifiers trained on many subjects) were not produced and the baselines of the ECG features used were not removed (by subtraction of the 1st value of each night). This was because the removal of baseline was considered necessary mainly when producing a global classifier, in order to normalise the ranges of the features by removing offsets.

Five-fold cross-validation was used in order to maximise the data-sets. This means that five classifiers were produced that were trained on different subsections of the data available and the classification performance averaged. This procedure of training 5 neural networks was repeated many times from different initial conditions and the best networks were chosen. 1750 networks were trained for each subject. This is broken down as: 70 different random initial conditions times 5 different hidden layer sizes times 5 different networks due to cross-validation. This task demanded a lot of processing power and many hours of simulation. This is a big disadvantage of neural networks over statistical classifiers such as LDA whose execution time is negligible.

The hidden layer size varied in [2 5]. For each of the above configuration four networks were trained each one with a different number of hidden neurons variable in [2 5]. This was attempted in order to identify the optimal network size for this classification problem. Table 5.7 shows a summary of the classification parameters.

Table 5.7: Parameters for per-patient classification of automatically extracted ECG features

Classification parameters	
Patients:	202, 204, 212, 216, 227, 229 (Adult type1 diabetics)
ECG leads used:	YY' from 3-lead orthogonal ECG
Feature extraction:	Automatic
T wave onset method:	intersection of T upslope with isoelectric line
T wave offset method:	tangent method (fully automatic)
ECG features used:	RR, RTc, Total Area under T wave, HAR (Symmetry of T wave)
Baseline removal method:	none
Number of output classes:	2 (euglycaemic – hypoglycaemic)
Hypoglycaemic threshold:	2.5 mmol/lt
Euglycaemic range:	(2.5 +∞)
Hypoglycaemic range:	[2.2 2.5]
Ranges excluded:	none (in order to maximise the amount of data used)
cross-validation method:	5-fold
outlier removal:	mean ± 3SD
Number of hidden layers:	1
Number of neurons in hidden layer:	Variable in [2 5]
Glucose sensing method:	MiniMed CGMS

The classification results for the six patients are presented in Table 5.8.

Table 5.8: Training and test classification results for six diabetic patients

	training			testing		
patient	accuracy	hitrate	specificity	accuracy	hitrate	specificity
202	89.30%	82.13%	96.47%	70.38%	72.92%	69.29%
204	78.47%	100.0%	77.22%	61.12%	67.00%	61.43%
212	79.00%	100.0%	79.00%	74.92%	75.76%	93.33%
216	73.97%	85.44%	69.26%	73.86%	86.21%	42.67%
227	88.95%	95.69%	92.29%	57.82%	58.86%	76.00%
229	86.47%	90.64%	88.72%	64.14%	43.33%	66.71%
mean	82.70%	92.32%	83.83%	67.04%	67.35%	68.24%
std	6.40%	7.53%	10.34%	7.05%	14.88%	16.70%
min	73.97%	82.13%	69.26%	57.82%	43.33%	42.67%
max	89.30%	100.0%	96.47%	74.92%	86.21%	93.33%

The accuracy, hitrate (sensitivity) and specificity are tabulated for the training and testing data. The table also presents summary results (mean, standard deviation, range). The average train and test accuracies are 82.70% and 67.04% respectively. The accuracy on unseen data is not impressive but shows that the use of neural networks for classification can be a promising approach. The highest test accuracy was observed for subject 212 (74.92%) and the lowest for 227 (57.82%).

By considering the results from Section 5.2.2 and the results from this section, some observations are made regarding the use of a global classifier as opposed to the use of classifiers customised per patient. It is obvious that this section's results are significantly better. The training accuracy for the global classifier is higher than the average training accuracy of the per-patient study but the generalisation on unseen data is very poor and in some cases worse than a random classifier. It is possible that the high performance on the training data may be due to overfitting. This study indicates that the approach of producing a custom classifier for each patient can overcome the severe inter-patient variability problems.

It must be noted here that not all the parameters, involved in the above two classification studies, were kept the same in the two studies. Some of the ECG features used were different. Two different algorithms for the calculation of the T wave symmetry were used in the two studies (Benhorin's symmetry vs HAR). Also the total area under the T wave was used in this study instead of the %Tmax/Tbaselinemax feature used in the global classifier study. Moreover the RTc interval was used in this study while QTc was used in the study of Section 5.2.2. Although different, these features describe the same cardiac process, and the main difference is a change in baseline[29]. Finally, the RR feature was exactly the same for both studies.

The other factor that could have affected the performance is that the maximum number of hidden neurons allowed in this study was 5 instead of 10 used in the global classifier study. This means that the strength of the neural networks was restricted more in the current study which is a step towards prevention of overfitting leading to better performance on unseen data.

[29] Although the main difference is the change in baseline, ambiguous Q point detection can introduce further differences. On the other hand R point detection is straight-forward and does not introduce problems in the comparison of the two features.

5.2.4 Comparison of two feature combinations for the per-patient classification of automatically extracted ECG features, by neural and statistical classifiers

This study was a continuation of the work presented in Section 5.2.3. A few more diabetic patients were recruited and added to the dataset and some additional ECG features were used. Moreover, Linear Discriminant Analysis (LDA)[30] was employed in order to provide a comparison between neural and statistical classifiers.

11 subjects from the dataset were included in this study, since more patients were made available for analysis after the completion of the previous studies. Five ECG features were used in this study namely: RR, RTc, T wave amplitude (Tampl), T wave skewness (skew) and T wave kurtosis (kurt). These features were extracted using automatic algorithms. The onset and end of the T wave were detected using the tangent method.

The 5 ECG features produced were combined in two combinations of 4 features namely RR, RTc, Tampl, skew and RTc, Tampl, skew, kurt. This was so because it was decided to keep the number of input features to a minimum. Feature vectors consisting of more than 4 features were not used. Neural networks were trained using the above feature combinations and comparisons were made in order to identify the best one.

Both ANNs and LDA were used for classification with five-fold cross-validation applied in both cases. Classifiers customised per patient were produced as before. The classification process was almost the same in this study and the one presented in Section 5.2.3, as can be seen in Table 5.9 that summarises the classification parameters. The only differences were the different features and the different numbers of patients used. Neural network and LDA classification results for the 11 subjects and for the feature combination of RTc, Tampl, skew, kurt are given in Tables 5.10 and 5.11 respectively. Similarly for the feature combination of RR, RTc, Tampl, skew, the ANN and LDA results are given in Tables 5.12 and 5.13.

[30] using the Mahalanobis distance metric

Table 5.9: Parameters for per-patient classification

Classification parameters	
Number of output classes:	2 (euglycaemic – hypoglycaemic)
Hypoglycaemic threshold:	2.5 mmol/lt
Euglycaemic range:	(4 8]
Ranges excluded:	(2.5 4] & (8 +∞)
ECG leads used:	YY' from 3-lead orthogonal ECG
Feature extraction	Automatic
ECG features (or feature combinations) used:	▪ RR, RTc, Tampl, skew
	▪ RTc, Tampl, skew, kurt
outlier removal	mean ± 3SD
Number of hidden layers	1
Number of neurons in hidden layer	Variable in [1 5]
Glucose sensing method:	MiniMed CGMS
Baseline removal method	None
Patients	Adult type1 diabetics

Table 5.10 (LHS): ANN classification results (RTc Tampl skew kurt) and Table 5.11 (RHS): LDA Classification results (RTc Tampl skew kurt)

patient	TRAIN			TEST		
	accuracy (%)	hitrate (%)	tnratio (%)	accuracy (%)	hitrate (%)	tnratio (%)
202	89.82	100	89.26	71.52	73.85	74.29
203	93.78	98.46	89.1	87.5	90.46	73.33
204	77.08	100	77.22	58.33	62	58.67
208	88.3	94.86	90.33	66	71	66
212	83.5	100	79	77.66	85.45	83.33
216	79.15	93.89	70.07	76.82	85.61	39.33
220	83.89	97.78	83.89	65.19	70.87	50
223A	82.19	96.19	79.76	69.11	84	68.33
227	93.17	100	89.33	62	65.86	68.67
229	78.21	81.88	78.21	64.19	60	65.14
244	86.67	100	86.67	73.28	80.6	58
mean	85.07	96.64	82.99	70.15	75.43	64.1
std	5.79	5.38	6.59	8.36	10.41	12.17
min	77.08	81.88	70.07	58.33	60	39.33
max	93.78	100	90.33	87.5	90.46	83.33

patient	TRAIN			TEST		
	accuracy (%)	hitrate (%)	tnratio (%)	accuracy (%)	hitrate (%)	tnratio (%)
202	83.42	86.99	79.85	69.62	72.58	62.86
203	91.35	92.05	90.64	82.92	86.52	65.33
204	70.56	66.11	75	50.67	26.67	74.67
208	73.4	70.17	76.63	63.67	46.67	80.67
212	100	100	100	89.96	92	40
216	77.28	82.28	72.28	69.88	71.67	63.33
220	68.89	90.83	46.94	87.1	89.77	0
223A	64.9	76.86	52.95	56.25	84.76	34.72
227	65.95	51.14	80.76	44.09	39.33	58.67
229	79.68	87.18	72.18	36.33	45	35.51
244	87.78	97.78	77.78	83.44	90.68	38.67
mean	78.47	81.94	75	66.72	67.79	50.4
std	11.24	14.75	14.92	18.24	23.95	23.18
min	64.9	51.14	46.94	36.33	26.67	0
max	100	100	100	89.96	92	80.67

By comparing Tables 5.10 and 5.11 it can be seen that the neural network results were superior to those from Linear Discriminant Analysis. The sensitivity and specificity on unseen data were 75.43% and 64.1% respectively for the ANN while they were 67.79% and 50.4% for LDA. Looking at the standard deviation, across patients, of the classification metrics it is realised that the ANN had more uniform performance across patients. In the case of LDA there was greater variation in the performance of the classifier among patients.

Table 5.12 (LHS): ANN classification results (RR RTc Tampl skew)

Table 5.13 (RHS): LDA Classification results (RR RTc Tampl skew)

patient	TRAIN			TEST		
	accuracy (%)	hitrate (%)	tnratio (%)	accuracy (%)	hitrate (%)	tnratio (%)
202	82.13	93.96	75	75.08	79.46	72.86
203	90.64	96.92	84.36	86.83	90.46	72.67
204	76.94	97.78	78.89	62.94	66	61.9
208	87	96.94	86.52	59.44	67.78	58.33
212	90	96	88	76.96	83.02	83.33
216	83.9	94.12	74.93	71.3	79.5	39.33
220	81.67	100	83.89	63.91	72.62	30
223A	87.86	98.1	87.38	63.92	67.62	63.89
227	84.62	99	80.81	60	70.65	62
229	89.87	86.51	96.92	65.62	38	70.33
244	88.75	95.28	91.11	85.67	89.29	84
mean	85.76	95.87	84.35	70.15	73.13	63.51
std	4.31	3.64	6.71	9.79	14.46	16.64
min	76.94	86.51	74.93	59.44	38	30
max	90.64	100	96.92	86.83	90.46	84

patient	TRAIN			TEST		
	accuracy (%)	hitrate (%)	tnratio (%)	accuracy (%)	hitrate (%)	tnratio (%)
202	85.07	90.51	79.63	75.91	84.08	58.93
203	88.21	90.64	85.77	75.67	76.22	72.67
204	64.58	74.17	55.00	56.22	84.67	32.38
208	69.17	51.76	86.57	59.69	37.11	85.00
212	98.00	100.00	96.00	66.91	67.86	43.33
216	83.90	84.56	83.24	67.48	70.30	54.67
220	64.86	93.33	36.39	93.43	95.82	20.00
223A	60.10	70.14	50.05	50.59	68.93	34.44
227	73.52	78.90	68.14	56.82	56.58	57.33
229	82.12	95.38	68.85	46.52	58.00	44.63
244	89.58	95.56	83.61	84.06	88.43	58.67
mean	78.10	84.09	72.11	66.66	71.64	51.10
std	12.29	14.31	18.33	14.51	16.85	18.66
min	60.10	51.76	36.39	46.52	37.11	20.00
max	98.00	100.00	96.00	93.43	95.82	85.00

Tables 5.12 and 5.13 show that for the feature combination of RR, RTc, Tampl, skew the ANN also gave superior classification compared to LDA. The test accuracy was 70.15% for the ANN and 66.66% for LDA. The standard deviation, across patients, of the classification metrics was again greater for the case of LDA.

Comparing the ANN classification results it can be seen that the two feature combinations consisting of four features each (RR-RTc-Tampl-skew & RTc-Tampl-skew-kurt) gave very similar results. The training accuracy differed by 0.69% while the test accuracy was identical (given two significant digits). Comparing the two classifiers used it is summarised that for both feature combinations the ANN were superior to LDA. There was a difference of approximately 4% on the test accuracy, in both cases.

5.2.5 LDA classification of ECG traces modelled by AutoRegressive Modelling

This section presents the classification results when the ECG traces were represented by the coefficients of Autoregressive (AR) models. This type of ECG trace representation was used as an alternative to the approaches where the ECG traces were segmented and ECG features were produced. As mentioned in Section 4.6, a third order AR model was employed which yields four model parameters (a_1, a_2, a_3 and β). The AR parameters (3

coefficients and offset) were classified by LDA. The classification results are given in Table 5.16.

Table 5.16: Classification of AR coefficients by LDA

Patient	Known data					Unseen data				
	accuracy (%)	hitrate (%)	false alarm (%)	tnratio (%)	missed hypos (%)	accuracy (%)	hitrate (%)	false alarm (%)	tnratio (%)	missed hypos (%)
P202	84.52	82.06	13.01	86.99	17.94	76.52	93.33	64.76	35.24	6.67
p203	99.17	100.0	1.67	98.33	0.00	88.58	91.23	21.43	78.57	8.77
p204	97.78	95.56	0.00	100.0	4.44	66.85	56.67	24.76	75.24	43.33
p208	93.98	91.17	3.20	96.80	8.83	85.03	87.50	18.67	81.33	12.50
p216	78.01	89.26	33.24	66.76	10.74	71.39	75.20	46.67	53.33	24.80
p220	79.44	100.0	41.11	58.89	0.00	89.99	91.70	70.00	30.00	8.30
p223A	83.67	86.57	19.24	80.76	13.43	65.20	86.00	64.29	35.71	14.00
P227	82.26	82.71	18.19	81.81	17.29	59.75	64.95	55.33	44.67	35.05
P229	82.82	73.46	7.82	92.18	26.54	52.86	31.67	44.43	55.57	68.33
P244	93.19	88.61	2.22	97.78	11.39	52.77	47.06	10.67	89.33	52.94
mean	87.49	88.94	13.97	86.03	11.06	70.89	72.53	42.10	57.90	27.47
std	7.77	8.33	14.11	14.11	8.33	13.89	21.58	21.73	21.73	21.58
min	78.01	73.46	0.00	58.89	0.00	52.77	31.67	10.67	30.00	6.67
max	99.17	100.0	41.11	100.0	26.54	89.99	93.33	70.00	89.33	68.33

Figure 5.11 presents a comparison of the two approaches of ECG representation, i.e. using ECG features and alternatively using AR coefficients to describe each cardiac cycle. The results in the case of using ECG features originate from the study presented in Section 5.2.4. The bars in Figure 5.13 represent training and test accuracy. Labels "ECGfeat-NNET" and "ECGfeat-LDA" correspond to the approaches of using ECG features classified by neural networks and linear discriminant analysis respectively. AR-LDA corresponds to the approach of using AR coefficients for ECG representation classified by linear discriminant analysis. The AR coefficients were not classified using neural networks due to time constraints. Comparing the LDA classification results for the two ECG representation approaches it is observed that the AR modelling yielded higher training and test accuracy. This gives some indication that the approach of modelling whole segments of interest of the ECG cycle is a promising one for detecting subtle cardiac abnormalities related to hypoglycaemia.

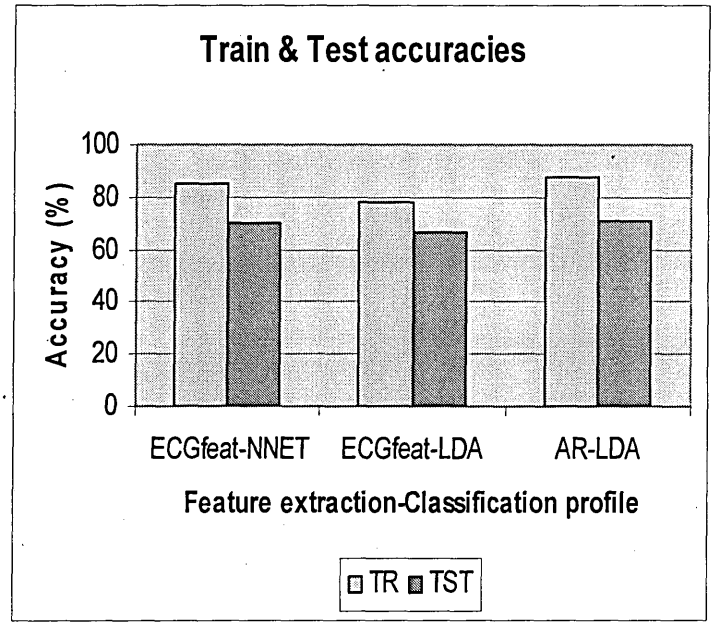


Figure 5.11: Summary results for the two ECG representation approaches

5.2.6 Investigation of improved preprocessing on the classification of ECG traces represented by the RTc and T amplitude features

This section presents a study where less ECG features were fed to the classifier but extra preprocessing was applied. Only the main two ECG features that were quantifying the changes dictated by the research hypothesis were used. The main changes are broken down to delayed Ventricular Repolarisation and flattened T waves. Therefore the two main features involved are the RTc (or alternatively QTc) and the T wave amplitude. The extra preprocessing step involved Moving Average filtering. The aim was to test whether using additional preprocessing techniques would improve the performance and whether it would allow reduction of the number of input ECG features needed. This section also discusses the effect of fluctuations and transient ECG feature changes on the classification performance. Classifiers used were tailored to the needs of each patient and global classifiers were not used. The details of this classification study are summarised in Table 5.17.

Table 5.17: Classification parameters

Classification Parameters	
Patients:	202 (both nights), 203-night1, 204, 212-night2, 227 (both nights), 244-night1 (Adult type1 diabetics)
ECG leads used:	YY' from 3-lead orthogonal ECG
Feature extraction:	Automatic
T wave onset method:	tangent method (fully automatic)
T wave offset method:	tangent method (fully automatic)
ECG features used:	RTc, Tampl
Baseline removal method:	none
Number of output classes:	2 (euglycaemic – hypoglycaemic)
Hypoglycaemic threshold:	2.5 mmol/lt or 3 mmol/lt
Euglycaemic range:	(2.5 +∞] mmol/lt (according to 2.5 mmol/lt threshold) (3 +∞] mmol/lt (according to 3 mmol/lt threshold)
Hypoglycaemic range:	[2.2 2.5] mmol/lt (according to 2.5 mmol/lt threshold) [2.2 3] mmol/lt (according to 3 mmol/lt threshold)
Ranges excluded:	none (in order to maximise the amount of data used)
ECG features used:	RTc, Tampl
cross-validation method	2-fold (data was partitioned to train and test set and this partitioning was repeated 1000 times after randomising the data each time)
outlier removal	mean ± 3SD
Glucose sensing method:	MiniMed CGMS

Effect of ECG feature fluctuations and transient ECG feature changes on Classification Performance

Pattern classification as it was carried out up to this point of the thesis, suffers from a number of problems discussed here. This section analyses how the fluctuations of ECG features are degrading classification performance. ECG feature values sometimes fluctuate significantly and this is unrelated to hypoglycaemia. Moreover, some ECG features change in value as hypoglycaemia occurs but often the change is a short-term one and the feature values recover to their previous ranges associated with euglycaemia, although the patient remains in hypoglycaemia. Such feature changes will be referred to as "transient ECG feature changes". The ECG features undergo changes in magnitude along with the onset of hypoglycaemia but such changes are transients and the feature values do not settle to a steady-state level for the duration of the hypoglycaemic period. Before the hypoglycaemic period finishes the feature values often return to their original values. This introduces ambiguity since the same feature magnitude can be encountered both under euglycaemia and hypoglycaemia. This phenomenon is illustrated in Figure 5.12

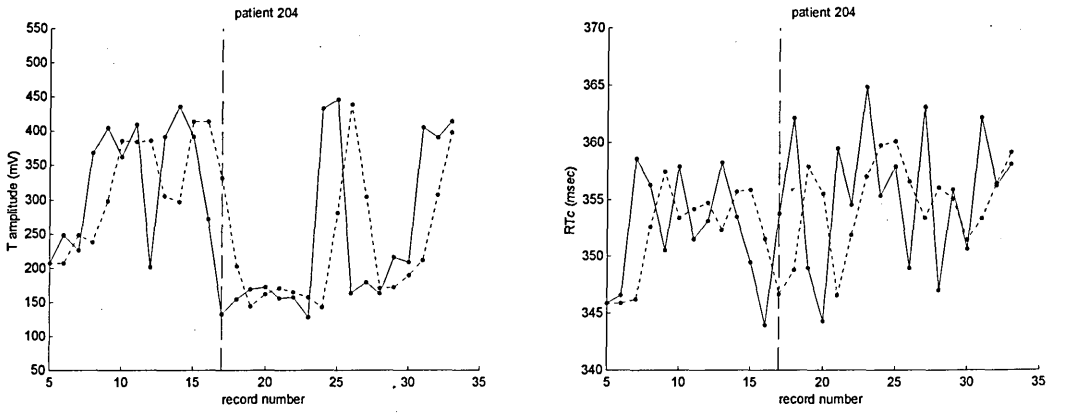


Figure 5.12: Tampl (LHS) and RTc (RHS) graphs (solid lines) including MA values (dotted lines) for subject 204

The LHS graph presents the Tampl feature (solid) and the RHS graph, the RTc feature (solid). The dotted lines describe the moving average of the two features. The vertical dashed line marks the onset of hypoglycaemia. Records to the left of the dashed line are normal while records to the right are hypoglycaemic. It can be seen on the Tampl graph that there is a sharp drop around the onset of hypo. The low Tampl value is preserved for a while but for records 24 and 25 the value of the feature returns to amplitudes, similar to those before the hypo event. This change happens although the patient remains in hypoglycaemia. A classifier will associate high T amplitudes (around 400-450 mV) with euglycaemia and later on it will be presented with similar amplitudes that correspond to hypoglycaemia. Such ambiguity is confusing the classifier and making its task very difficult. Extra ECG features added may resolve the ambiguities and improve classification.

Looking at the RTc graph, similar problems exist. The signal has a lot of fluctuations and this again causes ambiguities between feature values corresponding to euglycaemia and hypoglycaemia. For instance, record 13 has a similar value to record 19 although they correspond to two different clinical conditions. These ambiguities need to be overcome and the useful information in the feature extracted. For instance, the feature has an upward trend as the patient goes into hypoglycaemia which is useful for detection of the condition. Also there is a significant increase in feature value for a couple of successive sampling instants around the onset of hypo. However, if static pattern classification is carried out then this significant event will be masked by similar feature values occurring before and after it in the night.

Table 5.18 presents training and test classification results for patient 204 using LDA when only RTc and Tampl are used as inputs. The data was partitioned to training and test and classification was carried out by LDA. This process was repeated 1000 times and average performance calculated. This approach was chosen over the use of 5-fold cross-validation because in the latter case the cross-validation groups were very small for calculation of statistics. Splitting the data in half (i.e. 2-fold cross-validation) maximised the size of the train and test sets. Since this was repeated many times, and the classifiers were assessed on the average performance, we were confident that the data-set formation was not biasing the classification in any way. A hypoglycaemic threshold of 2.5 mmol/lt was used to define euglycaemia and hypoglycaemia.

Table 5.18 p204 LDA classification results using RTc and Tampl (raw features, no exclusion of records) for a hypoglycaemic threshold of 2.5 mmol/lt

	TRAIN			TEST		
	accuracy	hitrate	specificity	accuracy	hitrate	specificity
mean	71.67%	55.29%	86.00%	56.43%	43.86%	69.00%
std	10.31%	22.49%	10.99%	10.77%	21.99%	24.79%

It can be seen from the table that the metrics on the test data are poor with an accuracy of 56% approximately. Poor performance on this patient is due to the existence of ambiguous variations of the ECG features during the night that are confusing the classifier. Specifically on the Tampl feature, there are samples with similar amplitude that correspond to both clinical conditions and hence the classifier fails to distinguish between the two; this was discussed earlier and illustrated in Figure 5.14.

In order to investigate the effect of the fluctuating ECG features on the classification performance, the ambiguous records according to the Tampl feature were removed. The records removed were: 5, 6, 7, 12, 24, 25, 31, 32, 33 and LDA classification was repeated. The removal of these records is not set up aiming to improve the numbers in the classification performance. The justification for the removal of these records is as follows:

- records 5-7 are the first 3 records of the night corresponding to the transitional period of the patient falling asleep and including the early stages of sleep. These records can introduce ECG feature variations that are related to the transition into sleep but which are unrelated to hypoglycaemia.

- record 12, although not an outlier, is a record where a big change in feature value happens and is unrelated to hypoglycaemia. In some of the studies, such changes are assumed to be insignificant since they only occur at one sampling instant. This record could be suppressed during preprocessing and set to the value of the previous sample or a moving average value.
- Records 24 and 25 belong to the hypoglycaemic part of the night but the Tampl values correspond to levels of magnitude encountered under euglycaemia. These samples are removed in order to investigate the impact of such ambiguous records on the classification performance.
- Records 31-33 were removed for the same reason as records 24, 25.

The LDA classification results after removing the ambiguous feature vectors are tabulated in Table 5.19 (the results with the ambiguous records included, which were already presented in Table 5.18, are included again for easy comparisons). As it was expected, the results significantly improved after removal of the ambiguous records. The test accuracy increased from 56.4% to 81.2%.

Table 5.19: p204 LDA results using RTc and Tampl (with ambiguous records excluded)

		TRAIN			TEST		
		accuracy	hitrate	specificity	accuracy	hitrate	specificity
ambiguous records excluded	mean	93.70%	94.00%	93.40%	81.20%	85.80%	76.60%
	std	8.00%	9.64%	10.66%	12.08%	18.04%	28.58%
raw features	mean	71.67%	55.29%	86.00%	56.43%	43.86%	69.00%
	std	10.31%	22.49%	10.99%	10.77%	21.99%	24.79%

Moving Average Filtering

Manually removing ambiguous records was not a practical preprocessing step. In order to systematically tackle problems of ambiguous feature vectors, an extra preprocessing step was added before the classifier. This step involved calculating the moving average profile of each feature. This produced the necessary filtering that suppressed some of the fluctuations that were causing problems. The equation of the Moving Average filter used is given below in its generic form for an n^{th} order filter:

$$y(k)=\begin{cases}\dfrac{\sum_{i=1}^{n}x(k-i)}{n}, \text{ for } n<k \quad n,k\in\aleph^{+} & eq^{n}\ 5.7a\\ \dfrac{\sum_{i=1}^{k-1}x(k-i)}{k-1}, \text{ for } n\geq k \quad n\in\aleph^{+},k\in\aleph^{+}\wedge\ k>1 & eq^{n}\ 5.7b\\ x(k), \text{ for } k=1 & eq^{n}\ 5.7c\end{cases}$$

x(k) is the raw feature value at sample k, and *y(k)* is the MA filtered version.

It can be seen from the above equations that if there are not enough previous values to fill the moving window then a smaller window is used with as much data there is available. Figure 5.13 shows the graph of the raw Tampl feature (top) and the graphs of the filtered versions of Tampl using MA filters of a length 3 (middle) and 5 (bottom).

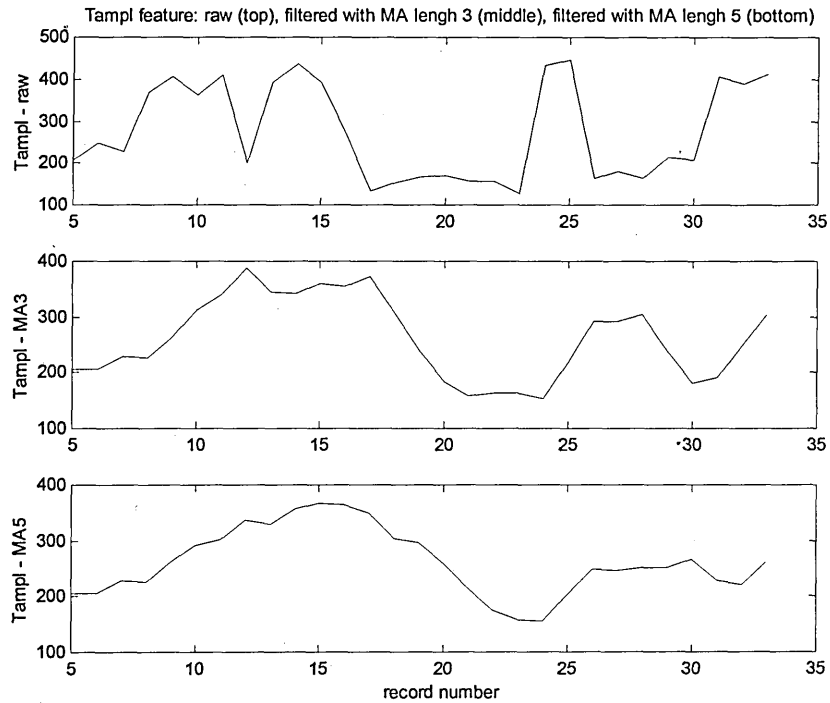


Figure 5.13: raw Tampl feature (top) and its filtered versions using MA filters of a length 3 (middle) and 5 (bottom).

A moving window is used and the mean value of the data in that window is calculated. The classification process could be run multiple times with different window sizes in order to tune the length of the window by trial and error. In-depth investigations on the optimal window size were not carried out. MA filters of length 3 and length 5 were tried. A length 3 MA filter was used since it gave better results in terms of classification accuracy. The classification results for the length 5 filter are not included. The LDA classification results for p204 when using the length-3 MA filter are given in Table 5.20. It must be noted that all the data that the patient contributed were used; no records were excluded.

Table 5.20: p204 LDA results using RTc and Tampl (after applying MA length 3 (on all records))

(hypoglycaemic threshold =2.5)						
	TRAIN			TEST		
	accuracy	hitrate	specificity	accuracy	hitrate	specificity
mean	89.39%	89.73%	89.09%	82.36%	83.27%	81.46%
std	7.31%	7.75%	9.10%	10.84%	21.28%	20.73%

The results are very close to those obtained when the ambiguous records were removed. Looking at the test data, the classification accuracy is very similar in both cases being higher when using MA filtering. The sensitivity has decreased but this is compensated by a greater increase in specificity.

Once there were indications that MA filtering could improve the classification performance, it was utilised in the classification process for a number of patients. Two hypoglycaemic thresholds were used. Classification was carried out for both hypoglycaemic thresholds to allow comparisons. The LDA classification results for the 8 patient-nights using MA preprocessing and the hypoglycaemic threshold set at 2.5 mmol/lt are given in Table 5.21.

Table 5.21: LDA results for hypo thresh=2.5, with length 3 MA preprocessing

	TRAIN			TEST		
patient	accuracy	hitrate	specificity	accuracy	hitrate	specificity
p202A	95.10%	97.10%	80.90%	66.70%	95.90%	37.60%
p202	88.10%	93.50%	86.20%	83.30%	81.90%	84.80%
p203night1	81.30%	59.60%	87.10%	59.10%	38.60%	79.60%
p204	89.40%	89.70%	89.10%	82.40%	83.30%	81.50%
p212nght2	80.10%	77.80%	92.60%	68.40%	77.30%	59.50%
p227night1	79.20%	85.40%	65.90%	63.80%	80.90%	46.70%
p227night2	87.60%	90.20%	83.40%	78.60%	80.90%	76.30%
p244night1	85.90%	91.50%	79.50%	76.70%	86.00%	67.50%
mean	85.80%	85.60%	83.10%	72.40%	78.10%	66.70%
std	5.40%	12.00%	8.20%	9.10%	16.90%	17.30%

Classification is repeated for a threshold of 3 mmol/lt and the results are tabulated in Table 5.22.

Table 5.22: LDA results for hypo thresh=3, with length 3 MA preprocessing

	TRAIN			TEST		
patient	accuracy	hitrate	specificity	accuracy	hitrate	specificity
p202A	89.90%	91.80%	81.30%	68.20%	86.20%	50.20%
p202	99.00%	100.00%	98.90%	83.70%	69.20%	98.10%
p203night1	81.50%	60.70%	87.00%	57.90%	37.80%	78.10%
p204	92.10%	88.40%	94.20%	79.30%	67.20%	91.50%
p212nght2	76.10%	74.60%	83.60%	64.50%	72.80%	56.10%
p227night1	80.90%	87.50%	69.60%	67.70%	82.30%	53.10%
p227night2	86.80%	91.10%	82.50%	77.00%	78.80%	75.30%
p244night1	86.00%	91.70%	79.60%	76.70%	86.10%	67.30%
mean	86.50%	85.70%	84.60%	71.90%	72.60%	71.20%
std	7.20%	12.30%	9.00%	8.70%	15.80%	17.80%

Comparing the two hypoglycaemic thresholds when using the MA-length3 filter, it is observed that the results are very similar. The only easily observable difference is in the test sensitivity and specificity. For the threshold at 3 mmol/lt, the sensitivity and

specificity have similar values while for the 2.5 mmol/lt the sensitivity is significantly greater than the specificity. These variations average out so the accuracy figures for the two hypoglycaemic thresholds are very similar. These can be visualised in the bar chart depicted in Figure 5.14.

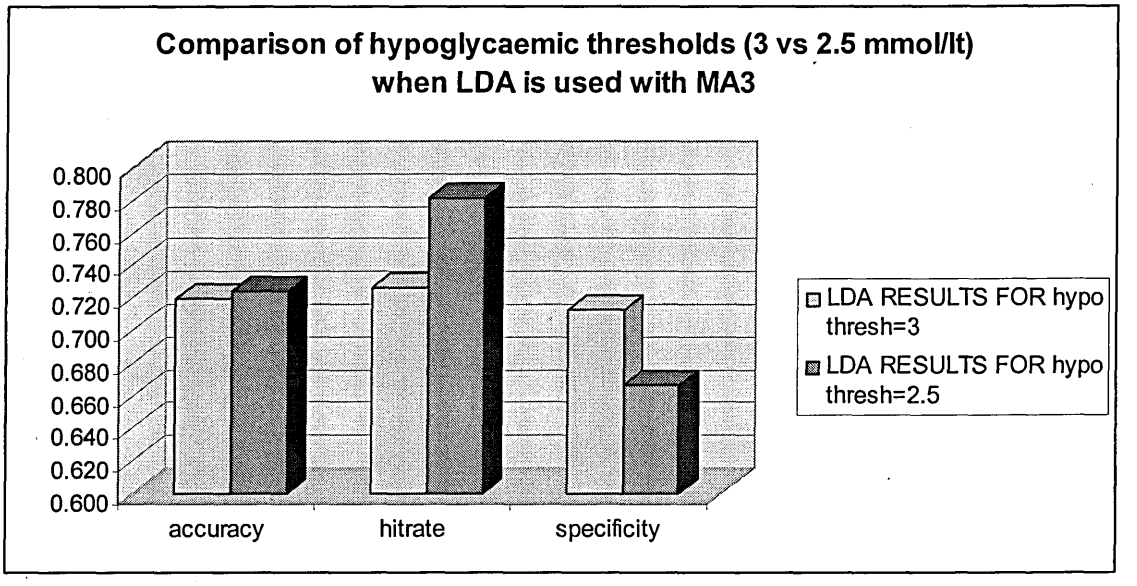


Figure 5.14 Comparison of hypoglycaemic thresholds

Concluding on the best hypoglycaemic threshold is not an easy task especially when only considering the metric of classification accuracy on unseen data. Although the two thresholds gave similar results, the one at 3 mmol/lt is selected for further use since it yielded more balanced performance towards euglycaemia and hypoglycaemia; that is, sensitivity and specificity were very close in value. The threshold at 2.5 mmol/lt gave very good performance in detecting hypo events but the performance in terms of detecting euglycaemic events was significantly lower.

Continuing the analysis of the effect of the introduction of MA filtering, the LDA classification results when no MA pre-processing is carried out and for a hypoglycaemic threshold of 3 mmol/lt are given in Table 5.23.

Table 5.23: LDA results for hypo threshold=3 (no MA preprocessing)

	TRAIN			TEST		
patient	accuracy	hitrate	specificity	accuracy	hitrate	specificity
p202A	93.00%	96.80%	75.60%	69.20%	94.20%	44.20%
p202	79.00%	100.00%	75.90%	65.80%	58.70%	73.00%
p203night1	81.80%	63.30%	86.70%	62.70%	44.70%	80.70%
p204	78.00%	67.90%	83.50%	61.10%	49.50%	72.60%
p212nght2	84.80%	82.00%	98.90%	68.90%	79.70%	58.20%
p227night1	62.50%	56.50%	72.80%	49.50%	46.70%	52.30%
p227night2	64.30%	49.60%	79.00%	49.20%	35.40%	63.00%
p244night1	81.20%	74.20%	89.30%	70.00%	64.80%	75.10%
mean	78.10%	73.80%	82.70%	62.00%	59.20%	64.90%
std	10.20%	18.20%	8.70%	8.50%	19.70%	12.60%

By inspecting Table 5.23 it is realised that the classification results without employing MA preprocessing are significantly inferior. The test sensitivity and specificity are 59.2% and 64.9% whereas they were 72.6% and 71.2% respectively when MA (length 3) was used. It can be concluded that the MA preprocessing step significantly improved performance by dealing with ambiguities in the ECG features and masking fluctuations unrelated to hypoglycaemia. Hence the task of the classifier was simplified significantly.

In order to allow further comparisons, a k-Nearest Neighbour (kNN) classifier was also employed as a benchmark for the LDA. A squared-Euclidean distance metric was used in the kNN. The kNN classifier was implemented using the "knnclassify" M file from the MATLAB bioinformatics toolbox. The number of nearest-neighbours in the classifier were set to 3 or 5 and we settled to k=3 since it gave better performance. The process of partitioning the data into training and testing was the same as with the LDA process. That is, the data was partitioned to training and testing sets and then fed to the kNN classifier; this process was repeated 1000 times and averages calculated.

Neural networks were not used in this classification study due to lack of processing power and time constraints. The kNN classification results with and without MA preprocessing for a threshold of 3 mmol/lt are tabulated in Tables 5.24 and 5.25 respectively.

Table 5.24: kNN results (k=3) for hypo threshold=3, with length 3 MA preprocessing

	TRAIN			TEST		
patient	accuracy	hitrate	specificity	accuracy	hitrate	specificity
p202A	91.38%	98.33%	60.10%	63.78%	87.30%	40.27%
p202	94.47%	63.30%	99.15%	74.52%	50.03%	99.00%
p203night1	82.04%	29.90%	95.95%	45.13%	1.00%	89.25%
p204	82.30%	71.38%	88.25%	69.92%	57.75%	82.08%
p212nght2	83.96%	95.67%	25.45%	52.41%	92.80%	12.03%
p227night1	86.80%	91.75%	78.31%	75.09%	85.72%	64.47%
p227night2	84.64%	84.50%	84.79%	66.38%	66.90%	65.85%
p244night1	86.50%	89.60%	82.96%	70.91%	74.03%	67.79%
mean	86.50%	78.10%	76.90%	64.80%	64.40%	65.10%
std	4.40%	22.90%	24.00%	10.70%	29.70%	27.90%

Table 5.25: raw kNN RESULTS (k=3) FOR hypo thresh=3

	TRAIN			TEST		
patient	accuracy	hitrate	specificity	accuracy	hitrate	specificity
p202A	86.08%	98.84%	28.63%	46.90%	91.67%	2.13%
p202	94.32%	57.40%	99.86%	69.00%	39.20%	98.80%
p203night1	85.73%	36.40%	98.88%	53.00%	13.63%	92.38%
p204	84.31%	72.75%	90.61%	61.64%	47.62%	75.67%
p212nght2	83.66%	95.54%	24.28%	52.41%	93.68%	11.15%
p227night1	71.27%	85.74%	46.46%	41.10%	62.83%	19.37%
p227night2	77.79%	87.80%	67.79%	55.70%	67.69%	43.71%
p244night1	77.42%	84.94%	68.83%	62.49%	72.33%	52.64%
mean	82.57%	77.43%	65.67%	55.28%	61.08%	49.48%
std	6.99%	21.13%	30.15%	8.99%	26.95%	37.07%

Using MA preprocessing in cascade with the kNN classifier improved the performance compared to when feeding the raw ECG features. The accuracy on unseen data increased from 55.28% to 64.8%. Further comparisons between the two classifiers will be presented in the Discussion Section (5.3).

Using a dynamic threshold to assess the magnitude of the ECG features

As indicated earlier, using the raw ECG feature values has the disadvantage that ambiguous events occur that confuse the classifiers. Besides using the MA signal instead of the raw one, an alternative approach was to use the raw ECG feature data but subtract the current MA value from the current feature value. This way the feature values would not be filtered (smoothed) which is what the MA process does, but at the same time, the fluctuations would be suppressed. Subtracting the MA value from the feature value effectively compares the current feature value with a moving threshold provided by the MA value. This, in theory, should absorb some of the fluctuations and reduce the reliance of the classifier on raw feature values.

The way this normalisation process works is that instead of using the raw feature values as the criterion of whether the ECG trace corresponds to normality or hypoglycaemia, the feature is compared to a dynamic threshold. The dynamic threshold equals the MA value at the current sampling instant. To put it simply, each input that the classifier receives is produced by comparing the current ECG value to a dynamic threshold (MA value) instead of using the absolute magnitude of the feature. So the classifier would have to classify the difference of the current feature value from the current MA value. The transformed ECG features for subject 204 are given in Figures 5.15 (Tampl) and 5.16 (RTc).

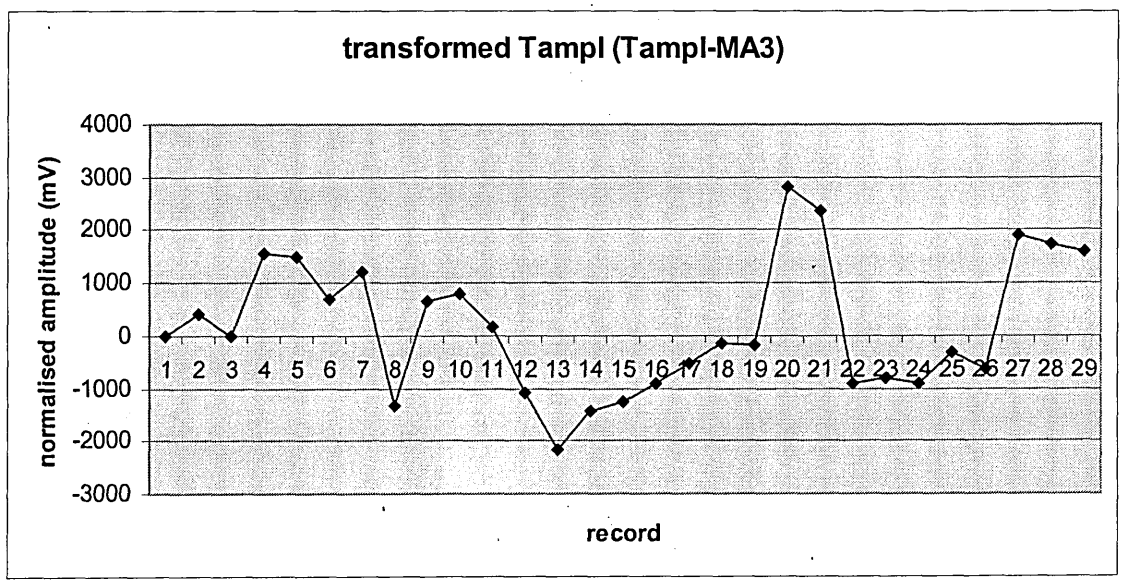


Figure: 5.15 Transformed (normalised) Tampl feature for subject 204

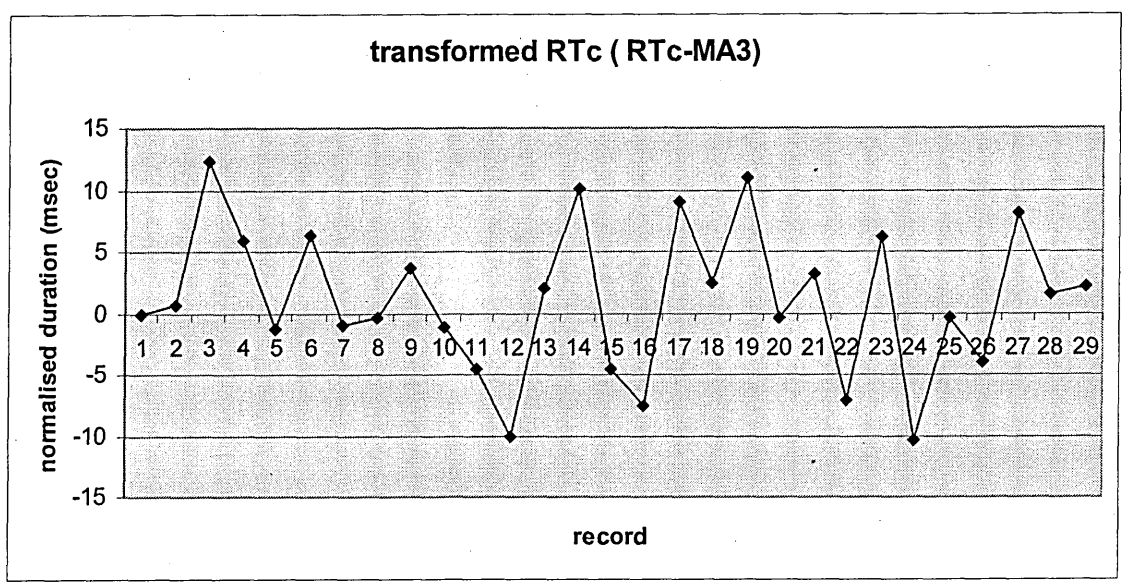


Figure: 5.16 Transformed (normalised) RTc feature for subject 204

It can be seen that after the normalisation by subtracting the MA signal, the features fluctuate around zero. This new transformed range of feature values would be useful if data were mixed together to produce global datasets for global classifiers to be trained on. However, this transformation will not guarantee the elimination of inter-patient variability.

LDA classification results using this approach (ECGfeat – MA) are presented in Table 5.26. A hypoglycaemic threshold of 3 mmol/lt was used and the MA window size used to calculate the dynamic thresholds was set to 3 as in the previous studies.

Table 5.26:LDA results for hypo thresh=3, adaptive features (ECGfeat-MA3) preprocessing

	TRAIN			TEST		
patient	accuracy	hitrate	specificity	accuracy	hitrate	specificity
p202A	80.57%	83.79%	66.88%	56.13%	82.10%	30.17%
p202	87.40%	80.93%	88.32%	69.43%	51.95%	86.90%
p203night1	77.60%	62.13%	81.73%	56.14%	34.13%	78.15%
p204	71.28%	33.20%	92.05%	51.13%	19.27%	83.00%
p212nght2	82.10%	82.31%	81.10%	58.61%	77.45%	39.78%
p227night1	66.88%	64.07%	71.71%	49.62%	55.30%	43.93%
p227night2	73.23%	65.89%	80.56%	61.95%	50.34%	73.56%
p244night1	62.27%	63.84%	60.47%	39.69%	43.66%	35.71%
mean	75.17%	67.02%	77.85%	55.34%	51.77%	58.90%
std	8.34%	16.46%	10.72%	8.86%	20.80%	23.61%

By inspecting the above table, it is realised that this preprocessing step did not live up to the expectations. The results appear to be worse than when the raw ECG features were used. In order to investigate the reasons behind the unsatisfactory performance of the new preprocessing step, a number of scatter diagrams of Tampl versus RTc were plotted. Figure 5.17 presents such a scatter diagram when the raw feature values are used. Figures 5.18 and 5.19 present the scatter diagrams when MA filters of length 3 and 5 were used, respectively. Finally, Figure 5.20 presents a scatter diagram when the dynamic threshold is used, based on a MA-length3.

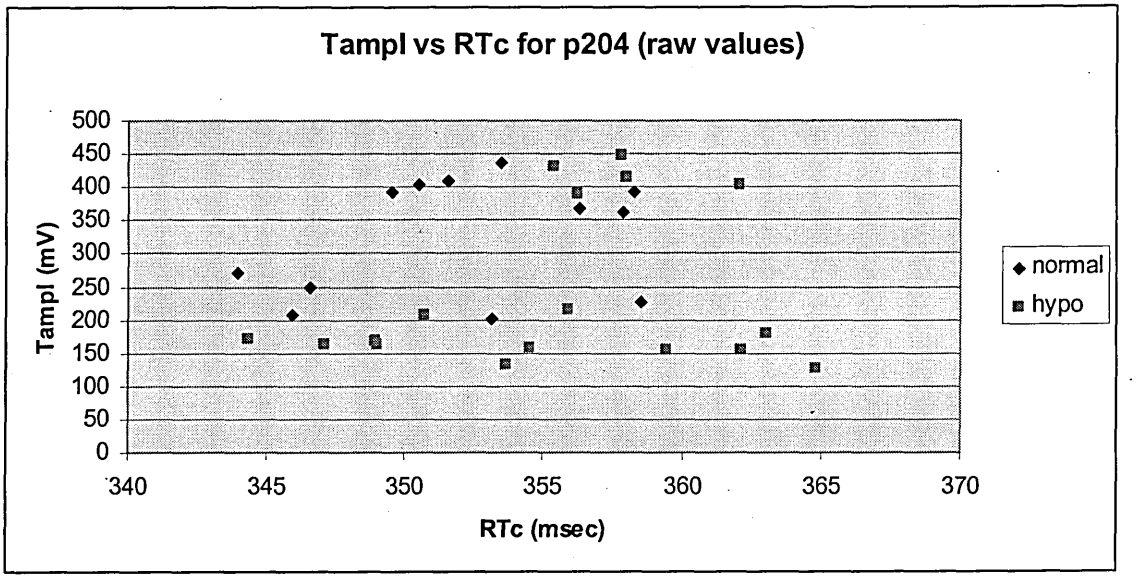


Figure 5.17: Tampl vs RTc scatter diagram for raw ECG feature values of patient 204

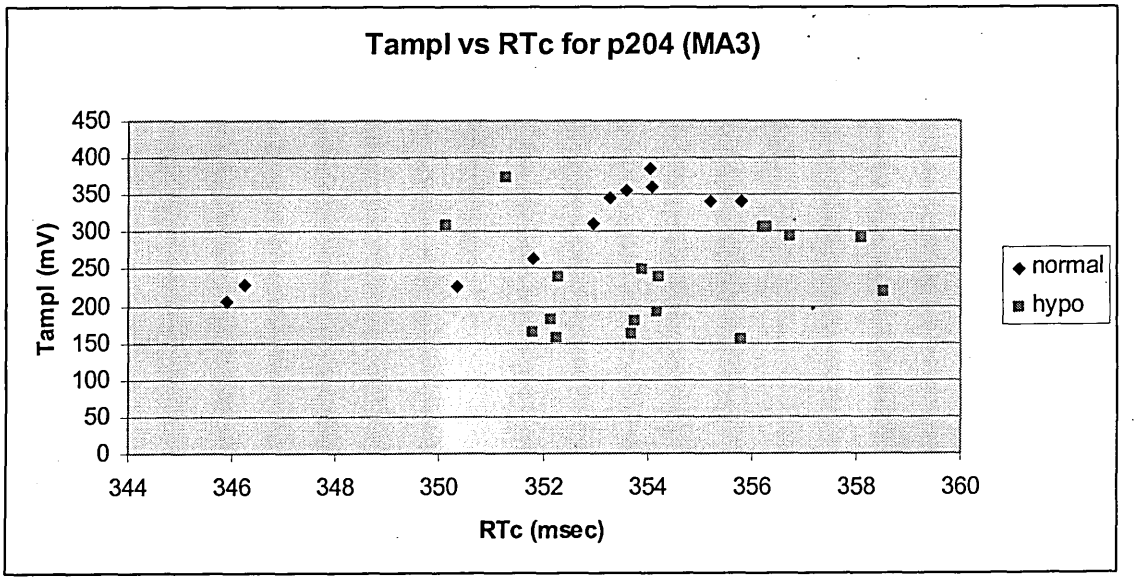


Figure 5.18: Tampl vs RTc scatter diagram for MA3 filtered ECG features (patient 204)

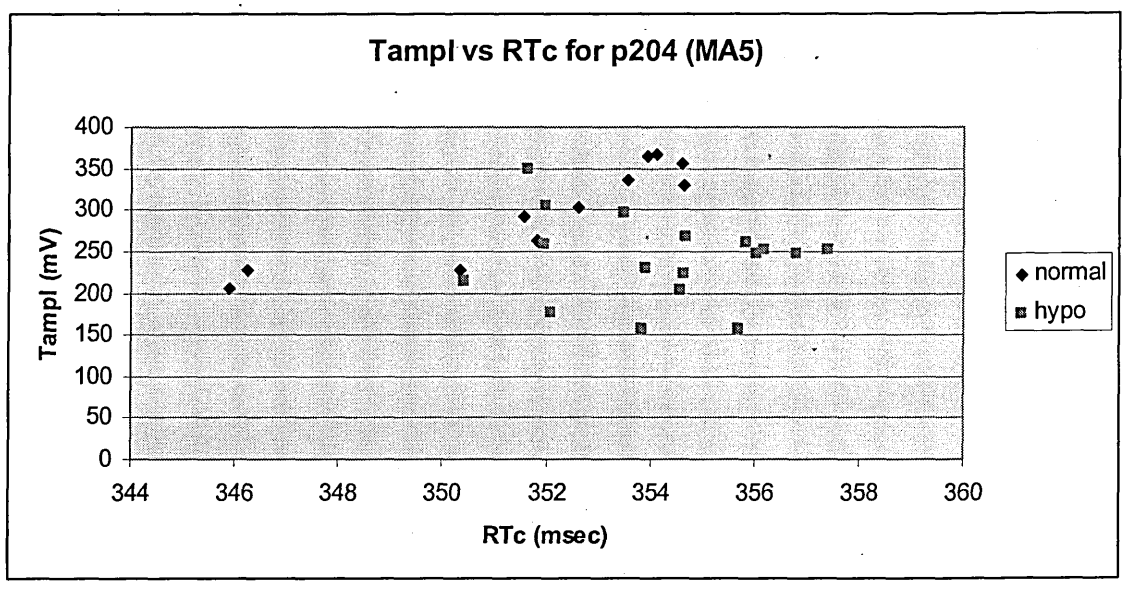


Figure 5.19: Tampl vs RTc scatter diagram for MA5 filtered ECG features (patient 204)

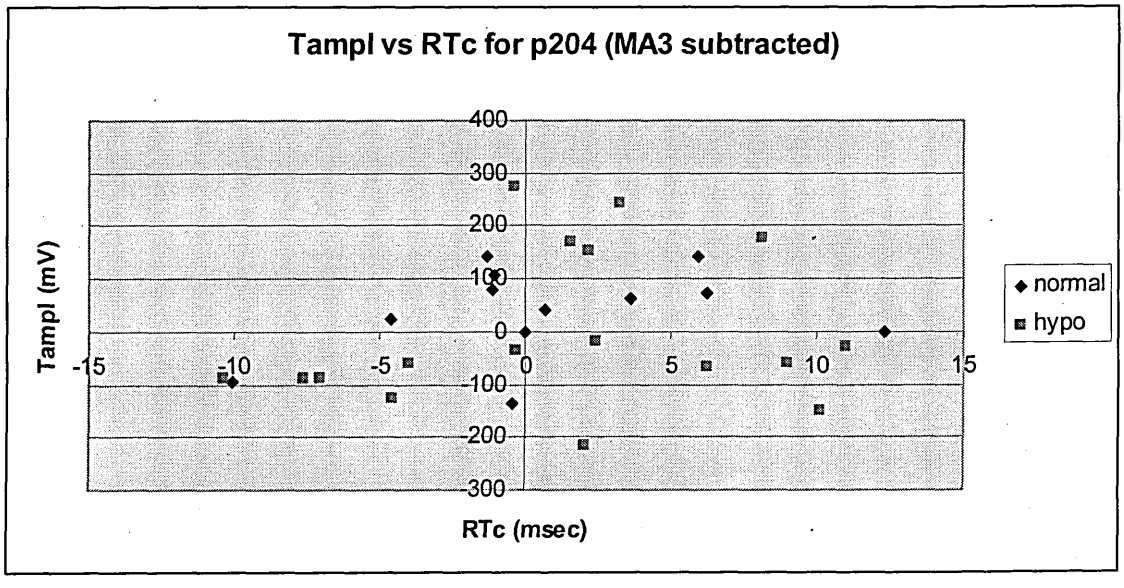


Figure 5.20 Tampl vs RTc scatter diagram for normalised ECG features by subtracting the MA3 signal (patient 204)

It is easily observed by comparing the scatter diagrams that filtering the data with a length-3 MA filter gave the most clear-cut sub-classes on the 2D classification surface. With the exception of two data-points, the two classes can be distinguished by a linear decision boundary. This leads to only two misclassifications (missed-hypo events) that lie in the euglycaemic class. Using a MA-length5 filter also gives useful results though they are inferior to when using a window of length 3. The two classes are closer together and more difficult to distinguish.

Using a dynamic threshold gives a classification task that is significantly more difficult to solve. The main euglycaemic sub-class occurs within the hypoglycaemic class. Using

the raw feature values also leads to a situation that is difficult to classify. To summarise the classification results presented earlier, a bar chart is given in Figure 5.21.

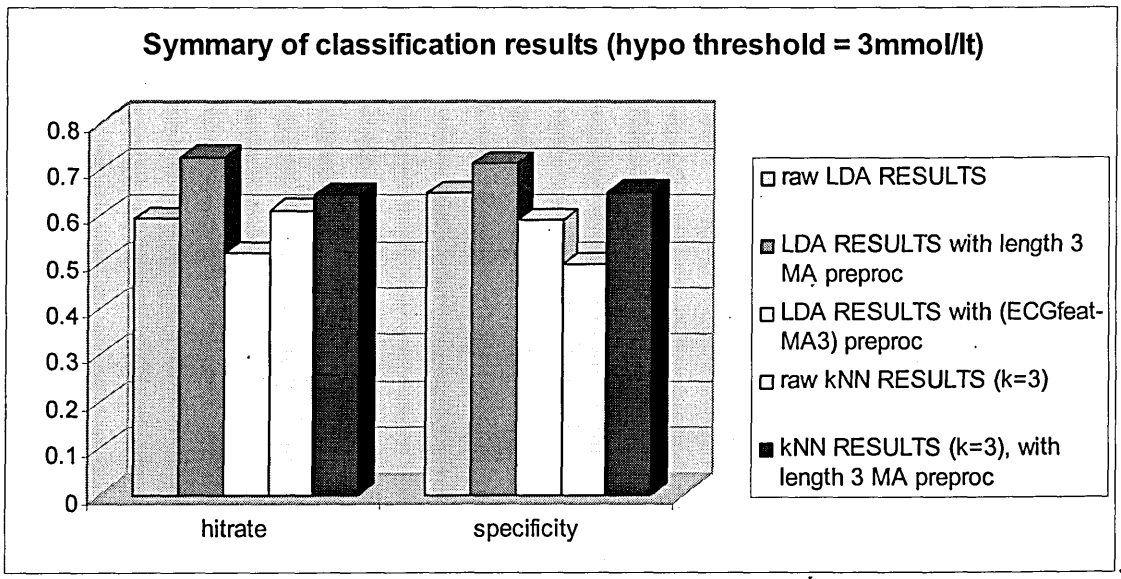


Figure 5.21: Summary of preprocessing-classification approaches

The best results were produced using the combination of MA (length 3) preprocessing and LDA. The second best approach was again using MA (length 3) preprocessing but followed by a k-Nearest Neighbour (k=3) classifier. The kNN classifier gave better results than the LDA when raw features were fed and also when the dynamic threshold was used.

Using a dynamic threshold defined by the current feature value minus the current MA value proved to be worse than using the raw ECG features! In principle, the dynamic threshold should be superior because it would lead to adaptivity in response to cardiac changes during the night. The results demonstrated that this was not the case in practice. However the idea of a dynamic threshold was successfully used as part of a monitoring system for hypoglycaemia detection, which will be presented in Chapter 6. The advantage of the system presented in Chapter 6 was that it was also using temporal information in combination with dynamic thresholds. This combination was effective in improving performance.

5.3 Discussion

In this chapter, pattern classification of time-averaged ECG signals was carried out. This is a static process since no time stamps were used for the feature vectors fed to the classifier. In the studies incorporating MA preprocessing, although we did not feed a

variable conveying temporal information to the system, we were effectively including some information from the temporal dimension since the filtering process effectively provided information about the samples previous to the one classified at each classification epoch. The same was done in the studies where the derivative of an ECG feature was used. The first derivative value was accompanying each ECG feature fed and was effectively providing information about the previous ECG sample. The derivative value used was providing information on whether the previous sample was at a higher/lower amplitude and by how much. The best results from this chapter reached a classification accuracy on unseen data of approximately 72%. This figure is promising and is improved in Chapter 6. Further improvements yielding robust detection of hypoglycaemia could lead to an online monitoring system for the bedside.

5.4 Conclusion

This chapter presented a number of classification studies for distinguishing between ECG traces corresponding to the conditions of euglycaemia (normal glucose levels) and hypoglycaemia. Multi-layer perceptrons were used along with statistical classifiers (LDA and kNN). A number of approaches towards tackling the problem were presented. Classifiers trained on global datasets versus classifiers tailored to the dynamics of specific patients were investigated. It could safely be concluded that producing a customised system for the needs of a given patient would be the best approach towards tackling the problem; the reason being the great inter-patient variability. Even if a globally trained classifier could solve the problem, it would be expected that a customised classifier would introduce a further improvement in performance.

A brief study on ECG trace representation by AR coefficients instead of the use of ECG features gave promising results and could be investigated further in the future. This was not done due to time constraints and also because it was thought that the quite recent approach of Action Potential modelling [Wohlfart 1987, Vila 2000] of ECG signals would be more suited to this work compared to AR modelling. Hence it would be wiser to perform further work on ECG trace representation by Action Potentials instead of doing so by AR coefficients. This is because the process of modelling by Action Potentials is more plausible biologically.

Finally, the use of extra preprocessing techniques that would aid in solving ambiguities in magnitude of ECG features corresponding to both clinical conditions, were

investigated. The best approach proved to be using the Moving Average signals instead of the raw ECG features. The approach of using a dynamic threshold to assess the normality of an ECG feature failed and gave results inferior to those when using the raw features.

A Knowledge-Based Monitoring System for Hypoglycaemia Detection

6.0 Introduction

This chapter focuses on the design of a monitoring and alarm system for detection of the onset of spontaneous nocturnal hypoglycaemia. Firstly it discusses the differences from previous classification approaches (MLP, LDA and kNN presented in Chapter 5) and then it presents the ECG features used and the way monitoring was carried out. The adaptivity of the system to ECG feature changes as time elapses is discussed and the rule-base is presented. The system was realised as an Expert System (using Crisp Logic) and also as a Fuzzy Inference System (using Fuzzy Logic). Both systems are presented, along with monitoring results, and their differences are highlighted.

6.1 Overview of monitoring system

The approach for detection of spontaneous nocturnal hypoglycaemia presented in this chapter differs from the classification approaches presented in Chapter 5, in that it simulates a patient monitoring scenario. The data available are of course offline, but the approach is very similar to that where a patient is monitored online on his bedside. The data are treated as if they were online data. The main difference to previous approaches is that temporal information is incorporated to the system. In previous studies, static pattern classification was carried out i.e. no temporal information was included. What was classified consisted of feature vectors corresponding to the two conditions (euglycaemia and hypoglycaemia) and the classifier was trying to recognise patterns and classify the feature vectors correctly. In this study each feature vector was fed to the system while the system was also using information from previous feature vectors in order to detect changes in the behaviour of the cardiac signal. In other words, the temporal dimension was also included and taken in account in the inference process.

This system was based on incorporation of basic knowledge by human experts. The Knowledge Base was constructed from vague guidelines by our medical collaborators besides our observations on the dataset. A number of basic rules were used to achieve monitoring of the patients. The initial Knowledge-Based System (KBS) produced was based on traditional Crisp Set Logic and will be referred to as "Expert System" (ES) since this term has been used in the field of AI to describe KBS based on Crisp Set Logic. An alternative version of the system was also produced that was using Fuzzy Logic. The Fuzzy Inference System (FIS) produced was using exactly the same rule-base as the ES. The FIS possessed the additional advantage that it could produce a degree of certainty to support the decision inferred. This degree of certainty was describing the strength of an alarm if one was raised and, in the case of no alarm raised, the extent to which the alarm sounding threshold was approached. Such information is useful for patients and clinicians using diagnostic systems.

The monitoring system is outlined in Figure 6.1. It has four inputs and three outputs. More information about the architecture will be given in Section 6.7.

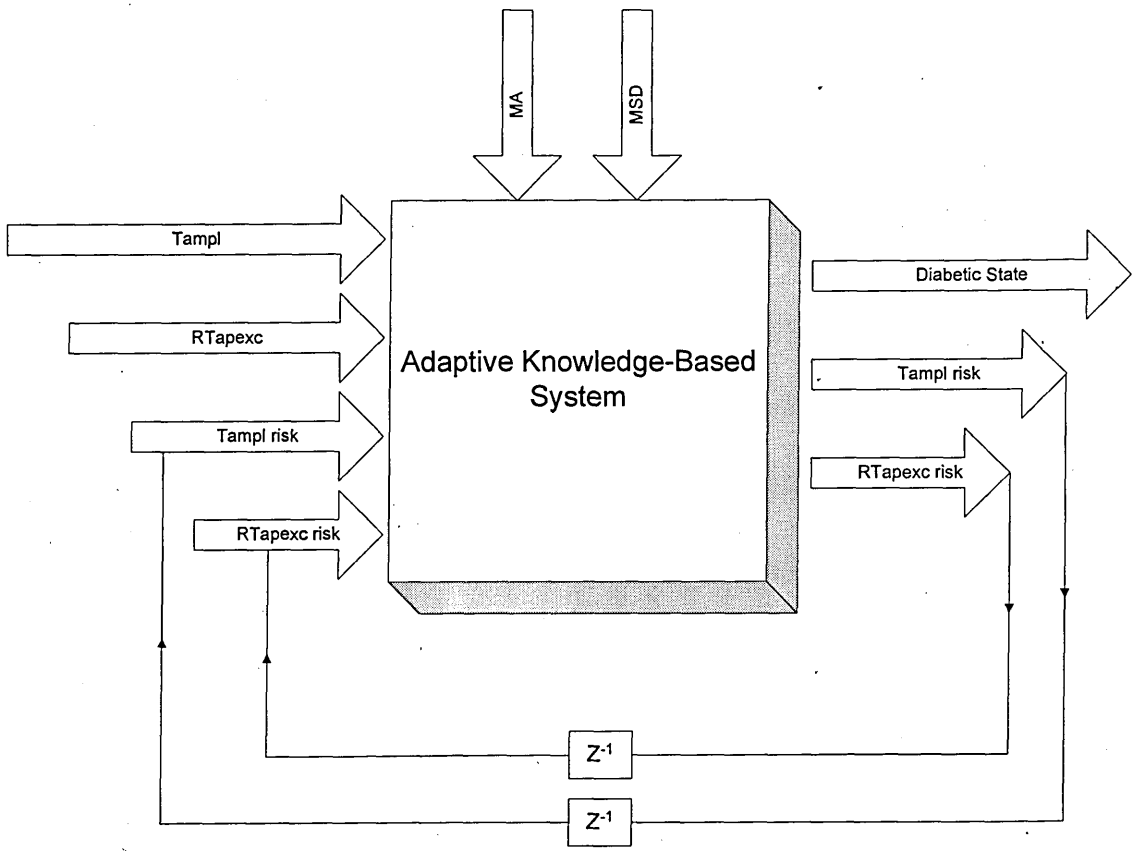


Figure 6.1: Illustration of the KBS

6.2 Dataset

The data used in this approach originated from the dataset on spontaneous hypoglycaemia presented in Section 3.3.1. The fact that the hypoglycaemic events are nocturnal and spontaneous makes this dataset suitable for a realistic monitoring study. If an online monitoring system was to be produced to monitor abnormal glucose levels by inspection of the ECG then it would be aimed at detecting spontaneous events only while the subject would be asleep. Such a system is not addressing the monitoring of patients when they are awake. The ECG traces used were signal-averaged (SAECG) as opposed to beat-to-beat in line with classification studies presented in chapter 5, also using SAECGs.

6.3 ECG features

Two ECG features were fed to the monitoring system. One of the features was the T wave amplitude (Tampl) and the other feature was a time interval feature, either RTapexc or RTc, describing the VR duration. RTapex is the time interval from the R peak up to the T peak (T apex) and RT the interval from R peak to T-end. Both these features were corrected for heart rate using Bazett's formula (eq^{n} 2.1) to produce the RTapexc and RTc features where the suffix "c" stands for "corrected". Either the RTapexc or the RTc feature was fed to the system and comparisons of the effect of each feature in system performance were carried out.

The reason for choosing the RTapexc feature, as an alternative to the RTc feature, is the lack of a robust and well-established T-end detection algorithm. As it has been stressed earlier, despite the large amount of research effort that has been devoted to the design of a robust algorithm, the gold standard in T-end detection is still the manual annotation by a human expert. The RTapex feature has the advantage of not using the T-end annotation in its definition. The drawback of this feature is that it will not describe late VR phenomena that may be reflected on the T-downslope. A prolongation in the QT interval will be manifested to a lesser extent on the RTapex feature compared to the RT. Although the T downslope is of great interest and the RTapex is incapable of representing the behaviour of this downslope, it was chosen because of issues of robustness. Weaknesses in T-end annotation algorithms may cause variations in QT and RT features that have no clinical significance and could be confusing in a classification or monitoring situation.

This led us to investigate the alternative of choosing a less informative but more robust feature. Besides testing the monitoring system when either using the RTc or the RTapexc features, the uncorrected versions (i.e. no heart-rate correction) of these features were also tested. The RTc and RTapexc feature profiles for patient 205-night1 are presented in Figure 6.2 to allow comparisons of the two definitions of describing VR duration. The top figure presents the raw values of RTc (blue) and RTapexc (black). The middle graph presents the values for the two features after the baseline (1st value of the night) has been removed. This was done to aid comparisons of the variations of the two features. The bottom graph contains the glucose variable for that night, with the dashed horizontal line denoting the 3 mmol/lt hypoglycaemic threshold. It is apparent that this patient did not go into hypoglycaemia.

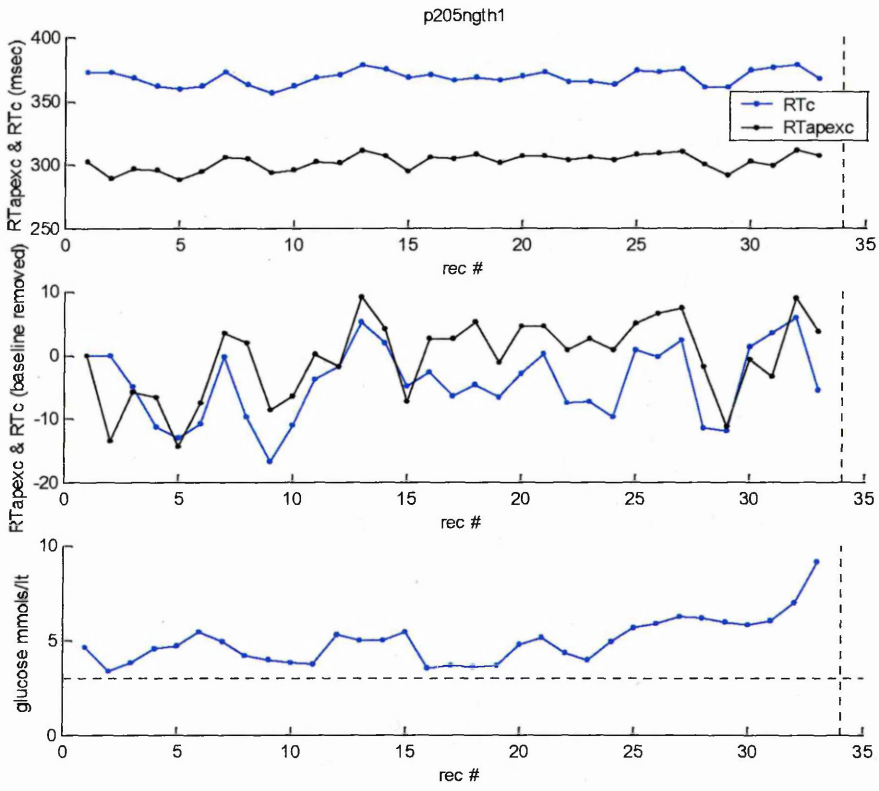


Figure 6.2: RTc and RTapexc for 205-night1

The other feature chosen besides RTapexc was the T amplitude, as mentioned earlier. This choice was because depletion in plasma potassium due to counter-regulatory responses occurring during hypoglycaemia affects in many cases the T wave by a drop in its amplitude, as discussed in Section 1.1.5. T wave flattening is apparent in many studies related to hypoglycaemia and is very frequently encountered in our datasets.

The number of input features used was kept small in order to achieve low complexity of the rule-base and the prototype system in general. The reason for choosing two features was to be in line with the main clinical hypothesis upon which this thesis is based. According to the hypothesis, only features quantifying the QT prolongation and T wave flattening would be used. The features used were Tampl and a time interval feature describing VR duration (either RTapexc or RTc). The above two features were chosen because of their clinical significance. They both reflect the counter-regulatory responses that can be encountered under hypoglycaemia. RTapexc (and also RTc) is related to the release of adrenaline and T amplitude to the drop in potassium. Therefore choosing to avoid features based on the T-end annotation leads to the use of the above two features. In theory, RTapexc (or RTc) and T amplitude should be sufficient in quantifying the flattening and prolongation of the T wave occurring under hypoglycaemia, which constitutes the main assumption of this work. ST segment changes and U wave morphology changes are extra events that may occur under hypoglycaemia and could be incorporated in the system as part of future work.

6.4 Monitoring

The pre-requisite for patients used in this monitoring study was that they should have had healthy glucose levels at the time they went to bed. Both hypoglycaemic and euglycaemic nights were used. Successful monitoring on an euglycaemic night would mean that an alarm should not be raised. Successful monitoring on a hypoglycaemic night would mean the detection of the onset of hypoglycaemia as closely as possible to the time it occurred. Once an alarm had been raised the monitoring would stop since the patient would have woken up. In a real monitoring situation the patient would treat himself (e.g. with carbohydrates) to restore the glucose levels to normal and go back to bed. The system would be reset and start monitoring again. A post-alarm continuation of monitoring was not considered in this study in order to simplify the problem and also because the duration of each night-recording (8 hours) was not enough to allow resuming the monitoring process. Also, since the data used was not online data, if an alarm was raised this would not correspond to the patient waking up. This means that there would be no restoration of glucose to normal levels after the alarm. Consequently, the cardiac function would not have been fully restored. This makes it an unrealistic situation to be used for monitoring.

6.5 Moving Window Applied on ECG Features to Achieve Adaptivity

During the monitoring process, abnormal changes in feature value as time elapsed were detected by comparison to an adaptive threshold. This threshold was based on a moving average value. A moving window was used containing a few samples prior to the current time instant. In some versions of the system, this was combined with a moving value of the standard deviation, calculated from the same moving window, to define an accepted range (Healthy Band) of feature values.

The equation of the Moving Average filter used is given below in its generic form for n^{th} order:

$$\bar{x}(k)=\begin{cases}\dfrac{\sum_{i=1}^{n}x(k-i)}{n},\ for\ n<k \quad n,k\in\aleph^{+} & eq^{n}\ 6.1a\\ \dfrac{\sum_{i=1}^{k-1}x(k-i)}{k-1},\ for\ n\geq k \quad n\in\aleph^{+},k\in\aleph^{+}\wedge k>1 & eq^{n}\ 6.1b\\ 0,\ for\ k=1 & eq^{n}\ 6.1c\end{cases}$$

x(k) is the raw feature value at sample *k*, and $\bar{x}(k)$ is the MA filtered version.

The equation for the calculation of the Moving value of the Standard Deviation (MSD) is given below in its generic form for n^{th} order:

$$y(k)=\begin{cases}\dfrac{\sum_{i=1}^{n}(x(k-i)-\bar{x}(k))^{2}}{n-1},\ for\ n<k \quad n,k\in\aleph^{+} & eq^{n}\ 6.2a\\ \dfrac{\sum_{i=1}^{k-1}(x(k-i)-\bar{x}(k))^{2}}{(k-1)-1},\ for\ n\geq k \quad n\in\aleph^{+},k\in\aleph^{+}\wedge k>2 & eq^{n}\ 6.2b\\ 0,\ for\ k\in\{1,2\} & eq^{n}\ 6.2c\end{cases}$$

x(k) is the raw feature value at sample *k* and $\bar{x}(k)$ is the MA value at sample *k* calculated from the moving window that spans up to *x*(*k*-1). *y(k)* is the standard deviation of the feature values that lie in the moving window. This standard deviation is not the deviation from a static mean but the deviation from the moving average.

In equations 6.1 and 6.2 presented above, the moving window does not include the current sample, i.e. the right-hand-side limit of the window is the sample previous to the current one. Inclusion of the current sample in the moving window was also tested. When the current sample is included in the window, the equations take the form:

$$\bar{x}(k)=\begin{cases}\dfrac{\sum_{i=1}^{n}x(k-i+1)}{n}, \text{ for } n<k \quad n,k\in\aleph^{+} & eq^{n}\ 6.3a\\ \dfrac{\sum_{i=1}^{k-1}x(k-i+1)}{k-1}, \text{ for } n\geq k \quad n\in\aleph^{+},k\in\aleph^{+}\wedge k>1 & eq^{n}\ 6.3b\\ 0, \text{ for } k=1 & eq^{n}\ 6.3c\end{cases}$$

$$y(k)=\begin{cases}\dfrac{\sum_{i=1}^{n}(x(k-i+1)-\bar{x}(k))^{2}}{n-1}, \text{ for } n<k \quad n,k\in\aleph^{+} & eq^{n}\ 6.4a\\ \dfrac{\sum_{i=1}^{k-1}(x(k-i+1)-\bar{x}(k))^{2}}{(k-1)-1}, \text{ for } n\geq k \quad n\in\aleph^{+},k\in\aleph^{+}\wedge k>2 & eq^{n}\ 6.4b\\ 0, \text{ for } k\in\{1,2\} & eq^{n}\ 6.4c\end{cases}$$

The only difference in equations 6.3 and 6.4 is the "+1" component in the numerator, which causes the current sample to be included in the window. The effect of including the current sample is discussed in Section 6.13.

For the T amplitude feature, a significant event related to the onset of hypoglycaemia would be an abnormal drop in amplitude while for the RTapexc feature we would look for an abnormal increase in feature value (corresponding to QT prolongation). This means that opposing changes in the two features are significant. A drop in T amplitude below the moving average value would be a significant event (risk factor)[31] and similarly for an increase in RTapex value. For an alarm to be raised, two successive in time significant events in both features are necessary. The use of two successive significant events was needed to avoid changes in feature value that would correspond to artefacts or other brief cardiac events unrelated to hypoglycaemia.

[31] An abnormal feature change will be referred to a "potential risk", "risk factor", or "significant event" interchangeably throughout the chapter and without further clarification.

Looking for significant events independently in each of the two features would lead to many false alarms. There were cases where abnormal changes in only one feature would be observed that were not related to hypoglycaemia. There are even cases where abnormal changes even in both features were unrelated to hypoglycaemia. Such an example is the first night of patient 205 presented in Figure 6.5 (Section 6.10). This patient exhibits both T flattening and RT prolongation which are unrelated to hypoglycaemia and this causes a false alarm. Using just one feature would mean that many more cases of feature changes unrelated to hypoglycaemia would lead to false alarms. To ensure that the changes were genuine, they were expected to happen in both features.

An alarm accurately raised at record 17 for patient 204, is illustrated in Figure 6.3. The graphs show the actual feature values, together with the moving average values (dotted lines). The vertical dashed lines mark the record where the alarm was raised.

The use of the moving average criterion was necessary for detection of significant changes in feature value. Because of inter-patient variability, an absolute threshold that if exceeded would raise an alarm, could not be used. Also because of intra-patient variability during the night, the baseline at the start of the night or any other fixed value could neither be used as a threshold that when exceeded corresponds to a significant event. An adaptive threshold, that would be changed as night progressed, had to be used to detect significant events. This was realised in practice as early approaches of using a non-adaptive KBS did not achieve satisfactory performance. The term "non-adaptive KBS" means that the thresholds used are static and hence the system uses a fixed definition of euglycaemia and hypoglycaemia. On the other hand, adaptive thresholds mean that the definitions of euglycaemia and hypoglycaemia change as time elapses.

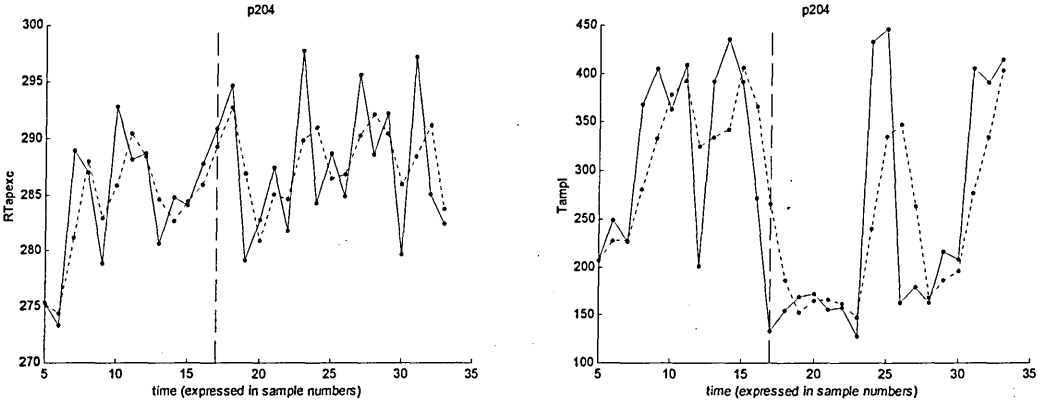


Figure 6.3: Tampl (LHS) and RTapexc (RHS) graphs (solid lines) including MA values (dotted lines) for subject 204

A reason justifying why adaptive thresholds were necessary to define the abnormal events is the apparent prolongation in QT, and hence RTapex, during sleep. Molnar et al [Molnar 1996] have studied the diurnal pattern of QTc and observed lengthening of QTc during sleep. Moreover an increase in T amplitude for a few samples at the beginning of sleep was observed in some patients of our dataset. This can be clearly seen for patient 204 in Figure 6.3 and patient 203 in Figure 6.4 presented in Section 6.10. If such an initial increase in amplitude occurs and later there is a drop due to hypoglycaemia, the drop may cause the feature to have similar values as those at the start of the night. So if the baseline at the start of the night was used as a threshold this would mean that the significant event due to a drop in feature value, could not be detected. However, an adaptive threshold will closely follow the dynamics of the T amplitude and RTapexc variables. It will mask the acceptable dynamic changes and help identify the abnormal ones.

6.6 Tuning of Window Size

The length of the window used for calculation of the moving average was varied and an optimal value was chosen. It was varied from a length of 1 sample up to 5 samples for both features. This was done by an exhaustive search of all combinations while testing the performance of the alarm system. The optimal values were different for a few patients although groups of patients having the same optimal values could be identified. The differences in optimal window sizes among patients should not be an obstacle for a custom alarm system that would be trained on the patient to be monitored prior to the monitoring period. An alarm system tailored to a specific patient can be permitted to learn and adjust for a period of time prior to the start of monitoring. The width of the Healthy Band (HB) was also tuned for optimal performance as will be discussed in Section 6.11.

6.7 Rule-Base

As shown in Figure 6.1, the system has 4 inputs and 3 outputs. Two of the inputs are the two ECG feature values (Tampl and RTapexc) at the current sample and one of the outputs (Diabetic State) represents the alarm state. The rest of the inputs and outputs refer to the significant events related to previous samples. The potential-risk outputs are fed back to the system as inputs and are used in the next monitoring epoch.

The two principal rules used for monitoring are presented below:

1. **IF** (Tampl is flattened) **and** (Tampl_prev is flattened) **and** (RTapexc is prolonged) **and** (RTapexc_prev is prolonged) **THEN** (DiabeticState is hypo)
2. **IF** (Tampl is normal) **or** (Tampl_prev is normal) **or** (RTapexc is normal) **or** (RTapexc_prev is normal) **THEN** (DiabeticState is eugly)

The suffix "_prev" stands for previous sample before the current one. "flattened" is defined as: Tampl < Tampl_MA, where Tampl_MA is the moving average of Tampl, based on a window size selected for optimal performance. Similarly, "prolonged" is defined as: RTapexc > RTapexc_MA, where RTapexc_MA is the moving average of RTapexc, based on a window size tuned for optimal performance. If the Diabetic State is "hypo" then an alarm is raised. In all other cases the Diabetic State is "eugly" i.e. the night is euglycaemic.

If the combination of moving average and moving value of standard deviation is used, as will be discussed in Section 6.11 onwards, then "flattened" is defined as: Tampl < (Tampl_MA - Tampl_MSD) where Tampl_MSD is the moving value of the standard deviation of Tampl and "prolonged" is defined as: RTapexc > RTapexc_MA + RTapexc_MSD, where RTapexc_MSD is the moving SD of RTapexc. It is emphasised that the system is adaptive to changes of ECG features only. The rest of the inputs and also the outputs of the system are static. The window used to calculate the MA and MSD values is kept fixed during monitoring. For the Tampl and RTapexc inputs a number of past values, dictated by the window size, are stored in internal registers for calculation of the MA and MSD values of each epoch. Besides the inputs to the KBS, the MA and MSD values are also used in the inference process and are illustrated with arrows, at the top of the block diagram, in Figure 6.1.

The two principal rules were provided for illustrating the concept behind the monitoring approach. They were presented for purposes of better readability since they were compact. The actual rule-base used in the system consists of eight rules since not all the eventualities are covered in the two principal rules. The latter present only how the Diabetic State output is calculated, but not the other two outputs. The expanded rule-

base fully complies with the compact rules presented above and is the one that corresponds to the system presented in Figure 6.1.

The exact rule-base, representing the knowledge of the system, is presented below:

1. **IF** (Tampl is flattened) **and** (Tampl_risk is high) **and** (RTapexc is prolonged) **and** (RTapexc_risk is high) **THEN** (DiabeticState is hypo)
2. **IF** (Tampl is normal) **or** (Tampl_risk is low) **or** (RTapexc is normal) **or** (RTapexc_risk is low) **THEN** (DiabeticState is eugly)
3. **IF** (Tampl_risk is low) **and** (Tampl is normal) **THEN** (Tampl_risk is low)
4. **IF** (Tampl_risk is low) **and** (Tampl is flattened) **THEN** (Tampl_risk is high)
5. **IF** (Tampl_risk is high) **and** (Tampl is normal) **THEN** (Tampl_risk is low)
6. **IF** (RTapexc_risk is low) **and** (RTapexc is normal) **THEN** (RTapexc_risk is low)
7. **IF** (RTapexc_risk is low) **and** (RTapexc is prolonged) **THEN** (RTapexc_risk is high)
8. **IF** (RTapexc_risk is high) **and** (RTapexc is normal) **THEN** (RTapexc_risk is low)

Rules 1-2 perform monitoring of the patient's ECG. Rules 3-8 are used for evaluating potential risk factors from previous feature samples. They increase the potential risk of a feature if the current feature value appears abnormal while the previous one appears to be normal or decrease the potential risk when the opposite happens. The weighting for all rules was set to 1 for both the ES and the FIS.

A look-up table is provided (Table 6.1) to demonstrate how the system works. It presents the input-output mappings for a simplified system using only two inputs, being the ECG features, and only one output being the Diabetic State. Since the actual system has four inputs it cannot be illustrated as a two-dimensional look-up table. The simplified system does not use information about the potential risks from previous samples and hence the number of inputs is reduced to two.

Table 6.1 Look-up table for two-input KBS (excluding potential-risk inputs)

RTapexc / Tampl	normal	prolonged
normal	euglycaemia	euglycaemia
flattened	euglycaemia	hypoglycaemia

It must be noted that abnormal feature values to the other extreme (i.e. elevated T waves or shortened RTapexc intervals) are treated as normal since they are not known to be related to hypoglycaemia.

6.8 Hypoglycaemic Threshold Used and Quantitative Evaluation of KBS Performance

In order to assess the performance of the monitoring system, the threshold defining the onset of hypoglycaemia needed to be defined. This was discussed in Section 4.2. In contrast to the classification studies, in this chapter, the onset of hypoglycaemia was defined using two different hypoglycaemic thresholds, one at 3 mmol/lt and another one at 2.5 mmol/lt. The onset according to both thresholds is tabulated in the results sections. This is because there is no single answer for which is the optimal hypoglycaemic threshold to be used. Various research studies have considered different thresholds as discussed in Section 4.2. Some patients could be symptomatic at 3 mmol/lt, while others would have to drop lower for changes on the ECG to be manifested. (There could even be cases where there is no manifestation on the ECG which is when our main research assumption would not hold leading to inability of hypo detection.) For this reason both thresholds are presented so each case can be judged individually. The abbreviation "eugly" in the 3rd column of the results-tables means that the night was euglycaemic. On the last column labelled "perf" a summary of the performance of the system is given. The following abbreviations are used to describe the performance which in some cases is also given descriptively:

- CA: Correct Alarm, i.e. True Positive (TP), used to denote correctly raised alarms
- CE: Correctly monitored Euglycaemic night i.e. True Negative (TN)
- FA: False Alarm for a night i.e. False Positive (FP)
- MH: Missed Hypo for hypoglycaemic night where no alarm was raised (i.e. False Negative)

For the version of the system incorporating the MSD criterion the result tables contain a few extra columns as will be seen in Table 6.3.

Two sets of results for each version of the system are presented. A global KBS system was produced which is tuned in such a way so as to perform optimal monitoring for all patients. In order to improve performance, the KBS system was further tuned to achieve customisation per patient. Such an approach is preferable since it achieves higher performance. Moreover it is still realistic for monitoring because, when a real-life monitoring system is produced it is feasible for it to be tuned on the patient to be monitored, for a period of time, before the actual monitoring starts. The rule-base, representing the knowledge of the system, was identical in both cases (global and custom KBS). The only parameters that were tuned were the window sizes for the two features and also the width of the Healthy Band as it will be seen in the version of the system incorporating moving values of the standard deviation.

Analysis of performance on the hypoglycaemic nights is not as straightforward as for the euglycaemic nights. For the latter, if no alarm is raised this means correct monitoring and when the contrary happens it means a false alarm. But when assessing the performance on hypoglycaemic nights the outcome is not binary (correct-alarm or missed-alarm). There are cases where an alarm is raised with a small deviation, in time, from the onset of hypoglycaemia. The cost of such a deviation must be assessed in each case. Such events are also presented descriptively, instead of just using percentages, in the results tables of this chapter.

The analysis of how early or late an alarm was raised, was carried out in terms of sample numbers and not in terms of the actual time duration. In the dataset used, ECG data were recorded every 15 minutes during the night so an alarm two samples early or late corresponds to half an hour in time which is a significant duration for a monitoring system, making it look inaccurate. However this is only due to the fact that the data were recorded every 15 minutes. This is the maximum temporal resolution of the dataset. So if an alarm is raised just one sample late this means 15 minutes late which is not a limitation of the alarm system but a limitation of the dataset. Such an alarm is actually the second best result after a "spot on" detection. In a different situation with more frequent sampling, one sample early or late would mean a shorter period of time. In order to be fair in the assessment of the alarm system, the analysis is considering inaccuracies in terms of sample numbers and not in terms of time duration. This means that we are not assessing limitations due to the temporal resolution of the dataset but only assessing the limitations of the monitoring system.

6.9 Approved patients from the dataset

This section discusses the choice of the patients from the dataset that were included in the study and gives justification for patients excluded. It also presents the assumptions formulated for certain patients. It must also be clarified here that the serial numbers for the ECG traces were 1-33 for the first night of a patient and 34-66 for the second night.

The datafiles not included in the study are listed below:

- 208nght1 was not included because the night started in hypoglycaemia.
- 216nght1 was not included because hypoglycaemia started very early (glucose dropped below 2.5 mmol/lt at record 4). This made it very difficult for the KBS to perform monitoring since there was not enough healthy data for the system to start defining its normal and abnormal thresholds.
- 216nght2 was not included because hypoglycaemia also started early (glucose dropped below 2.5 mmol/lt at record 39 i.e. the sixth record in the night). Moreover the glucose was fluctuating in and out of hypo throughout the night. The patient went in and out of hypo 5 times.
- Patient 220 (both nights) was not included because the ECG exhibited many inverted T waves. Because the temporal position of a normal T peak varies from that of the peak of an inverted T wave, the transition from a normal T wave to an inverted and vice versa meant fluctuations in the time-interval feature describing the QT which can be confusing for the monitoring system. Once more sophisticated ECG features could be incorporated in the system, this patient could be included in future studies.
- Both nights of patient 223 started as hypoglycaemic so this patient was excluded.
- Patient 225 (both nights) was rejected because of extremely bad quality of the ECG recording.
- Patient 226 (who only contributed a single night) could not be used because of the existence of many ambiguous components in the ECG. The patient exhibited biphasic T waves, non-standard ST segments and, in certain records, almost inexistent T waves.
- 229nght1 was not used because the onset of hypo happened very early (record 2), with the patient recovering for a while and then going hypo again.
- 229nght2 started as hypo so it was also rejected.
- Both nights of patients 234 were rejected because of the bad quality of the ECG.

Apart from the excluded patients an assumption had to be made for one of the approved patients:

For patient 201A (night 2) the glucose dropped below 3 mmol/lt (it did not go below 2.5 mmol/lt) at the last record of the night (record 66). This night was classed as euglycaemic and this late event was ignored. Since there was no data past this event it is unknown what happened after that and also the glucose did not drop below 2.5 mmol/lt which would constitute a stronger candidate to be classed as hypoglycaemic event to take into account. The biomedical scientist performing data acquisition and ECG annotation in the Diabetic Clinic of the RHH[32] had also classed this night as euglycaemic. In clinical studies a night would normally be defined as hypoglycaemic only if the glucose was in the hypoglycaemic region for a significant period of time (i.e. 30 minutes).

6.10 Monitoring Results using MA

The performance of the alarm system when only the MA criterion was used is tabulated in Table 6.2. The table contains the following information: the patient number and corresponding night, the record at which the onset of hypoglycaemia occurred, the record at which the alarm was raised, the optimal MA window sizes for the two features and the assessment of performance ("perf"). "TamplWS" stands for Tampl-WindowSize and similarly for RTapexcWS. When the hypoglycaemic-onset entry of the table ("hypo-onset@rec") is zero the night was euglycaemic. Similarly when the alarm output ("alarm@rec" field) is zero, no alarm was raised by the system. It must be stressed here that the current feature value at each time instant was included in the calculation of MA. Alternatively, the adaptive statistics can be calculated using a window spanning up to the previous sample, as will be seen in later sections.

The table contains the results from the customised monitoring systems. Global results are not presented for this early system. Both a global and a set of customised monitoring systems will be given in Section 6.11 presenting the system that incorporates the MSD criterion. The total nights monitored were 32, contributed by 19 patients. An alarm was classed as acceptable if it satisfied any of the two hypoglycaemic thresholds considered (2.5 mmol/lt or 3 mmol/lt). Out of the 32 nights used, 12 nights were monitored accurately out of which, 3 were hypoglycaemic, with the alarm being raised at the

[32] Cath Davies

correct sample number by the KBS, and 9 were euglycaemic where no alarm was raised by the monitoring system. This means that the accuracy, sensitivity and specificity of the KBS were 37.5%, 33.3% and 39.13% respectively. There were also 14 nights where a false alarm occurred and 6 nights where hypoglycaemic events were not detected.

Table 6.2: Performance of alarm system based on RTapexc and T amplitude featrures

#	patient	hypo onset@rec	alarm@rec	TamplWS	RTapexcWS	perf
1	201Anght1	eugly	0	3	2	CE
2	201Anght2	66(<3)	0	3	2	CE
3	202nght1(202A)	23(<3), 24(<2.5)	26	4	3	CAwithin2
4	202nght2(202)	41 (<2.5)	41	3	2	CA
5	203nght1	11 (<2.5)	20	3	2	CAwithin9
6	203nght2	eugly	0	3	2	CE
7	204 (nght1)	17(<3), 18(<2.5)	17	3	2	CA
8	205nght1	eugly	26	3	2	FA
9	205nght2	eugly	48	3	2	FA
10	207nght1	eugly	0	2	2	CE
11	207nght2	eugly	43	2	2	FA
12	208nght2	eugly	64	3	2	FA
13	209nght1	eugly	20	4	4	FA
14	209nght2	50(<3), 55(<2.5)	50	4	4	CA
15	210 (single night)	eugly	0	3	2	CE
16	212nght1	eugly	12	3	2	FA
17	212nght2	58(<3), 59(<2.5)	45	5	2	CAwithin13
18	215nght1	eugly				FA
19	215nght2	eugly				FA
20	218nght1	eugly	0	4	2	CE
21	218nght2	eugly	0	4	2	CE
22	221 (single night)	eugly	8	3	2	FA
23	222 (single night)	eugly	0	3	2	CE
24	p227nght1	21(<3), 22(<2.5)	20	2	4	CAwithin1
25	p227nght2	39 (<2.5mmol)	44	2	3	CAwithin5
26	p230nght1	eugly	14	3	2	FA
27	231nght1	eugly	14	3	2	FA
28	231nght2	eugly	45	3	2	FA
29	232nght1	eugly	14	3	2	FA
30	232nght2	eugly	0	3	3	CE
31	p244nght1	18(<2.5) 23(<2.5)	22	2	5	CAwithin1
32	p244nght2	eugly	41	2	5	FA

If we consider a deviation from the hypoglycaemic onset of up to and including 2 samples as acceptable then there are 3 more hypoglycaemic nights monitored correctly giving a total of 15 nights where the alarm system performance was acceptable with the sensitivity reaching 66.6%. Such nights were: 202A (alarm 2 samples late), 227nght1

(alarm 1 sample early) and 244nght1 (alarm 4 samples after a very brief hypoglycaemic event and 1 sample before the main hypoglycaemic onset).

A discussion of a few interesting cases of patients will be presented here. For patient 203 (night 1) which is a hypoglycaemic night, the alarm is raised late by 9 samples (hypo occurs at record 11). This happens because there is no manifestation of an abnormal event on the ECG at the onset of hypoglycaemia. Regarding the T amplitude feature, the opposite of what is expected happens. That is an increase in amplitude of the T wave. It can be seen in Figure 6.4 presenting the actual (solid) and Moving Average (dotted) profiles for the Tampl feature that there are 4 successive increases in T amplitude from record 9 up to record 13. This means that no flattening is manifested in the proximity of record 11 which is when hypoglycaemia occurs. The vertical dashed line on the figure marks the record where the alarm was raised.

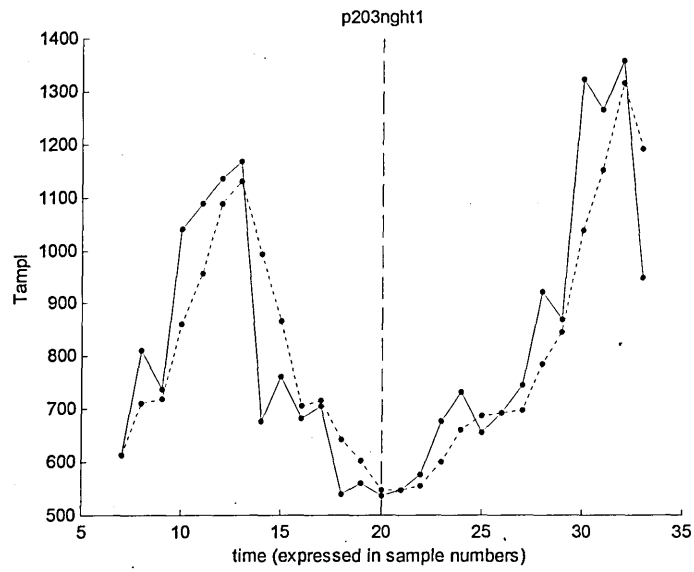


Figure 6.4: Tampl for 203nght1 showing delayed alarm (9 samples)

Investigating the glucose profile it is realised that the night started as hyperglycaemic with the glucose at 12.24 mmol/lt at the start of the night (record 7). By record 11, i.e. within an hour, the glucose had dropped to 2.2 mmol/lt. This is a very steep descent, going from one extreme (hyperglycaemia) to the other (hypoglycaemia). This very rapid drop in glucose could give an explanation as to why hypoglycaemia is manifested on the T amplitude feature, and detected by the KBS, very late. A rapid drop in glucose is expected to take longer before affecting the heart. On the other hand, the delay between slow changing glucose and the subsequent effect on the heart is expected to be shorter [PD2].

For patient 205 (night 1) a false alarm is raised because a drop in T wave amplitude, below the MA value, occurs for records 25 and 26 and also a prolongation in RTapexc, above the MA value occurs for these records as seen in Figures 6.5 and 6.6. (Abnormal changes in the two features also happen for record 27 but this does not affect anything since the alarm has been raised from the previous record.) These changes were very similar to the changes that happen during a hypoglycaemic event, which is why an alarm was raised. It seems impossible to avoid this false alarm using a monitoring system based only on the above two features. There is a lot of fluctuation in both features that cannot be masked simply by using the MA criterion, and this leads to a false alarm.

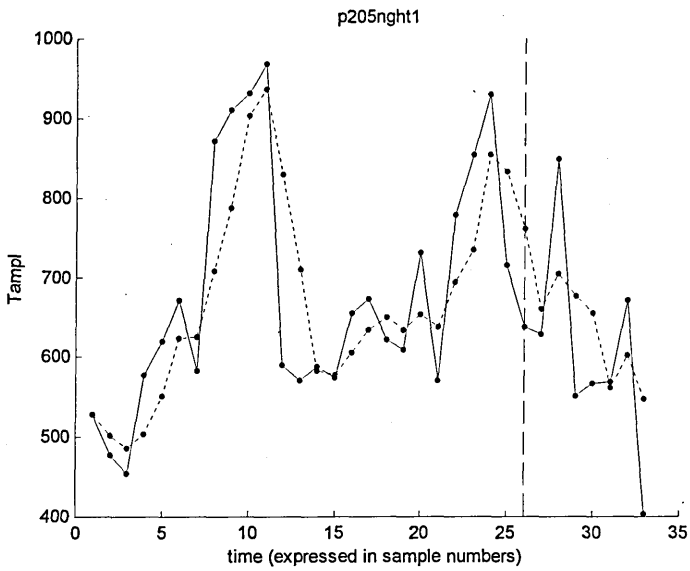


Figure6.5: Tampl graph for 205nght1 showing alarm @ rec26

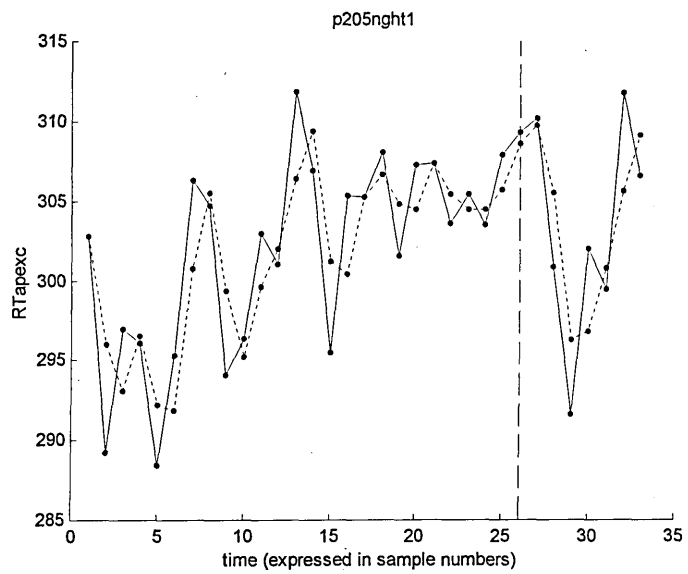


Figure 6.6: RTapexc graph for 205nght1 showing alarm @ rec26

In Section 6.5 it was stressed that an alarm system based on only one ECG feature would not be robust. The above alarm system was also assessed when only the T amplitude feature was used. For subjects 202 (night 2) and 204 it detected hypoglycaemia correctly while it was within the three samples for subject 209 (night 2). Such a system though has many false alarms. This was expected but assessment of a system based only on the T amplitude feature was carried out to see whether some

hypoglycaemic events occurring in the dataset, would be detected. The next section will present a monitoring approach using moving values of mean and standard deviation.

6.11 Monitoring System using MA and MSD

6.11.1 Monitoring Results

An improved approach based on the system presented in Section 6.10 was to incorporate the moving value of the standard deviation of the feature values in the system. In that case a significant event will occur if a feature value lies outside the mean ± standard deviation (SD) region. For the T amplitude feature, a reduction in amplitude below mean - SD is significant while an increase above mean + SD is ignored since only the flattening of the T wave has clinical significance. The opposite holds for the RTapexc feature where we are monitoring increases in feature value above mean + SD while RTapexc shortening below mean - SD is ignored. At each monitoring epoch the standard deviation was calculated from exactly the same moving window of data as the mean.

The main difference between the monitoring system presented in this section and the one using MA only (Section 6.10) was the inclusion of the MSD criterion. All other aspects of the system were the same (inputs, outputs, rule-base etc).

The moving window in the MA system was including the current feature value at the instant of monitoring for calculation of the mean. In the MA&MSD system both the approaches of including and not including the current sample were tested. Further experimentation (outlined in Figure 6.10) included the assessment of the RTc feature instead of the RTapexc and also the use of the uncorrected versions of both the RTc and the RTapexc features. Moreover the approach of freezing the window values once a potential risk occurred, in order to make the alarm system more sensitive to feature changes, was tested as will be described in Section 6.13. Finally the Receiver Operating Characteristic (ROC) approach was employed as presented in Section 6.11.2 in order to further improve the performance of the global KBS by tuning its parameters.

The accuracy, sensitivity and specificity of the global KBS using MA and MSD were: 78.13%, 22.22% and 100% respectively[33]. The results from the customised systems per patient are tabulated in Table 6.3. Apart from the standard fields, the table also contains the values of the optimal parameters (Window Sizes and Healthy Band Widths) for each patient. "scTampl" is a scaling factor by which the standard deviation is multiplied so that the width of the Healthy Band can be varied. The default parameter is 1 which effectively uses MA±MSD as the Healthy Band. Similarly for "scRTapexc". The entries of the table highlighted in **bold** denote cases where the parameters of the custom system differ from those of the global system.

In most cases, more than one set of tuning parameters corresponds to the optimal performance tabulated for each patient. For instance, for patient 202-night2, the acceptable parameters ranged in 1, 1, [0.5 2], [0.5 0.7] for TamplWS, RTapexcWS, scTampl and scRTapexc respectively[34]. Wherever there was agreement between the global and the custom parameters, the global parameters were shown in Table 6.3. This was done in order to identify the maximum number of patients showing agreement to a global set of parameters.

A significant improvement in performance of the system was observed once the standard deviation criterion was incorporated. The reason behind the improvement is the fact that an acceptable range of feature values was defined at each sampling instant. A significant event occurred only if the feature magnitude exceeded this range. According to the previous approach where only the moving average was used, any reduction below the mean in the case of T amplitude or increase for RTapexc would be a significant event. This leads to many false alarms because a lot of fluctuation in feature values occurs and such fluctuation is not necessarily corresponding to hypoglycaemia. A typical example of a false-alarm being rectified by the inclusion of MSD is 205-night1 that was discussed in Section 6.10. The system using MA falsely raised an alarm because T wave flattening and QT prolongation was observed. The improved system was able to infer that the above flattening and prolongation was within acceptable ranges.

---

[33] Current sample not included in moving window and freezing of window not allowed.

[34] For scTampl and scRTapexc a range of values is given.

Table 6.3: Alarm system results using MA & MSD (for RTapexc and Tampl features)

MA & MSD Results (not including current sample and no freezing of Window allowed)									
No	patient	gl<3	gl<2.5	alarm@rec	perf	TamplWS	RTapexcWS	scTampl	scRTapexc
1	p201Anght1	0	0	0	CE	1	4	1.9	0.7
2	p201Anght2	66	0	0	CE	1	4	1.9	0.7
3	p202A	23	24	**10**	**CAwithin13**	**1**	**1**	**1.9**	**0.7**
4	p202	41	41	41	CA	1	4	1.9	0.7
5	p203nght1	11	11	**18**	**CAwithin7**	**4**	**1**	**0.7**	**0.5**
6	p203nght2	0	0	0	CE	1	4	1.9	0.7
7	p204	17	18	17	CA	1	4	1.9	0.7
8	p205nght1	0	0	0	CE	1	4	1.9	0.7
9	p205nght2	0	0	0	CE	1	4	1.9	0.7
10	p207nght1	0	0	0	CE	1	4	1.9	0.7
11	p207nght2	0	0	0	CE	1	4	1.9	0.7
12	p208nght2	0	0	0	CE	1	4	1.9	0.7
13	p209nght1	0	0	0	CE	1	4	1.9	0.7
14	p209nght2	50	55	50	**CA**	**5**	**2**	**0.9**	**0.9**
15	p210	0	0	0	CE	1	4	1.9	0.7
16	p212nght1	0	0	0	CE	1	4	1.9	0.7
17	p212nght2	58	59	**50**	**CAwithin8**	**4**	**2**	**1**	**1**
18	p215nght1	0	0	0	CE	1	4	1.9	0.7
19	p215nght2	0	0	0	CE	1	4	1.9	0.7
20	p218Anght1	0	0	0	CE	1	4	1.9	0.7
21	p218Anght2	0	0	0	CE	1	4	1.9	0.7
22	p221	0	0	0	CE	1	4	1.9	0.7
23	p222	1	0	0	CE	1	4	1.9	0.7
24	p227nght1	21	22	21	**CA**	**2**	**5**	**1**	**0.7**
25	p227nght2	39	39	39	**CA**	**1**	**4**	**0.7**	**0.7**
26	p230nght1	0	0	0	CE	1	4	1.9	0.7
27	p231nght1	0	0	0	CE	1	4	1.9	0.7
28	p231nght2	0	0	0	CE	1	4	1.9	0.7
29	p232nght1	0	0	0	CE	1	4	1.9	0.7
30	p232nght2	65	0	0	CE	1	4	1.9	0.7
31	p244nght1	18 23	18 23	**10**	**CAwithin8**	**1**	**1**	**0.5**	**0.5**
32	p244nght2	0	0	0	CE	1	4	1.9	0.7

The summarised results (accuracy, sensitivity and specificity) of the customised systems were: 87.5%, 55.56% and 100% respectively. Once customisation of the system is allowed, the performance (predominantly the sensitivity) increases significantly. Out of the 32 nights used, originating from 19 patients, all 23 euglycaemic nights were monitored accurately. Out of the 9 hypoglycaemic nights, 5 were monitored accurately with the alarm being raised in the exact sample where the onset of hypoglycaemia occurred. For the remaining of the hypoglycaemic nights, alarms were still raised but subject to a time-deviation from the onset of hypoglycaemia. For 203-night1 the alarm was raised late by 7 sampling instants. For 212-night2 and 244-night1 the time-deviation was 8 samples and for 202A it was 13 samples.

It must be stressed here that any patients not complying with the clinical assumptions of this work will not comply with the rule-base used in the monitoring system. Such patients will not exhibit QT prolongation and T wave flattening under hypoglycaemia and will not be monitored successfully by the system. Although this may be a problem for a low-glucose-monitoring point of view, it is not one from a hypoglycaemia-related arrhythmia-detection point of view. Patients not exhibiting the above changes on the ECG are probably exhibiting normal VR (i.e. normal cardiac function) and may not be in danger of arrhythmia or sudden death although the glucose is very low[35].

The fact that the KBS yields low sensitivity does not mean that it is inadequate as a monitoring system. From the above discussion it becomes apparent that the KBS will raise alarms only for those patients exhibiting the assumed ECG changes. Patients not exhibiting the changes will, probably, not be in danger. The fact that the KBS raises alarms at the correct sample for a few subjects verifies the fact that it is an adequate system for monitoring. The KBS uses the minimal amount of input ECG features. Further improvements in performance will require the inclusion of more features, characterising ST segment changes and the presence or not of U waves.

6.11.2 Receiver Operating Characteristic (ROC)

The parameters of the KBS (width of healthy bands, and window sizes) were tuned by applying the Receiver Operating Characteristic approach which is a graphical method for simultaneously maximising the sensitivity and specificity by selecting appropriate values of the parameters of the system. The ROC curve when varying the width of the healthy band[36] is presented in Figure 6.7. The horizontal axis corresponds to "1-specificity" while the vertical one corresponds to "sensitivity". Optimal performance occurs for data-points as close as possible to the top left-hand corner of the figure. The value of the corresponding scaling parameter (that adjusts the width of the healthy band) is plotted next to each data-point. There is a trade-off between high sensitivity and high specificity and ROC allows choice of the best pair given the requirements of the user of the technique. All 4 parameters (window sizes and healthy bands independently for 2 ECG features) were tuned by ROC. Such an ROC graph is depicted in Figure 6.8.

---

[35] Very low glucose may not necessarily mean abnormally low glucose, from the arrhythmogenesis point of view.

[36] in the case where the same width is used for both features

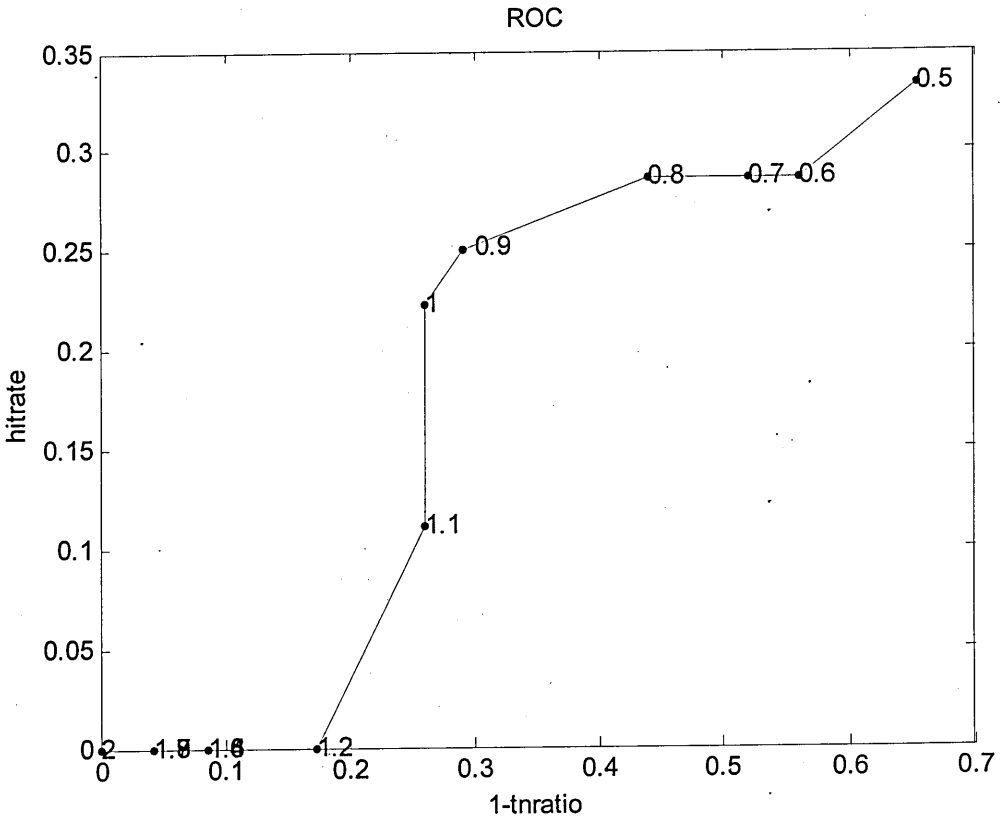


Figure 6.7: Receiver Operating Characteristic for tuning one parameter

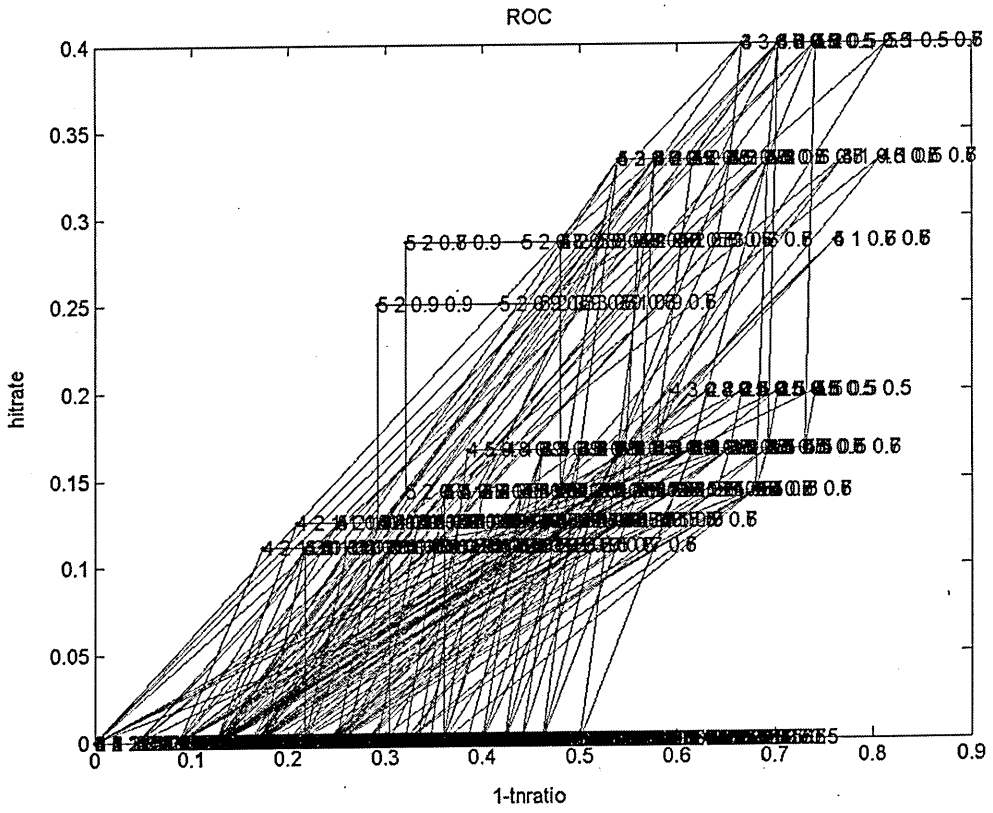


Figure 6.8: Receiver Operating Characteristic for tuning four parameters

All possible combinations of the 4 parameters were used in the ROC analysis. The window sizes varied in the interval [1 5] and the scaling factors adjusting the healthy band, in [0.5 2]. The graph looks "messy" because the area of the graph is swept up and down because of the variation of multiple parameters. The optimal set of parameters can easily be identified by zooming in. The ROC graphs were plotted in the MATLAB environment that was allowing the user to zoom in and read the parameters next to the coordinate point of interest on the graph. Figure 6.9 illustrates a section of an ROC graph after zooming in to read the parameters of interest. In many cases more than one set of tuning parameters were yielding the same performance causing overlapping points on the graph.

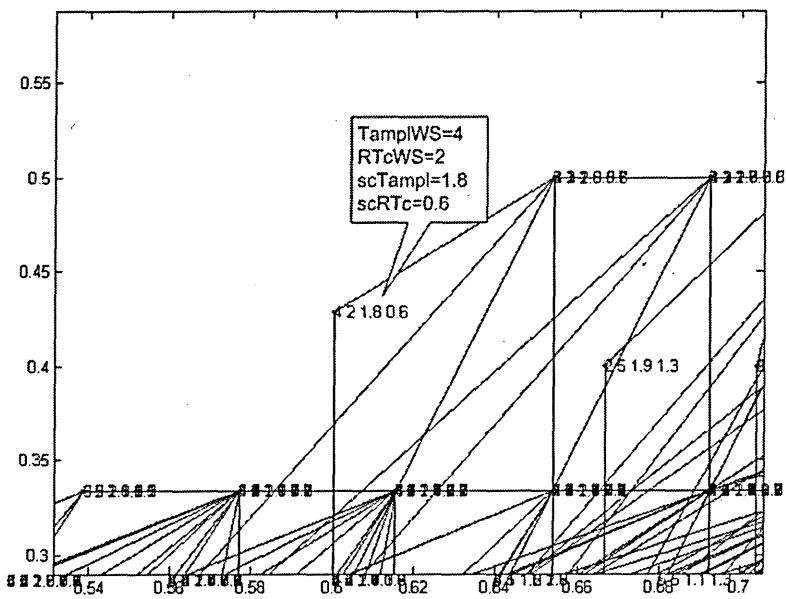


Figure 6.9: ROC graph after zooming-in.

For a number of different configurations of the KBS, there were many parameter-combinations yielding pairs in the bottom right corner of the ROC graph[37]. In a classification problem this corresponds to very bad classification performance. Nevertheless, such a classifier can be very useful when inverting its output. In our case, although many pairs occurred in the bottom-right corner, this was not useful because the KBS performs monitoring in time, as opposed to static pattern classification and therefore its output cannot be inverted.

[37] A weak classifier, will have data-points on the 45 degress line of the ROC graph, corresponding to 50% sensitivity and 50% specificity (i.e. a random classifier).

6.12 Outline of experiments

As mentioned earlier, a number of experiments were carried out in the development process of the knowledge-based monitoring system with some of them having been presented already. Before proceeding any further in the discussion of these experiments, a diagrammatic outline is given in Figure 6.10 to aid the reader in following the presentation.

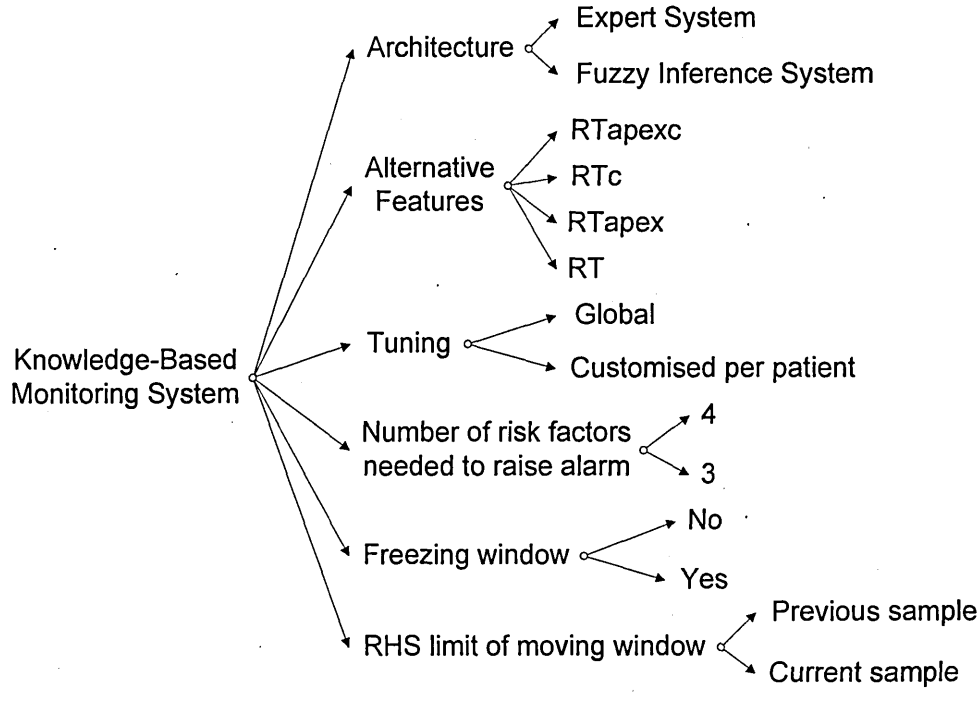


Figure 6.10: Outline of experiments carried out during the KBS development.

The issues that appear in the figure that have not been discussed yet are: the use of alternative versions of the time-interval feature that describes VR duration (RTapex, RT etc), the use of fewer risk factors as a requirement to raise an alarm, and the approach of freezing the window once a risk factor has occurred.

6.13 Modifications of KBS (FreezeW, up2current)

Besides adjusting of the monitoring system by means of varying the WS and HB parameters, a few more techniques were attempted in order to increase the performance. Such experimentation was carried out only on the MA&MSD system, since this was the main system produced for patient monitoring.

One dilemma encountered was whether to include the current feature value, at each sampling instant, in the calculation of the MA and MSD values or to use a window of past data spanning up to the previous sample. Inclusion of the current sample causes the

system to be less sensitive in feature changes over time. For instance a sharp increase in feature value from one sample to the next is more likely to be detected as abnormal if the current value is not included in the window. Including the current high value will affect the value of the MA and MSD and this high increase will be masked to some extent. Both options described above were attempted. The MA system presented in Section 6.10 was using a window that included the current value. On the other hand, the KBS yielding the results presented in Table 6.3 (MA&MSD) did not include the current sample in the moving window. Comparing the two approaches for the MA&MSD system led to the conclusion that exclusion of the current sample in the moving windows was the optimal approach. The monitoring results were inferior to those presented in Table 6.3 so they are not presented here.

Another idea that was implemented into the system was to "freeze" the MA and MSD values, once a significant event would occur i.e. to keep the window stationary. This would make the system more sensitive to abnormal features changes. The way this approach worked was to keep the current values of MA and MSD fixed for either feature once a potential risk would be raised. These values would be kept in a buffer for as long as potential risks were detected. The window would continue to move only if the potential risks would be reset to zero. After implementing this approach it was realised that it made the system very sensitive to feature changes and did not really introduce an improvement in performance. Results from this approach are again not included since they were inferior to those tabulated in Table 6.3.

6.14 Feature combinations including RTapex, RT and RTc

Further experimentation with the KBS led to the use of alternative features, instead of the RTapexc, that also produce estimates of the VR duration. The input ECG features to the system were always kept to two and Tampl was always one of them. The RTapexc feature was replaced in turn by RTapex, RT and RTc in equal tests of the monitoring system. RTapex and RT are the uncorrected versions (i.e. no decorrelation from the RR interval) of RTapexc and RTc respectively.

It has already been stressed that the RTc is a superior predictor of delayed VR compared to RTapexc since it also describes late VR phenomena reflected on the T downslope that RTapexc cannot describe. The only reason that the RTapexc was initially used in the

system was the caution towards the robustness of the T wave end annotation algorithm[38] used. Using the RTapexc feature we were more confident that the feature changes would exclusively be due to a cardiac event, whereas with the RTc feature they could either be due to a cardiac event or due to a weakness of the algorithm under more difficult circumstances (i.e. noise, artefacts etc). In later stages of development of the KBS, the RTc feature was used instead of RTapexc to test the quality of the former as a VR predictor besides testing the performance of the T-end annotation algorithm.

The experimentation with the RTc feature did not improve the global KBS. However, when considering customised-per-patient systems it helped identify a few hypos more closely. For subject 202A hypoglycaemia started at records 23-24. It was detected at record 18 (5-6 samples early) when using RTc while it was detected at record 10 when using RTapexc. Therefore the use of the RTc feature, although not achieving detection of the hypo, gave an alarm closer to the onset of hypoglycaemia. For 203nght1, when RTc was used the hypo was detected at record 15 (only 4 samples late). When RTapexc was used the earliest it could be detected was record 18 (7 samples late). For 212nght2, the alarm was still raised at record 50 when RTc was used and the chosen parameter combination (TamplWS, RTapexcWS, TamplHB, RTapexcHB) for RTapexc was (4, 2, 1, 1) while the closest one when RTc was used was (4, 2, 1, 0.9). For 244nght1 the hypo could not be detected at all when the RTc feature was used. When RTapexc was used, an early alarm was raised at record 10 (8 records early). When assessing the system on "spot on" alarms the use of RTc did not yield any improvement. However, if a deviation of up to 5 samples was considered as acceptable, then the use of RTc increased the sensitivity from 55.56% to 77.78%.

Besides the experimentation with the RTc interval, the RT and RTapex features were examined. These tests were carried out to see how useful the heart-rate-correction was. It had been observed that the correlation coefficient between RTapexc and RR did not reach satisfactorily low values, which means that the HR correction did not manage to decorrelate the two variables very well. According to the correlation coefficient, HR correction was better for the RTc interval compared to the RTapexc. Doubts about the quality of HR correction led to the use of the uncorrected versions of the features. After

[38] Throughout the study of producing a KBS system, the tangent method was used to annotate T wave end.

extensively testing the system, it was concluded that the uncorrected versions did not introduce any improvements.

In order to be concise and not overcrowd this section with tabulated results, the experimentation with the extra features described above was only presented descriptively, outlining only the improvements introduced and excluding the remaining results.

6.15 Using three significant events to raise alarms

Up to this point, the rule-base of the monitoring system was such that it would require two successive significant events (potential risks) in both features to raise an alarm i.e. 4 significant events in total. In an attempt to make the system more sensitive to alarms, the scenario where only 3 events would be enough was examined. Such events would either be two potential risks from the previous cycle and one on either feature on the current cycle (2+1) or one potential risk from the previous cycle and two, on both features, on the current cycle (1+2). This means that instead of four, three of the events that the system was looking for in its previous version would be enough to raise an alarm. The results using this analysis are presented below for the RTapexc and the RTc features[39].

When the RTapexc feature was used the following improvements were observed:

1. For 202A the hypo was detected at record 19 and also at record 26 depending on the choice of tuning parameters (hypo onset was at records 23-24). This is a significant improvement compared to when using 4 significant events (very early alarm at record 10).
2. For 203-night1 the alarm was raised at record 17 i.e. the use of only 3 events raised the alarm one sample closer to the onset of hypo (at record 11).
3. For 212-night2 the alarm was raised at record 51 which is again one sample closer to the onset of hypo compared to the use of 4 events.
4. For 244-night1 the alarm was raised at record 22 i.e. 4 samples after the onset of a brief period of hypo and one sample before the onset of the main hypo period. This is a significant improvement.

---

[39] For the configuration where current sample not included in moving window and freezing of window disabled.

If a deviation of up to 2 samples is acceptable then the use of 3 significant events for the RTapexc feature increased the sensitivity from 55.56% to 77.78%. This improvement in sensitivity is the same quantitatively with that achieved by the use of RTc instead of RTapexc when 4 significant events were used.

The results when the RTc feature was used along with 3 significant events, are analysed in more depth and the full set of results for the customised ES are presented in Table 6.4. Again the results in bold denote departures of the tuning parameters from those values corresponding to the global KBS.

Table 6.4: Alarm system results (for features: RTc, Tampl) when using 3 significant events

patient	gl<3	gl<2.5	alarm@rec	perf	TamplWS	RTcWS	scTampl	scRTc
p201Anght1	0	0	0	CE	5	1	2	1.7
p201Anght2	66	0	0	CE	5	1	2	1.7
p202A	23	24	**27**	**CAwithin3**	**5**	**1**	**1.7**	**1.7**
p202	41	41	37	**CAwithin4**	5	1	2	1.7
p203nght1	11	11	**15**	**CAwithin4**	**5**	**1**	**1.7**	**1.7**
p203nght2	0	0	0	CE	5	1	2	1.7
p204	17	18	18	CA	5	1	2	1.7
p205nght1	0	0	0	CE	5	1	2	1.7
p205nght2	0	0	0	CE	5	1	2	1.7
p207nght1	0	0	0	CE	5	1	2	1.7
p207nght2	0	0	**0**	**CE**	**4**	**1**	**2**	**1.7**
p208nght2	0	0	0	CE	5	1	2	1.7
p209nght1	0	0	**0**	**CE**	**5**	**2**	**2**	**1.7**
p209nght2	50	55	**53**	**CAwithin2**	**5**	**2**	**1.8**	**0.7**
p210	0	0	0	CE	5	1	2	1.7
p212nght1	0	0	**0**	**CE**	**4**	**1**	**2**	**2**
p212nght2	58	59	**58**	**CA**	**3**	**4**	**2**	**0.8**
p215nght1	0	0	0	CE	5	1	2	1.7
p215nght2	0	0	0	CE	5	1	2	1.7
p218Anght1	0	0	0	CE	5	1	2	1.7
p218Anght2	0	0	0	CE	5	1	2	1.7
p221	0	0	0	CE	5	1	2	1.7
p222	0	0	0	CE	5	1	2	1.7
p227nght1	21	22	**21**	**CA**	**5**	**1**	**0.9**	**1.7**
p227nght2	39	39	**39**	**CA**	**5**	**1**	**0.9**	**1.7**
p230nght1	0	0	**6**	**FA**	5	1	2	1.7
p231nght1	0	0	0	CE	5	1	2	1.7
p231nght2	0	0	0	CE	5	1	2	1.7
p232nght1	0	0	0	CE	5	1	2	1.7
p232nght2	0	0	0	CE	5	1	2	1.7
p244nght1	18	18	**21**	**CAwithin3**	**5**	**1**	**1.4**	**1**
p244nght2	0	0	**41**	**FA**	5	1	2	1.7

When alarms are classed as correct only if they are raised on the exact record of hypo-onset then the accuracy, sensitivity and specificity of the system are: 78.13%, 44.44%,

91.30% respectively. However if alarms raised within 4 samples are classed as acceptable then the accuracy and sensitivity reach 93.75% and 100% respectively.

The use of 3 significant events made the system more sensitive to ECG feature changes. Although the detection of hypos was improved, two false-alarms were raised in euglycaemic nights. Such nights were p230nght1 and p244nght2. Specifically for the case of p230nght1 the alarm happened too early, that is at record 6 while the first valid record of the night was record 4. This means that the system had only two past samples available to use for setting the thresholds used for monitoring. If a restriction would be set for monitoring to start later, e.g. after the fifth record of the night so that the system would have a chance to adapt, this false alarm could be avoided.

Regarding the hypoglycaemic nights, the following improvements were observed when using 3 significant events:

1. For p202A the hypo was detected at record 18 (in agreement with the case where 4 significant events were used) and also at record 27 depending on the choice of tuning parameters. The alarm at record 27, which is the closest to the onset of hypo, could not be produced when 4 events were used.
2. For p203nght1 the alarm was still raised at record 15 i.e. the use of only 3 events did not introduce any improvement.
3. For p212nght2 the alarm was raised at record 58 which is a significant improvement since the blood glucose dropped below 3 mmol/lt at that exact record.
4. For 244nght1 the alarm was raised at record 21 i.e. 3 samples after the onset of a brief period of hypo and 2 samples before the onset of the main hypo period. Use of three significant events also resulted in a valuable improvement since no alarm was raised at all when 4 events were used.

If a deviation of up to 4 samples is acceptable then the use of 3 significant events for the RTc feature increased the sensitivity from 55.56% to 100%, as mentioned earlier.

6.16 A Monitoring System Incorporating Fuzzy Logic

During the process of the development of the monitoring system, Fuzzy Logic was also considered. A Fuzzy Inference System (FIS) can be seen as a generalisation of an expert system where continuous degrees of membership, as opposed to binary ones, are used. This feature makes Fuzzy Logic a powerful tool due to the smooth transition from one membership function to the next. As mentioned in Section 2.5.1, previous work on hypoglycaemia detection by a FIS was encountered in the work of Hastings et al [Hastings 1998] who presented a prototype hypoglycaemia detector that was using peripheral physiological responses (sweating and HR) to falling blood glucose.

The expert system used for monitoring, in previous sections, was converted to a FIS. The rule-base of the system was kept the same but fuzzy logic was introduced to replace the crisp logic previously used. This introduced the advantage that, when an alarm was raised it would be raised with a degree of certainty in the interval [0.5 1]. Similarly for the cases where no alarm was raised the output would lie in [0 0.5) and the user would be able to see how close to the threshold of 0.5 the output had reached. Providing a degree of certainty when raising an alarm and also informing how close to an alarm the system has reached, when one is not raised, is useful for clinicians and users of the system. Such a degree of certainty was also produced for the Tampl and RTapexc potential risk outputs besides the "Diabetic State" output.

A "Mamdani" system was used where both the antecedent and the consequent parts of each rule are fuzzy. The MATLAB fuzzy logic toolbox (ver 2) was utilised. The characteristics of the system are given in Table 6.5 presented below:

Table 6.5: FIS parameters

type:	**Mamdani**
AND Method:	**min**
OR Method:	**max**
Implication Method:	**min**
Aggregation Method:	**max**
Defuzzification Method:	**centroid**

The system is very similar to the expert system presented earlier. The only difference is that fuzzy logic is introduced. The inputs and outputs are kept the same. The fact that the rule-base is also the same justifies the reason why a Mamdani system was used.

6.16.1 Membership Functions (MFs)

Triangular and trapezoidal membership functions (MFs) were used throughout. Three MFs were used for the Tampl and RTapexc inputs. The left-most and right-most MFs were trapezoidal while the middle one was triangular. For the rest of the variables, two trapezoidal MFs were used. An illustration of the MFs for Tampl (LHS) and RTapexc (RHS) can be seen in Figure 6.11. The linguistic values used to label the MFs are

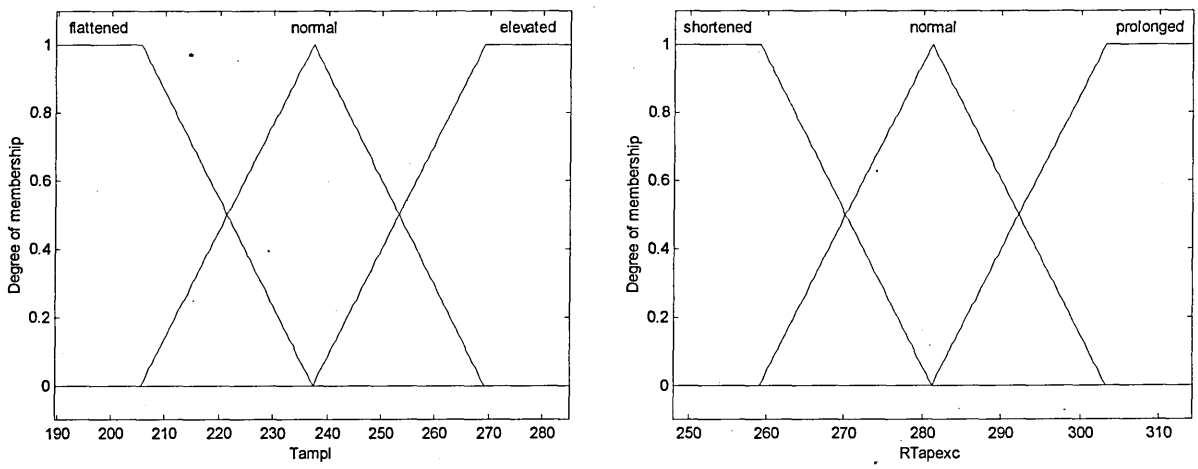


Figure 6.11: MFs for Tampl (LHS) and RTapexc (RHS)

apparent in the figures. For Tampl they are: "flattened", "normal" and "elevated" and for RTapexc: "shortened", "normal" and "prolonged". In line with the rule-base of the Expert System, only Tampl magnitudes belonging to the "flattened" MF and RTapexc magnitudes belonging to the "prolonged" MF would raise an alarm. The universe of discourse of Tampl is in mV while for the RTapexc it is in msec.

The system was adaptive, similar to the Expert System, and the MFs were updated on every monitoring epoch based on the calculated MA and MSD. The shapes of the MFs was kept fixed (i.e. triangular, trapezoidal) but the positions of the peaks (or flat segments) and also the slopes were varied. For the two ECG features (Tampl, RTapexc) using 3 MFs, the middle one (triangular) had its peak at the MA value calculated for that monitoring epoch and the cross-over points left and right of the peak occurred at MA-MSD and MA+MSD. The LHS and RHS trapezoidal MFs were reaching a membership degree of 1 at MA-2*MSD and MA+2*MSD respectively. The MFs for the

potential-risk variables and also the "DiabeticState" output had a universe of discourse in [0 1] and a cross-over points at 0.5 and were not adapted during monitoring.

Tuning of window sizes and healthy bands was carried out in the same way as in the Expert System. When the healthy band width was tuned, the cross-over points for the middle MF would lie at MA-k*MSD and MA+k*MSD where k is the scaling factor used to achieve variable MF width. In such a case, the LHS and RHS trapezoidal MFs were reaching a membership degree of 1 at MA–2*k*MSD and MA+2*k*MSD respectively. The MFs for the potential-risk variables and also the "DiabeticState" were not adapted as already mentioned.

The 3D surface for inputs Tampl and RTapexc and for the DiabeticState output is shown in Figure 6.12. The inputs and outputs relating to the potential risks are not included in order to achieve visualisation in 3 dimensions.

The 3D surfaces look identical because the axes are adjusted separately for each graph to allow better visualisation. The adaptivity of the FIS can be seen if the range of the axes in the two graphs is observed. For both input features the healthy range is smaller at the time of the alarm

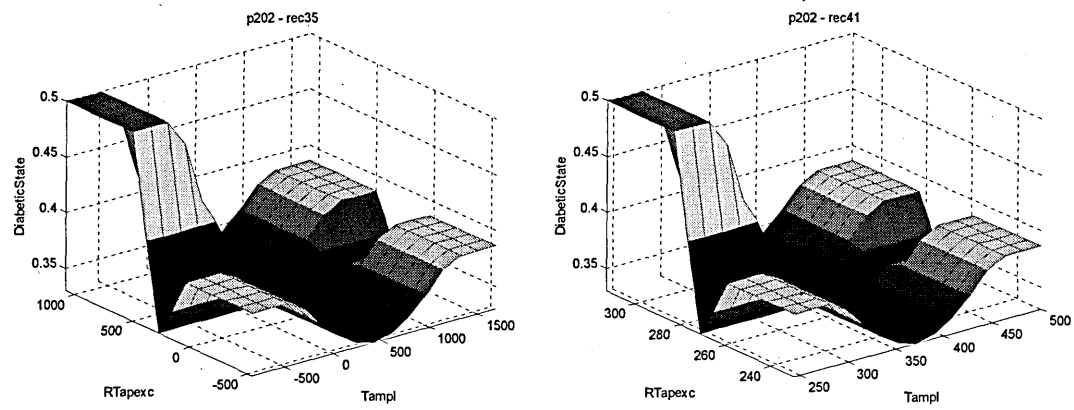


Figure 6.12: 3D surface for 2 input ECG features and DiabeticState output, at the start of the night (LHS) and at the time of the alarm (RHS)

Table 6.6 contains the linguistic values used to label the MFs of all the variables used in the system. The column labelled "middle MF" does not have a linguistic value for the variables using only 2 MFs.

Table 6.6: Linguistic values of all MFs

variable	LHS MF	middle MF	RHS MF
Tampl	flattened	normal	elevated
RTapexc	shortened	normal	prolonged
Tampl_risk	low	----------	high
RTapexc_risk	low	----------	high
DiabeticState	low	----------	high

6.16.2 Fuzzy Results

The performance of the fuzzy monitoring system is presented in Table 6.7. The table layout is similar to previous tables containing results on the Expert System. The only difference is that the outputs are not binary but continuous in the interval [0 1]. The threshold of 0.5 separates the two classes of euglycaemia and hypoglycaemia in the DiabeticState output. The same threshold separates the two classes of low and high risk in the other two outputs. The alarm strength at each sampling instant is presented at the relevant column. The system presented in Table 6.7 is the global FIS for parameters (1, 4, 1.9, 0.7) for (TamplWS, RTapexcWS, TamplHB, RTapexcHB)[40]. It corresponds to the Expert System presented in Table 6.3. Features Tampl and RTapexc were fed and the same tuning parameters were used. Also in both systems, the moving window did not include the current value and also freezing of the moving windows was not allowed while 4 significant events were needed to raise an alarm. The extra fields in Table 6.7, which were not included in Table 6.3 are "alarm strength", "TamplRisk" and "RTapexcRisk". "alarm strength" corresponds to the "DiabeticState" output and the other two refer to the potential risks, from the previous record, associated with features Tampl and RTapexc.

As expected the results produced by the two systems (ES and FIS) are exactly the same, since the only difference was the introduction of Fuzzy Logic. The enhancement that the FIS introduced was the fact that a degree of certainty for the alarm output and the potential-risk outputs is provided at every monitoring epoch. This is very useful in quantifying the possible risk of hypoglycaemia onset both when an alarm is raised and

[40] The tuning parameters are not tabulated since they are fixed (global system presented) and also because of space considerations on the page.

when otherwise. Since the results were the same, further discussion specifically on the performance of the FIS will not be provided in this section.

Table 6.7: Fuzzy monitoring system performance

No	pat	gl<3	gl<2.5	alarm@rec	alarmstrength	TamplRisk	RTapexcRisk	perf
1	p201Anght1	0	0	0	0.33	0.37	0.54	CE
2	p201Anght2	66	0	0	0.35	0.42	0.37	CE
3	p202A	23	24	0	0.37	0.47	0.48	MH
4	p202	41	41	41	0.5	0.53	0.6	CA
5	p203nght1	11	11	0	0.44	0.64	0.54	MH
6	p203nght2	0	0	0	0.37	0.5	0.5	CE
7	p204	17	18	17	0.51	0.5	0.61	CA
8	p205nght1	0	0	0	0.5	0.58	0.47	CE
9	p205nght2	0	0	0	0.35	0.5	0.42	CE
10	p207nght1	0	0	0	0.34	0.4	0.42	CE
11	p207nght2	0	0	0	0.36	0.5	0.49	CE
12	p208nght2	0	0	0	0.35	0.5	0.39	CE
13	p209nght1	0	0	0	0.38	0.61	0.5	CE
14	p209nght2	50	55	0	0.33	0.42	0.39	MH
15	p210	0	0	0	0.33	0.5	0.37	CE
16	p212nght1	0	0	0	0.35	0.36	0.43	CE
17	p212nght2	58	59	0	0.35	0.54	0.36	MH
18	p215nght1	0	0	0	0.33	0.5	0.37	CE
19	p215nght2	0	0	0	0.34	0.37	0.42	CE
20	p218Anght1	0	0	0	0.43	0.59	0.64	CE
21	p218Anght2	0	0	0	0.37	0.52	0.5	CE
22	p221	0	0	0	0.35	0.5	0.43	CE
23	p222	0	0	0	0.44	0.47	0.62	CE
24	p227nght1	21	22	0	0.37	0.41	0.5	MH
25	p227nght2	39	39	0	0.35	0.39	0.6	MH
26	p230nght1	0	0	0	0.34	0.39	0.47	CE
27	p231nght1	0	0	0	0.33	0.39	0.5	CE
28	p231nght2	0	0	0	0.34	0.38	0.42	CE
29	p232nght1	0	0	0	0.43	0.64	0.61	CE
30	p232nght2	0	0	0	0.34	0.52	0.38	CE
31	p244nght1	18	18	0	0.37	0.63	0.5	MH
32	p244nght2	0	0	0	0.36	0.38	0.44	CE

6.17 Discussion

A concluding discussion on the research direction of producing a KBS for hypoglycaemia monitoring is presented in this section.

6.17.1 Knowledge-Based Monitoring System versus Neural Networks

The approach presented in this chapter significantly improved the performance of the task of detecting the onset of spontaneous nocturnal hypoglycaemia. The use of a Knowledge-Based approach proved superior to neural (MLP) and statistical (LDA and

kNN) classifiers previously used. This is due to two main reasons. First, the incorporation of temporal information into the KBS so that the system was consulting previous ECG feature samples besides the current ones when inferring the symptomatic status of a patient; and second, the incorporation of human-expert knowledge.

The incorporation of temporal information meant that the problem tackled was no longer a static pattern classification problem as in the case of MLPs, LDA and kNN but a time-series analysis problem with the system being adaptive as time elapsed. This gave a competitive advantage to the KBS in performing better. As seen in the data used in this project, a lot of transient events occur due to dynamic changes of the ECG signal. The pattern classifiers were unable to make use of these transients and the opposite effect was seen; that of confusing the classifier due to the varying baselines of the feature vectors used. It is suspected that the performance of ANNs would be significantly improved upon the incorporation of the time variable. This was not possible with the existing datasets since the amount of data was not sufficient for training time-lagged neural networks. Bearing in mind that the data fed to an ANN should ideally be representative of all classes to be classified it becomes obvious that in our dataset the data is further reduced when using ANNs because not all the euglycaemic nights, which are a lot more than the hypoglycaemic, can be used. On the other hand, when using the Knowledge-Based approach, the system can be assessed on all data available since no euglycaemic nights are left out.

The other reason for giving the KBS a competitive advantage was the incorporation of human-expert knowledge. Not only had this made it possible to produce a working system using a dataset significantly smaller to what a time-lagged neural network would require, but it also guided the system in identifying only the significant ECG changes and ignoring the useless ones. This comprises a typical problem in neural network research. The ANN must decide on its own what comprises structure and what noise in the data which is a very difficult task and requires very lengthy datasets to achieve this. For the case of the KBS the rule-base dictated, for instance, that abnormal T wave elevations could be ignored and similarly for abnormal RTc shortening. It also dictated that abnormal changes, in the right direction, had to be successive in time. This knowledge was extremely useful in boosting the performance of the monitoring system.

Besides the above arguments, the use of a KBS had the advantage that it required tiny amounts of processing power during its development, compared to ANNs. ANNs required a large amount of time to train. This was because a different ANN was produced for each patient and also because multiple ANNs were trained for a given patient each one starting from different random initial conditions. Taking into account that a few ECG feature combinations were fed to the ANNs, the time taken to train all the ANNs needed was multiplied by the number of feature combinations used.

Another problem with neural networks was the danger of overfitting. There are high chances, in the process of training ANNs, to produce a not so useful ANN that cannot generalise well on unseen data. Overfitting was not a problem with the KBS. Even when the system was customised to a specific patient, the internal structure was meaningful to a human observer. This transparency of the system would allow an observer to study why a KBS is performing well on a given patient but not on another one.

Finally the incorporation of human-expert knowledge was very useful for researchers in the field, especially clinicians. A rule-base allows them to understand how the system operates, and also trust it is doing a wise thing. ANNs, being black-box models, were often faced with extreme caution and mistrust in the biomedical community; this is not the case for KBS. Besides the fact that the knowledge of the KBS is formulated in a meaningful way for human-experts it is also useful in validating clinical assumptions. In our case, a clinical assumption was provided[41] by our medical collaborators, this assumption was then formulated in a rule-base and successful use of this rule-base provided the necessary feedback to the clinicians to support and strengthen their assumption.

A drawback of using Knowledge-Based systems is that the expert knowledge must be available and it must be successfully coded into the system. The knowledge acquisition and the knowledge representation processes in the KBS context can be difficult tasks to perform.

[41] That of the flattening and prolongation of the T wave under hypoglycaemia

6.17.2 Detection of the onset of hypoglycaemia versus detection of life threatening arrhythmias.

The findings of this research regarding ECG analysis and monitoring can either be directed towards addressing the problem of the detection of the onset of hypoglycaemia or towards tackling the problem of hypoglycaemia-related life threatening arrhythmias. These two problems are related since severe nocturnal hypoglycaemia may lead to cardiac arrhythmias.

The first challenge seems a lot more difficult since monitoring the levels of glucose is attempted indirectly through analysis of the patient's ECG. Abnormally low glucose causes adrenaline release and potassium depletion which both affect the ECG and it is only through these ECG changes that the drop in glucose can be detected. If there is no manifestation of the dropping glucose on the ECG, then the abnormal drop cannot be detected. This is a problem for a hypoglycaemia-detection system since the hypos not reflected on the ECG will not be detected. However, the cost of not detecting these hypos is suspected to be very low from an arrhythmia-prevention point of view. This is because abnormally low glucose not causing delayed VR will probably not be dangerous as it will probably not lead to arrhythmogenesis. Hence a system detecting only those hypos that are manifested on the ECG could be useful in preventing nocturnal deaths of diabetics related to the "Dead in Bed" syndrome.

Investing further research effort in the system proposed in this chapter may lead towards a monitoring system for "Dead in Bed" prevention. The KBS produced was in many cases able to raise an alarm at the correct sample. If a fatal cardiac arrhythmia is to be developed, even a late alarm can be invaluable in saving the patient. There are a few examples from the dataset where patients (202, 204, 227, 244) were under hypoglycaemia for a few hours with the glucose being at 2.2 mmol/lt or below[42] and the person remained healthy. The most prominent case was that of 202-night2 where the patient was at 2.2 mmol/lt for 4.5 hours and below 3 mmol/lt for 5.5 hours and remained healthy. This gives an indication that even an alarm raised a few hours late could be invaluable in saving the patient from the occurrence of fatal cardiac arrhythmias. A long period of hypoglycaemia will probably have to occur before the genesis of a dangerous

[42] The actual values below 2.2 are not known because of a limitation of the MiniMed sensor as it was stressed in section 3.2.2.

arrhythmia and in that long period of hypoglycaemia there is high probability that ECG changes will occur, in the form of T flattening and QT prolongation, which will be detected by the monitoring system.

6.17.3 Patient-Oriented Customisation

As mentioned earlier in the chapter, two approaches were followed in producing a KBS. A global KBS was produced that was aimed at having optimal performance when monitoring all patients i.e. it was challenged to tackle both inter-patient and intra-patient variability in the ECG features. Moreover, customised systems were used for the elimination of inter-patient variability. These customised systems would focus on the dynamics of the specific patient to be monitored. This was expected to improve performance. The customisation involved the adjustment of the Window-Size and Healthy-Band parameters. All other aspects of the system were fixed.

A global system would be more useful in producing a generic model for the monitoring problem discussed. Such a generic model is useful academically. However tackling the real-life problem of patient monitoring will require customised systems. The increased performance of the customised systems on the dataset contributed in validating their use.

6.17.4 Static Pattern Classification Performance versus Monitoring System Performance

The performance of the monitoring approach was superior to that of the pattern classification approaches. However, this is not easily visible on the performance metrics because the performance is assessed differently in the two cases.

This becomes apparent by comparing two sets of results from the KBS and the MLP. The results from the customised KBS presented in Table 6.3 were: 87.5%, 55.56% and 100% for accuracy, sensitivity and specificity respectively[43]. The MLP classification results (on unseen data and when MLPs were customised per patient) from Section 5.2.4 were 70.15%, 75.43%, 64.10% for accuracy, sensitivity and specificity respectively.

[43] when not including current sample in moving windows, with no window freezing allowed and when assessing alarms as acceptable only if they were raised on the exact sample.

From the above metrics, the KBS results look superior in terms of the specificity only. However, they are superior in terms of the sensitivity as well.

Quantifying performance in the Monitoring studies is different to that applied in Pattern Classification. In the monitoring studies, performance is assessed on a patient-night by patient-night basis. In the classification by MLPs and LDA the performance is assessed on a pattern (i.e. ECG feature vector) by pattern basis. In the monitoring study, a patient is either monitored correctly or not, i.e. binary outcome. On the other hand, when using MLPs and LDA the accuracy per patient lies in the interval [0 100]%. To produce the overall performance metrics, for the MLPs presented above, the per-patient metrics are averaged.

The above discussion becomes clearer once a patient is inspected. The MLP test results for 204 were 58.33%, 62.00% and 58.67% for accuracy, sensitivity and specificity respectively. The KBS raised an alarm on the correct record for this patient. If patient 204 is analysed on a pattern by pattern basis then the KBS will give a sensitivity and specificity of 100%. Similarly for patient 227 (both nights merged together) the MLP yielded 62.00% 65.86% 68.67% while the sensitivity and specificity (on both nights) would be 100% by the KBS. Similarly, whenever a euglycaemic night was monitored correctly by the KBS the metrics would be 100% whereas the MLP result was always inferior.

In order to be able to compare the performance of the MLP and KBS approaches, the KBS performance will have to be assessed on a pattern-by-pattern basis and the results averaged. This was not followed since assessing the performance of the KBS in a pattern-by-pattern basis is not very informative in a monitoring study.

6.17.5 Optimal KBS Configuration and Optimal ECG Features

The monitoring system configuration according to which the moving window was frozen once a potential risk occurred, in an attempt to make the system more sensitive to abnormal changes, did not introduce an improvement. Similarly the inclusion of the ECG feature values from the current monitoring epoch into the moving window, did not introduce any improvement.

Regarding the ECG features, the use of the uncorrected versions of RTapexc and RTc did not yield any improvement in the performance of the system. The optimal results were produced by employing heart-rate correction on the features describing VR duration (RTapexc, RTc). As mentioned earlier, the customised KBS results when using the RTapexc were 87.5%, 55.56% and 100% for accuracy, sensitivity and specificity respectively. The use of the RTc instead of the RTapexc did not improve these results when assessing spot-on alarms. However, if a deviation of 5 samples is considered acceptable when assessing alarms, the use of the RTc increased the sensitivity from 55.56% to 77.78%.

The requirement of three significant events (risk factors) for an alarm, instead of four, gave the following metrics: 78.13%, 44.44%, 91.30% when alarms are classed as correct only if they are raised on the exact record of hypo-onset. However if alarms raised within 4 samples are classed as acceptable then the metrics reach 93.75%, 100% and 91.30%.

6.18 Conclusions

This chapter focused on the design of a Knowledge-Based System for the detection of the symptomatic status of patients experiencing spontaneous nocturnal hypoglycaemia. In this chapter the research focus moved from ECG pattern classification to monitoring of patients by monitoring of their ECG during the night. Offline monitoring of patients from the dataset was carried out in an approach simulating an online monitoring situation. The system was monitoring the time-series of two ECG features (T amplitude and VR duration) and was adapting itself as time elapsed. The KBS was realised both as an Expert System and a Fuzzy Inference System.

A short Rule-Base was produced to formulate the Knowledge-Base of the system. It consisted of eight rules. Such a Rule-Base can be very useful for clinical experts and the fact that the one developed in this work is very concise allows easy inspection by clinicians. The Knowledge-Base of the system was based on the clinical hypothesis (presented in Section 1.1.5) according to which abnormally low glucose encountered in hypoglycaemia is reflected on the ECG in the form of T wave flattening and QT prolongation. The performance of the monitoring system strongly supports this hypothesis raising optimism for its validation, once more data can be captured to allow future research.

The need for customised systems on each patient to be monitored was emphasised during this work. The research task of producing a monitoring system can be facilitated significantly once inter-patient variability can be overcome. This can be achieved by producing a tailor-made monitoring system aimed at the specific patient to be monitored. To achieve this, a future monitoring system based on the prototype presented in this chapter will need a period of learning and customisation on the patient to be monitored before the actual monitoring will start.

Although the proposed system focused on the detection of the symptomatic status of hypoglycaemia, it could also be a candidate after certain modifications in tackling the task of detecting the onset of cardiac arrhythmias leading to Sudden Death. If a monitoring system for Sudden Death was to be produced based on the approach presented in this chapter, then it should be tuned only to detect excessive ECG changes otherwise many false-alarms would be raised.

Further discussion of the impact of the Knowledge-Based monitoring system on the initial aims and objectives of this doctorate work will be presented in the next chapter (Chapter 7) which will conclude the thesis and provide recommendations for further work.

Final Discussion, Conclusions and Further Work

7.0 Introduction

This thesis focused on the investigation of the relationship between hypoglycaemia and cardiac function and also the detection of the onset of hypoglycaemia using solely ECG information. This chapter summarises and concludes the thesis. First of all it presents the main research challenges related to the physiological conditions addressed. It discusses the extent to which the goals initially set were achieved and the impact they have on the investigation of the clinical problem addressed. It proposes a way forward for further investigation of the problem by outlining a number of suggestions.

7.1 Discussion of Research Challenges

This section discusses various issues that make this research work a challenging study and many of the serious obstacles encountered during this project. Some of the obstacles could not be fully tackled in this research work and are subject to further investigation. The challenges are organised in subsections according to their relevant areas.

7.1.1 Issues related to Data Acquisition and Dataset

A novel dataset was used as a basis for this research work and allowed some new insights into the relationship between hypoglycaemia and the ECG. It was observed that the ECG changes due to hypoglycaemia were short-time transients (Sections 5.2.6 and 6.14.1). This suggests the inclusion of temporal information in the classification system. Unfortunately the data available was not sufficient for training time-lagged neural networks[44]. Such neural networks require long recordings of data. Because of this, only static neural classifiers (i.e. no time stamps), were used which is believed to have caused a compromise in performance.

[44] Networks receiving time-series inputs.

Another issue was that the two conditions studied, i.e. euglycaemia and hypoglycaemia, were not equally represented in the recordings since the hypoglycaemic events in the dataset were spontaneous. There is a significant advantage in investigating spontaneous hypoglycaemia namely that it is a realistic complication as opposed to experimental hypoglycaemia which is artificially induced. The drawback is that since the diabetic patients had to experience hypoglycaemia naturally, there were more recordings containing normal events than hypoglycaemic. Referring back to the task of training neural classifiers, the training data used should be representative of all conditions to be classified. The existence of less hypoglycaemic records did not allow maximal use of the data and restricted in some cases the size of the training files.

As mentioned in Section 3.1, there exist a number of online databases [URL 13] containing ECG records with annotations from human experts. In many research projects, such databases are the only source of data. Many researchers use these databases to design and evaluate feature extraction algorithms, classify cardiac beats etc. In our case such databases were not useful since the accompanying glucose data at the time the ECG was captured was not recorded. Moreover, the data on the online databases did not generally originate from diabetic patients, i.e. the patient group we were studying. Therefore, special acquisition of both ECG traces and glucose data was necessary in our case. Our data was less than what researchers working on online databases would have available. Moreover, online databases often contain manual annotations on the ECG cycles from more than one clinical expert, which were not available in our case.

The glucose sensor had a limitation which was that the minimum value it could record was 2.2 mmol/lt (Section 3.2.2). There were cases where the glucose was falling below this value but it was only recorded as 2.2 mmol/lt. In many cases, interesting dynamics of the glucose variable may have been lost because of this.

Another limitation of MiniMed CGMS was that it was recording glucose in the subcutaneous tissue (interstitial fluid) while the standard approach is to record glucose in the blood stream. Recording glucose in the subcutaneous tissue is not optimal in terms of accuracy since the readings may differ from those of blood glucose. Moreover there is a delay, of approximately 10 minutes, between subcutaneous tissue glucose and blood glucose. The MiniMed software is programmed to automatically correct this

delay but the effectiveness of this correction could be an issue to be investigated and could be affecting the studies carried out using MiniMed. The existence of a delay between interstitial fluid glucose and blood glucose may have affected to some extent our studies although the impact on our results cannot be quantified easily.

Due to processing power, memory size and battery life considerations the HOME system was only capturing the YY' orthogonal lead from the 3-lead ECG. Although the most prominent ECG changes due to hypoglycaemia are expected to be manifested on the YY' lead, recording of more leads would be useful. This is mainly because the dispersion of ECG features across ECG leads is a useful feature and can provide extra information. A popular dispersion feature is the QT dispersion (QTd) that was found to be informative in the detection of the Long QT syndrome [Benhorin 1990].

Due to the memory considerations in the HOME system, the ECG was only recorded every 15 minutes although glucose readings were available every 5 minutes and this obstructed the full use of all the glucose data available. Moreover, having more frequent ECG records during the course of the night would probably have improved the performance of the Knowledge-Based monitoring system. Hypoglycaemic events were in many cases detected by the system a few samples later or earlier than the actual onset. Having more frequent ECG recordings during the night would have reduced the time-deviation of alarms from the actual hypo onset. Nowadays, new trends in hardware technology allow for more storage at low cost which is promising for acquisition of more data and making better use of each patient recruited. Such hardware considerations will be discussed in the future work section of this chapter.

7.1.2 Cardiac function

Delay between changing glucose and cardiac function

This research investigates the relationship between cardiac function and abnormally low glucose levels in Type 1 diabetic patients. Assuming that such a relationship exists, it is suspected that there will be a delay between the changing blood glucose and the subsequent effect on the cardiac function. The glucose will have to drop to abnormally low levels for a period of time before the cardiac function will be affected. This delay is undefined [PD1, PD2] and for the studies in this research it was assumed to be zero. For

the KBS which incorporated the temporal dimension, the delay can be reflected on the time-deviation between alerts given and the actual onset of hypoglycaemia.

The delay may depend on the rate of change of the glucose [PD2] which complicates the relationship between glucose and cardiac function even further. For sudden glucose changes there may be a longer delay, while there may be a shorter one for smooth changes. The slower the change in glucose, the faster the corresponding manifestation of changing glucose on the cardiac function will be, since the heart will be affected sooner by slow changing glucose [PD2]. Faster changes in glucose will take longer to fully manifest on the cardiac function.

Transient cardiac changes in response to hypoglycaemia

Another issue that complicates the research question is that the changes in cardiac function that have been observed in response to hypoglycaemia do not reach a steady state for the duration of hypoglycaemia. In most cases, transient changes are observed. For instance, the value of a certain ECG feature might change after the onset of hypoglycaemia and later on this feature will recover its initial approximate value although the hypoglycaemia will still be present. This is illustrated in Figure 7.1.

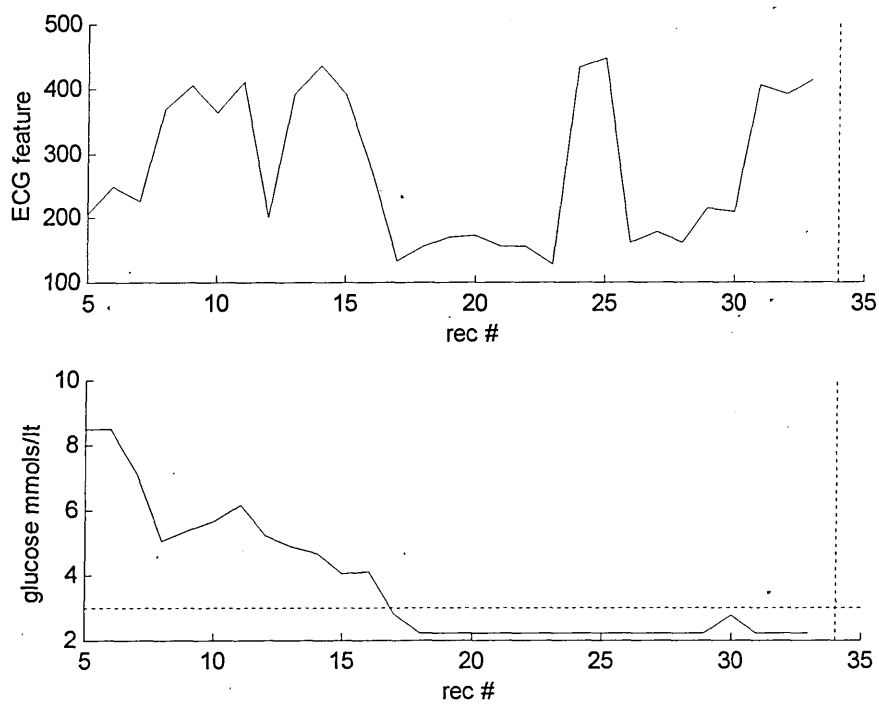


Figure 7.1: T wave amplitude over time (top graph) & Glucose profile over time (bottom graph)

The figure shows a significant feature change for records 17-23 followed by a recovery in feature magnitude although the glucose level remains at abnormally low values. The

feature magnitude drops again at record 26 followed by a further recovery in its magnitude towards the end of the recording period.

7.1.3 Reduced counterregulatory responses in subsequent hypoglycaemic events

It has been clearly established [Davis 1991, 1992], [Heller 1991] that repeated exposure to hypoglycaemia can reduce counterregulatory responses to subsequent hypoglycaemia by as much as 50% in healthy humans and Type 1 diabetic subjects [Davis 2000], [Cryer 1992]. Table 7.1 from [George 1995] quantitatively presents four different bodily responses to experimental hypoglycaemia obtained during three hypoglycaemic events on days 1, 3 and 8 of the study.

Table 7.1: Quantitative representation of bodily responses on subsequent hypoglycaemic events

	Adrenaline	Sweat	Tremor	Symptoms
Day 1	++++++	+++++	++++	++++++
Day 3	+++	++	++++	+++
Day 8	++++	+++	++++	+++++

The study involved 8 non-diabetic subjects over 8 days. Focusing on the adrenaline response to hypoglycaemia, which is the most related to cardiac function, it is clear that the response is reduced on the second hypoglycaemic event (Day 3) while it has partly recovered on the third hypoglycaemic event of Day 8. The study concluded that the physiological responses to hypoglycaemia are affected differentially by antecedent hypoglycaemia with sweating and adrenaline responses remaining impaired for at least 5 days [George 1995].

The above characteristic of the human body may have degraded the accuracy of the static pattern classifiers used in this study and also the performance of the Knowledge–Based monitoring system. Future monitoring systems for hypoglycaemia detection should be calibrated according to the frequency of hypoglycaemic events the patient is experiencing. The occurrence of a hypo event should be stored in memory and taken into account as part of the detection of subsequent onsets of hypoglycaemia. For instance the sensitivity of the system to feature changes should be altered depending on how many hypo events have occurred in the past days. After a hypoglycaemia-free period has elapsed, e.g. 5 days according to [George 1995], then the system would be reset to its default sensitivity towards ECG feature changes.

7.1.4 Feature extraction

One of the great challenges in ECG feature extraction is identifying robustly the end of the T wave either manually or automatically. The current gold standard is manual annotation by a human expert. However manual annotation can still be an ambiguous task under non-standard signal conditions. A number of automatic algorithms have been developed for marking the T wave end but no automatic algorithm can yet be accepted as a gold standard. This research focused on T wave morphology analysis and marking the T wave end was an essential task in the process. The problems in the inexistence of a robust algorithm affected this research. Time constraints did not allow the research to focus on the design of a robust T wave end algorithm so existing algorithms had to be used. The problem of marking the T-end robustly under all signal conditions and confidently undertaking the annotation tasks that clinical experts still carry out remains unresolved and subject to further investigation.

Manual T end annotation and the use of existing T end annotation algorithms is sufficient for many feature extraction tasks. However, in our case very accurate annotation of the T end is required so that even the most subtle QT variations can be uncovered and hence the existence of an accurate algorithm is essential.

Robust detection of the Q point is another problem to be overcome. This point is often masked in noise. In this research this was overcome indirectly by using the R point instead of Q. The QT interval is one of the main ECG features used and, since the R point was used to replace Q, the RT interval was considered instead of the QT. The RT interval is sufficient for this research since it still describes the process of ventricular repolarisation. Although the QT feature has been used traditionally, the RT interval is also encountered in the literature [Porta 1994, 1998].

7.1.5 Heart-rate-correction

Heart rate correction is another area requiring further investigation. The QT interval is a subsection of the RR interval (the instantaneous heart rate) and hence the two are correlated. In hypoglycaemic studies, among other studies, the changes in the QT interval must be examined in isolation to changes in heart rate. Prolongations in QT are significant but they have to be genuine and not due to a prolongation in RR. Therefore there is a need to decorrelate the RR and QT intervals and produce a heart-rate-corrected QT interval (QTc). A few approaches [Puddu 1988, Rautaharju 1993, Ahnve

1985] have been proposed and the most commonly used is Bazett's formula [Bazett 1920] which is also used in this research as mentioned earlier. None of the correction algorithms are 100% accurate and the QTc is found to still be correlated to the heart rate. Low quality of correction can be a limiting factor in the effective use of the QT (and RT) interval. Advancements in the field of hear-rate-correction will be constructive for our research as well as for the wider area of ECG signal analysis.

7.1.6 Diurnal pattern of QTc

The QT and QTc intervals undergo dynamic changes over 24 hours which is another issue to be taken into account when attempting to assess QTc prolongation. The QT varies, not only with heart rate, but also with gender [Lepeschkin 1951] and with time of day [Sarma 1990], [Ong 1993]. Molnar et al have concluded that there is a distinct transient increase in QTc during the first hour after awakening, when the longest hourly mean QTc of the day occurs [Molnar 1996]. Concluding their paper [Molnar 1996 pp82], they also suggest that: "caution should be exercised when categorizing a single clinically measured QTc interval as prolonged. The occurrence of long QT intervals in normal subjects underlines the importance of assessing the QT interval within the clinical context."

By considering the above findings it becomes obvious that the diurnal pattern of QTc is another characteristic that needs to be addressed when designing monitoring systems for the detection of hypoglycaemia. Treating the QTc interval as a static variable that would only change due to hypoglycaemia or other clinical conditions is not a sufficient approach. The healthy diurnal variation of QTc must be taken into account.

7.2 Summary of Achievements

The contribution to knowledge from this research was a detailed analysis of the ECG signal for further examination of the relationship between cardiac function and spontaneous hypoglycaemia and also for the detection of the latter indirectly through the ECG. The main achievement of the thesis was the demonstration of the relationship between spontaneous hypoglycaemia and cardiac function in Type 1 diabetic patients. This was demonstrated mainly by detection of the onset of hypoglycaemia using solely information from the ECG signal. When assessing alarms as correct only if they were raised at the exact time of hypo-onset the optimal configuration of the KBS yielded the following monitoring results: 78.13%, 44.44%, 91.30% for accuracy, sensitivity and

specificity respectively (Section 6.12.5). However if alarms raised within 4 samples from the hypo-onset were classed as acceptable then the metrics reached 93.75%, 100% and 91.30% for accuracy, sensitivity and specificity respectively (Section 6.12.5). These results were achieved after fine-tuning the KBS for the needs of each patient as stressed in chapter 6.

A number of approaches for ECG representation and classification were examined. In more detail:

A comparative study of geometric methods for marking the T wave end was carried out using data from Type 1 diabetic patients (Section 4.4.2). It was concluded that among all the algorithms studied, the tangent method was the one that correlated the most with the annotations from the clinical expert. Based on this finding, the tangent method was chosen for annotation of the T end.

A number of ECG features were extracted and assessed (Section 4.5). Two novel ECG features, inspired from the third and fourth central moments of statistical theory, were introduced for the evaluation of T wave symmetry and morphology. The concept behind an existing feature was modified accordingly to produce a third feature for assessing T wave symmetry. The new feature was based on the ratio of the two areas under the T wave to the left and right of the T peak. Two-tailed t-tests (Section 4.5.1) run separately for each patient, indicated that the above three features assessing T wave morphology underwent statistically significant changes between the conditions of euglycaemia and hypoglycaemia in some patients.

AutoRegressive (AR) modelling was employed for characterisation of post-QRS ECG segments (Section 4.6). The use of AR coefficients was investigated for modelling the ECG segment of interest (from the R peak to the end of the cycle). Comparisons revealed that the use of the AR approach for ECG characterization gave similar results to the use of ECG features.

The Multi-Layer Perceptron (MLP) Neural Network was assessed for the classification of the extracted ECG features in a number of studies (Chapter 5). The MLP performance was compared against the Mahalanobis classifier (Linear Discriminant Analysis). The k-Nearest Neighbour (kNN) classifier was also assessed (Section 5.2.6).

It was found that the MLPs had slightly better performance than LDA. However, a significant advantage of using statistical classifiers is that the user does not need to involve in the laborious process of tuning a neural network architecture for optimal performance. Moreover, the processing power needed for the use of the aforementioned statistical classifiers was, in our case, significantly less compared to that required for training neural networks.

A methodology was proposed according to which a diagnostic system can be implemented for hypoglycaemia monitoring. This consists of an ECG representation stage in cascade with a classification stage as was presented in Chapter 4. Following the proposed methodology, a Knowledge-Based System (KBS) was designed for monitoring offline data from diabetic patients in a manner that simulated an online patient-monitoring scenario (Chapter 6). Two versions of the system were produced namely an "expert system" and a system based on fuzzy logic. Employing a KBS yielded the highest performance among all the techniques used. This system was able to monitor correctly, patients that were consistent with the initial hypothesis i.e. exhibiting QT prolongation and T wave flattening during hypoglycaemia. A significant difference of the system compared to the neural and statistical classifiers was that it incorporated temporal information, while the latter were performing static pattern classification. Through the above system, it was demonstrated that two ECG features were sufficient for detection of hypoglycaemia in those patients that manifested both QT prolongation and T wave flattening in response to abnormally low blood glucose.

Moreover, the above system contributed in formulating the vague knowledge of the principal ECG changes under hypoglycaemia in the form of rules of natural language. This was informative for medical researchers and provided feedback to the clinical experts who formulated the initial hypothesis and contributed the initial guidelines for the knowledge-base.

An interesting result was that the diagnostic performance reached in this research work was achieved by using only one ECG lead. The Y lead from the 3-lead ECG was used since it is placed in the direction of maximal flow of current through the heart. Using more information from more leads is expected to increase the performance as it will be highlighted in the discussion of directions for future research (Section 7.4).

Finally, data analysis on the available datasets highlighted the existence of high inter-patient variability (Section 4.5). This suggests that the performance of a monitoring system would be boosted by allowing customisation to the specific patient to be monitored. Differences in ECG behaviour were also observed among different night-recordings of the same patient which suggests that a robust monitoring system should also be made adaptive to ECG changes as time elapsed. Moreover it should be able to start monitoring from a number of different initial conditions, i.e. different ECG signature at the start of the night, without this affecting its performance.

Hypothesis testing (Student's t-test) on all ECG features extracted, proposed that apart from the inter-patient variability, different features may be robust predictors of hypoglycaemia among different patients (Section 4.5.1). The condition of hypoglycaemia may not be sufficiently manifested using the same features on all patients.

Investigation of the ECG and glucose profiles also indicated that the ECG responses to hypoglycaemia are expressed in the form of transient events (Sections 5.2.6 and 6.14.1) and hence, the incorporation of a temporal dimension in a monitoring system will be essential for robust detection of the condition.

7.3 Addressing the Research Question

The main conclusion of this thesis is that there exists a relationship between spontaneous hypoglycaemia and cardiac function in some patients. This is in line with the observations and conclusion of a clinical study focusing on QTc interval changes in response to spontaneous hypoglycaemia [Robinson 2004]. In our work it is also demonstrated that hypoglycaemia is reflected on other ECG features studied, besides the QTc and that the onset of hypoglycaemia can in many cases be detected automatically through analysis of the ECG.

This above relationship was initially established in studies of experimental, i.e. artificially induced, hypoglycaemia. The fact that the same relationship holds, although more subtly, in the case of spontaneous hypoglycaemia is a significant result with possible impact towards the investigation of unexplained nocturnal deaths of young diabetic patients.

Besides the strengthening of the research hypothesis, we conclude that we can detect the onset of hypoglycaemia, subject to a time-deviation, in those patients that exhibit T wave flattening and QT prolongation. Detection of the condition with a time-deviation of up to one hour[45] from its onset is perfectly legitimate [PD1]. This is an important result since it is achieved purely by using information from the ECG. Clinical studies have focussed on the assessment of the cardiac changes due to hypoglycaemia and also on the correct discrimination between patients with the Long QT syndrome and healthy control subjects. However, this to the best of our knowledge is the first time that detection of the symptomatic status of hypoglycaemia is attempted purely by analysis of the ECG.

Regarding the design of a patient monitoring system for hypoglycaemia, we have identified and proposed that such a system must be adaptive as time elapses and tailored to the patient to be monitored. Adaptivity is necessary to overcome day-to-day intra-patient variability (due to QTc diurnal pattern, frequency of hypos affecting responses to subsequent hypos etc). Customisation per patient is necessary to overcome problems due to inter-patient variability (variations in gender, age, fitness level, duration of diabetes, level of glycaemic control and so on). Customisation per patient may involve the selection of the ECG features to be used since there are indications that different features may by related to hypo for different patients. Ideally feature selection should be automated and be done online and adaptively rather than a priori since it is likely that features manifesting the onset of hypo could in some cases be affected by the day-to-day variability of the patient.

7.4 Further Work

This section proposes a few directions for further work that were either not materialised in this research program or they comprise a logical continuation of the work carried out in this thesis.

A sensible step in further work would be to use more features in the knowledge-based monitoring system. This would allow the validation of genuine ECG changes due to hypoglycaemia, leading to suppression of some false alarms and detection of hypo events subject to a shorter time-deviation. Moreover it would aid detection of hypo

[45] i.e. 4 samples in the current dataset.

events associated with very subtle ECG changes that may not be sufficiently reflected on the QTc and T amplitude features, i.e. in the cases where hypoglycaemia is almost asymptomatic. Features assessing the symmetry and morphology of the T wave, developed in this research, will be useful additions to the system. The RR interval (instantaneous heart rate) and also features characterising U waves and ST segment changes would be very informative. Generally, improvements in the field of ECG signal processing, involving both ECG annotation and heart-rate correction algorithms, are expected to positively influence our research.

7.4.1 Action Potential modelling

Modelling of the T and U waves using Action Potentials is a powerful approach for ECG characterisation and may prove promising for investigation of the manifestation of hypoglycaemia on the ECG. A number of studies have been carried out on Action Potential modelling of ECGs [Wohlfart 1987, Malik 1989, Padrini 1995, Vila 2000]. Action potentials could be used for modelling of the ECG segment of interest and then traditional feature extraction could be carried out on the modelled signal, which is less noisy than the original one. Alternatively the model parameters describing the T and U waves could be investigated to identify whether they differ between the conditions of euglycaemia and hypoglycaemia. The latter is particularly interesting as a direction for future work. Results from the Autoregressive modelling of ECG traces, presented in this thesis, indicated that the modelling of whole segments of the ECG cycle is a promising approach. Action Potential modelling has the advantage, over AR modelling, that it is a biologically plausible approach and hence further improvement in performance is anticipated.

7.4.2 Beat-to-beat ECG analysis

Signal-averaged ECG (SAECG) cycles were exclusively used in this research work. The investigation of the raw (beat-to-beat) ECG cycles could be a direction for further work. Such an investigation would look for significant events due to hypoglycaemia than were not reflected on the SAECG cycles as they may have been filtered out through the averaging process. The challenge of using beat-to-beat cycles is that the level of noise is very high and any type of processing, such as annotation and feature extraction is extremely difficult and requires very robust algorithms.

7.4.3 Using more ECG leads

The diagnostic performance reached in this research was achieved by using only one ECG lead. As stressed earlier, the Y lead is the best choice when using a single-lead system. Incorporating more leads from the standard 12-lead ECG is expected to improve performance and will allow the use of extra ECG features based on the dispersion across the ECG leads of the features already used (e.g. QTc dispersion). Moreover, the availability of more leads allows the application of Blind Source Separation (BSS) techniques for reducing the level of noise and artefacts that contaminate the ECG. Application of BSS techniques for noise reduction could lead to the reduction of the level of Signal-Averaging applied. This is desirable since the raw signal is more informative and because the Signal-Averaging process can introduce distortion. In addition, application of BSS for reduction of motion artefacts from the skeletal muscle could possibly allow use of the monitoring system while the patient would be awake. The current system is not addressing the monitoring of patients while they are awake. There was no data available to perform such testing but performance is expected to be lowered due to the increased amount of muscular activity in the skeletal area.

7.4.4 Baseline wandering

Baseline wandering on the beat-to-beat ECG signals has been observed in our dataset. A question is raised as to whether the existence of baseline wandering conveys a message related to the existence of hypoglycaemia. The baseline wandering was only investigated by visual inspection and it was hard to draw any conclusions as to whether it is a predictor of hypoglycaemia. An interesting direction for further work would be to devise algorithms that can quantify the level of baseline wandering and then investigate the importance of wandering to the detection of hypoglycaemia.

7.4.5 Datasets

The current dataset constitutes the current state of the art in the studies of the relationship between hypoglycaemia and cardiac function. To the best of our knowledge this lies among the first datasets using continuous glucose monitoring, besides ECG recordings, in the studies of nocturnal hypoglycaemia. However, acquisition of more data will be a significant enhancement to the study.

ECG traces recorded more frequently during the night will probably increase the chances of the current system in giving a hypoglycaemic alert within shorter time from the actual onset. In the current dataset, an alarm raised with one sample deviation corresponds to 15 minutes in time but this is only due to the time-resolution of the dataset. Moreover, data capture from more leads is very informative as stressed earlier. Current trends in glucose sensing technologies allow increased range of glucose values recorded. The GlucoDay continuous glucose monitoring system by Menarini Diagnostics [URL 16] allows recording of glucose levels down to ≅1.1 mmol/lt which is very informative for investigation of severe hypoglycaemic events. The glucose meter used in this study could not record glucose concentrations below 2.2 mmol/lt (Section 3.2.2) and may have masked the severity of some hypoglycaemic events detected.

More data will be necessary for further testing of the existing monitoring system and also in the process of upgrading and improving it. More data will also be essential when embarking on the research direction of using time-lagged neural network architectures for patient monitoring. Such architectures require extensive datasets. Using time-lagged neural networks is promising for detecting transient changes that we may have not been able to observe and formulate in the rule-base produced.

Upon satisfactory performance, the functionality of the system can be extended so that it can be attached to diabetic patients and perform online monitoring. Extensive testing will also be necessary followed by the necessary modifications before a monitoring system can be approved for patient use.

Regarding the existing ECG acquisition hardware (HOME system) it could be improved by upgrading the battery, internal memory etc, because at the moment not all the glucose data available per patient recruited is used[46]. Current technology can also offer new devices with increased capability in affordable prices. Modern palmtop computers with 12-lead ECG interfaces are available in the market. Current technology significantly facilitates acquisition at the patient's own environment. When the HOME

[46] A MiniMed CGMS probe used for subcutaneous glucose measurements lasts for at least 3 days. However, each acquisition of the dataset was including only two nights worth of ECG which means that not all the glucose data available could be used. Moreover, the glucose variable was sampled every 5 minutes while ECG was captured every 15 minutes. Because of this not all the glucose samples during a given night were used.

system was implemented the ECG interface attached to it had to be custom-designed for the needs of the research while today it can be bought off-the-shelf and has significantly reduced size. Such devices were presented in Section 2.5.2.

7.5 Final Remarks

This thesis focused on the analysis and interpretation of SAECG signals for the detection of spontaneous nocturnal hypoglycaemia. A number of approaches for ECG representation were employed and some novel ECG features were introduced. The modelled ECGs were classified according to their corresponding glucose levels. Besides the approaches used for static ECG pattern classification (ANN, LDA and kNN), a KBS was designed to perform patient monitoring. It was developed and tested on offline data in a manner that simulated an online monitoring scenario. The KBS was rule-based and the knowledge-base was formulated within guidelines from clinical experts.

To conclude the chapter it is stressed once more that analysis of the ECG is a promising approach for the detection of the symptomatic status of hypoglycaemia in Type 1 diabetic patients. For a few patients studied in this work, hypoglycaemia could be detected at the exact sample it occurred during the night only by looking at the patient's ECG. This is an important result and provides new insights on the effects of hypoglycaemia on the cardiac function.

Further improving the performance of the monitoring system may lead to a commercial alarm system in the long term future. Such a system could be used for nocturnal hypoglycaemia detection or for the detection of Sudden-Death-related cardiac arrhythmias. The software engine behind such a monitoring system could be used in any clinical situation where ECG is already captured and could be incorporated in commercial ECG monitors.

References

(Citations in the body of the thesis and in the following list are given using the first author name and the publication date. If more than one publication date is given in a citation then more than one paper from the same author is cited. For organisations and when an author name is not known, the acronym of the organisation name is used. Internet references are cited using the acronym "URL" followed by a serial number. Personal discussions with field experts are cited using the acronym "PD" followed by a serial number.)

ADA 1995	American Diabetes Association, (1995), "Self-monitoring of blood glucose", Diabetes Care, vol 18, pp 408-413.
Ahnve 1985	Ahnve S, (1985), "Correction of the QT interval for heart rate: review of different formulas and the use of Bazett's formula in myocardial infarction" Am Heart J, vol 109, pp 568-74.
Aikens 1983	Aikens J S, (1983), "PUFF: An expert system for interpretation of pulmonary function data", Computers and Biomedical Research, vol 16, pp 199-208.
Al-Fahoum 1999	Al-Fahoum A S, Howitt I, (1999), "Combined wavelet transformation and radial basis neural networks for classifying life-threatening cardiac arrhythmias", Med. Biol. Eng. Comput., vol 37, pp 566–573.
Alperin 1986	Alperin N, Sadeh D, (1986), "An improved method for on-line averaging and detecting of ECG waveforms", Computers Biomed Res, vol 19, pp 193–202.
AMA 1988	American Medical Association, (1988), Arch Intern Med, vol 148, pp 1859-1863.
Balda 1977	Balda R A, Diller G, Deardorff E, Doue J, Hsieh P, (1977) "The HP ECG analysis program" Trends in Computer Processed Electrocardiograms, pp 197-205
Bazett 1920	Bazett H C (1920), "An analysis of time-relations of the electrocardiogram", Heart 1920;7;353-70
Benhorin 1990	Benhorin J, Merri M, Alberti M, Locati E, Moss A J, Hall W J, Cui L (1990), "Long QT syndrome. New electrocardiographic characteristics" Circulation, vol 82, pp 521 - 527.
Berbari 2000	Berbari E J, Steinberg J S, (2000), "A Practical Guide to the Use of the High-Resolution Electrocardiogram" Futura Publishing Company, Inc. Armonk, NY.
Bernstein 2002	Bernstein R K, (2002), "Diabetes solutions. A complete guide to achieving normal blood sugars". Available from http://www.diabetes-normalsugars.com/ articles/meter_history.pdf - Last accessed 12/10/2004

Bland 1986 — Bland J M, Altman D G, (1986), "Statistical methods for assessing agreement between two methods of clinical measurement", The Lancet, vol 8

Buchanan 1983 — Buchanan B G, Barstow D, Bechtel R, Bennet J, Clancey W, Kulikowski C, Mitchell T M, Waterman D A, (1983), "Constructing an Expert System. in Building Expert Systems", F. Hayes-Roth, D. Waterman, and D. Lenat, eds., Addison-Wesley, Reading, Mass.

Bystricky 2002 — Bystricky W and Safer A, (2002), "Modelling T-end in Holter ECGs by 2-Layer Perceptrons", Computers in Cardiology, vol 29, pp 105-108

Campbell 1991 — Campbell I W, (1991), "Dead in bed syndrome: a new manifestation of nocturnal hypoglycaemia?" Diabetic Med, vol 8, pp 3-4.

Cheyne 2002 — Cheyne E H, Cavan D A, Kerr D (2002), "Performance of a Continuous Glucose Monitoring System During Controlled Hypoglycaemia in Healthy Volunteers", Diabetes Technology and Therapeutics, vol 4, no 5, pp 607-613.

Cryer 1992 — Cryer P E, (1992), "Iatrogenic hypoglycemia as a cause of hypoglycemia-associated autonomic failure in IDDM", Diabetes vol 41, pp 255-260.

Daskalov 1999 — Daskalov I K, Christov I I, (1999), "Automatic detection of the electrocardiogram T-wave end" Med. Biol. Eng. Comput., vol 37, pp 348–53.

Davis 1998 — Davis E A, Keating B, Byrne G C, Russell M, Jones T W (1998), "Impact of improved glycaemic control on rates of hypoglycaemia in insulin dependent diabetes mellitus", Arch Dis Child, vol 78, pp 111-115.

Davis 1991 — Davis M R, Shamoon H, (1991), "Counterregulatory adaptation to recurrent hypoglycemia in normal humans" J Clin Endocrinol Metab, vol 73, pp 995-1001.

Davis 1992 — Davis M R, Mellman M J, Shamoon H, (1992), "Further defects in counterregulatory responses induced by recurrent hypoglycemia in IDDM", Diabetes vol 41, pp 1335-1340

Davis 2000 — Davis S N, Mann S, Galassetti P, Neill R A, Tate D, Ertl A C, Costa F, (2000) Effects of differing durations of antecedent hypoglycemia on counterregulatory responses to subsequent hypoglycemia in normal humans" Diabetes, vol 49, pp 1897-1903.

DCCT 1991 — The DCCT Research Group, (1991), "Epidemiology of severe hypoglycaemia in the Diabetes Control and Complications Trial", Am J Med, vol 90, pp 450-459.

DCCT 1993 — The Diabetes Control and Complications Trial Research Group, (1993), "The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus". N. Engl. J. Med, vol 329, pp 977-986.

Feigenbaum 1977 — Feigenbaum E A (1977), "The art of artificial intelligence – Themes and case studies of knowledge engineering", 5th international j oint c onference o n a rtificial i ntelligence (IJCAI), pp.1014-1029.

Fleming 1994 — Fleming D R, (1994), "Accuracy of blood glucose monitoring for patients: what is it and how to achieve it", Diabetes Educator, vol 20, pp 495-500.

Foster 1999 — Foster S A, Goode J V, Small R E, (1999), "Home blood glucose monitoring" The Annals of Pharmacotherapy, vol 33, no 3, pp 355-363.

George 1995 — George E, Harris N, Bedford C, Macdonald I A, Hardisty C, Heller S R (1995), "Prolonged but partial impairment of the hypoglycaemic physiological response following short-term hypoglycaemia in normal subjects", Diabetologia, vol 38, pp 1183-1190.

Gholam-Hosseini 1998 — Gholam-Hosseini H, Nazeran H, (1998), "Detection and extraction of the ECG signal parameters", 20th Annual International IEEE/EMBS Conference, pp127-130, Oct. 29 - Nov. 1, 1998, Hong Kong

Greenhow 1993 — Greenhow S G, Linkens D A, Asbury A J, (1993), "Pilot study of an expert system adviser for controlling general anaesthesia", British Journal of Anaesthesia, vol 71, issue 3, pp 359-365.

Gross 2000 — Gross T M, Bode B W, Einhorn D, Kayne D M, Reed J H, White N H, Mastrototaro J J, (2000), "Performance evaluation of the MiniMed continuous glucose monitoring system during patient home use", Diabetes Technol Ther. Spring, vol 2, no 1, pp 49-56.

Gross and Mastrotrotaro 2000 — Gross T M, Mastrotrotaro J J, (2000), "Efficacy and reliability of the continuous glucose monitoring system", Diabetes Technology and Therapeutics, vol 2, suppl 1, pp 19-26.

Gross and Ter Veer 2000 — Gross T M, Ter Veer A, (2000), "Continuous glucose monitoring in previously unstudied population subgroups" Diabetes Technology and Therapeutics, vol 2, suppl 1, pp 27-34.

Hansen 1993 — Hansen K, Duck S (1993), "Sleep Sentry: Evaluation in paediatric patients for detection of nocturnal hypoglycaemia" Diabet Care, vol 6, pp 597-600.

Harris 1996 — Harris N D Baykouchev S B, Marques J L, Cochrane T, George E, Heller S R, Ward J D, (1996) "A portable system for monitoring physiological responses to hypoglycaemia", Journal of Medical Engineering and Technology, vol 20, no 6, pp 196-202.

Harris 2000 — Harris N D, Ireland R H, Marques J L B, Hudson S, Davies C, Lee S, Robinson R T, Heller S R, (20000 "Can changes in QT interval be used to Predict the Onset of Hypoglycaemia in Type 1 Diabetes?" Computers in Cardiology vol 27, pp375-378.

Haykin 1994 — Haykin S, (1994), "Neural Networks A comprehensive foundation" Macmillan Publishing.

Hedén 1997 — Hedén B, Öhlin H, Rittner R, Edenbrandt L, (1997), "Acute myocardial infarction detected in the 12-lead ECG by artificial neural networks," Circulation, vol. 96, pp. 1798–1802.

Heinemann 1995 — Heinemann L, Meinhold J, Kunze W, (1995), "Hypoglycaemia detection by ECG recording", Diabetes Care, vol 18, pp 139–140.

Heller 1991 — Heller S R, Cryer P E, (1991), "Reduced neuroendocrine and symptomatic responses to subsequent hypoglycemia after one episode of hypoglycemia in nondiabetic humans", Diabetes, vol 40, pp 223-226.

Hoss 2001 — Hoss U, Kalatz B, Gessler R, Pfleiderer H, Andreis E, Rutschmann M, Rinne H, Schoemaker M, Haug C, Fussgaenger R D, (2001) "A Novel Method for Continuous Online Glucose Monitoring in Humans: The Comparative Microdialysis Technique". Diabetes Technology and Therapeutics, vol 3, no 2, pp 237-243.

Ireland 1998 — Ireland R H, Robinson R T C E, Heller S R, Harris N D, (1998) "QT Measurement for TU fused ECG Morphology as exhibited during hypoglycaemia", 20th Annual International IEEE/EMBS Conference, pp 240-242, Oct. 29 - Nov. 1, Hong Kong.

Ireland 2000 — Ireland R H, Robinson R T C E, Heller S R, Marques J L B, Harris N D, (2000), "Measurement of high resolution ECG QT interval during controlled euglycaemia and hypoglycaemia" Physiol. Meas. vol 21 pp 295-303.

Ireland 2001 — Ireland R, "ECGLAB Version 2.06" Instruction Manual, Diabetes Research, Dept of Medical Physics and Clinical Engineering, Royal Hallamshire Hospital, The University of Sheffield, May 2001, (UNPUBLISHED MATERIAL)

Jadeh 1965 — Jadeh L (1965), "Fuzzy Sets" Inf Control, vol 8, pp 338-353.

Jager 1996 — Jager F, Moody G B, Taddei A, Antolic G, Zabukovec M, Skrjanc M, Emdin M, Mark R G (1996), "Development of a Long Term Database for Assessing the Performance of Transient Ischemia Detectors", Computers in Cardiology, vol 23, pp. 481-484.

Jager 1998 — Jager F, Moody G B, Taddei A, Antolic G, Smrdel A, Glavic B, Emdin M, Marchesi C, Mark R G, (1998), "Research Resources for Development and Evaluation of Transient Ischemia Detectors", Proc. Computer-Aided Data Analysis in Medicine (CADAM 98), Informatica Medica Slovenica, vol 5(1,2), pp 45-54, ISSN 1318-2129.

Jager CiC98 — Jager F, Moody G B, Taddei A, Antolic G, Emdin M, Smrdel A, Glavic B, Marchesi C, Mark R G, (1998), "A Long-Term ST Database for Development and Evaluation of Ischemia Detectors", Computers in Cardiology, pp 301-304, IEEE Press. ISSN 0276-6547.

Jager 2000 — Jager F, Taddei A, Emdin M, Antolic G, Dorn R, Moody G B, Glavic B, Smrdel A, Varanini M, Zabukovec M, Bordigiago S, Marchesi C, Mark R G, (2000), "The Long-Term ST Database: A Research Resource for Algorithm Development and Physiologic Studies of Transient Myocardial Ischemia" Computers in Cardiology, vol 27, pp 841-844.

Jager 2003 — Jager F, Taddei A, Moody G B, Emdin M, Antolic G, Dorn R, Smrdel A, Marchesi C, Mark R G, (2003), "Long-term ST database: a reference for the development and evaluation of automated ischaemia detectors and for the study of the dynamics of myocardial ischaemia", Med. Biol. Eng. Comput, vol 41, pp 172–182.

Kenney 1951 — Kenney J F, Keeping E S, (1951), "Mathematics o f S tatistics", Pt. 2, 2nd ed. Princeton, NJ: Van Nostrand.

Kirk 1998 — Kirk J K, Rheney C C, (1998), "Important features of blood glucose meters" Journal of the American Pharmaceutical Association, vol 38, pp 210-219.

Kundu 1993 — Kundu M, Nasipuri M, Basu D K, Bhattacharya A, (1993), "A reasoning system for on-line interpretation of ECG signal", IEEE Region Ten Conference (TENCON '93), vol 2, pp 626-630, Beijing, 19-21 Oct.

Kundu 1994 — Kundu M, (1994), Studies on some AI techniques for interpretation of ECG waveforms, Doctoral Dissertation, Computer Sc and Eng Dept, Jadavpur University, India.

Kundu, Nasipuri 1994 — Kundu M, Nasipuri M, Basu D K, (1994), "A rule-based expert system for real-time interpretation of ECG patterns", Proceedings of International Seminar held on occasion of Ninth Indian Engineering Congress at Calcutta, vol 2, pp. 827-838, 15-20 Dec.

Lagerholm 2000 — Lagerholm M, Peterson C, Braccini G, Edenbrandt L, Sörnmo L, (2000), "Clustering ECG Complexes Using Hermite Functions and Self-Organizing Maps", IEEE Transactions On Biomedical Engineering, vol 47, no 7, pp 838-848.

Laguna 1990 — Laguna P, (1990), "New Electrocardiographic Signal Processing Techniques: Application to Long-term Records", PhD dissertation, Science Faculty, University of Zaragoza.

Laguna, Thakor 1990 — Laguna P, Thakor N V, Caminal P, Jané R, Yoon H R, (1990), "New Algorithm for QT interval Analysis in 24-hour Holter ECG: performance and applications", Medical & Biological Engineering and Computing, vol 28, pp 67-73.

Laguna 1994 — Laguna P, Jané R, Caminal P, (1994), "Automatic Detection of Wave Boundaries in Multilead ECG Signals: Validation with the CSE Database" Computers and Biomedical Research, vol 27, pp 45-60.

Laguna 1997	Laguna P, Mark R G, Goldberger A, Moody G B, (1997), "A Database for Evaluation of Algorithms for Measurement of QT and Other Waveform Intervals in the ECG", Computers in Cardiology, vol 24, pp 673-676.
Lazzara 1992	Lazzara R, (1992), "Cellular basis for the U wave and its translation to the surface ECG", J Electrocardiol, vol 24, pp 44–5.
Lee 2003	Lee S, Harris N D, Robinson R T, Yeoh L, Macdonald I A, Heller S R, (2003), "Effects of adrenaline and potassium on QTc interval and QT dispersion in man", Eur J Clin Invest, vol 33, no 2, pp 93-8.
Lepeschkin 1951	Lepeschkin E, (1951), "Modern Electrocardiography", Baltimore: Williams and Wilkings, vol 1, 180-8
Li 1995	Li C, Zheng C, Tai C, (1995), "Detection of ECG characteristic points using wavelet transforms", IEEE Trans Biomed Eng, vol 42, no 1, pp 21-8.
Linkens 1990	Linkens D A, Hacisalihzade S S, (1990), "Computer control systems and pharmacological drug administration: a survey", J Med Eng Technol, vol 14, no 2, pp 41-54.
Malik 1989	Malik M, Camm A J, (1989), "Computer model of cardiac repolarisation processes and of the recovery sequence", Comput Biomed Res, vol 22, pp 160-180.
Marques 1994	Marques J L B, (1994), "High-resolution Electrocardiogram analysis", PhD thesis, University of Sheffield.
Marques 1997	Marques J L B, George E, Peacy S R, Harris N D, MacDonald I A M, Cochrane T, Heller S R, (1997), "Altered ventricular repolarisation during hypoglycaemia in patients with diabetes", Diabetic Medicine, vol 14, pp 648-654.
MathWorks Fuzzy	MathWorks Inc, (1999), Fuzzy Logic Toolbox for Use with MATLAB.
MathWorks Statistics	MathWorks Inc, (1999), Statistics Toolbox for Use with MATLAB.
McCall 1987	McCall A L, Mullin C J, (1987), "Home Blood Glucose Monitoring: Keystone for Modern Diabetes Care", Medical Clinics of North America, vol 71, pp 763-787.
McCulloch 1943	McCulloch W S, Pitts W, (1943), "A logical calculus of ideas immanent in nervous activity", Bulletin of Mathematical Biophysics, vol 5, pp 115-133.
McLachlan 1981	McLachlan E M, (1981), "Fundamentals of electrocardiography", Oxford University Press.
McLaughlin 1995	McLaughlin N B, Campbell R W F, Murray A, (1995), "Comparison of automatic QT measurement techniques in the normal 12 lead electrocardiogram", Br Heart J, vol 74, pp 84-89.

McLaughlin 1996 McLaughlin N B, Campbell R W, Murray A, (1996), "Accuracy of four automatic QT measurement techniques in cardiac patients and healthy subjects", Heart, vol 76, pp 422-426.

Mehta 2002 Mehta M, Vincze G, Lopez D A, (2002), " Emerging Technologies in Diabetes Care", U.S. Pharmacist, vol 27, pp 11.

Minsky 1969 Minsky M, Papert S, (1969), "Perceptrons: an introduction to computational geometry", Published MIT Press.

Minsky 1988 Minsky M, Papert S, (1988), "Perceptrons - Expanded Edition: An Introduction to Computational Geometry", Published MIT Press.

Molnar 1996 Molnar J, Zhang F, Ehlert F A, Rosenthal J E, (1996), "Diurnal pattern of Qtc: How long is prolonged? Possible relationship to circadian triggers of cardiovascular events", J Amer College Cardiol, vol 27, pp 76-83.

Monsod 2002 Monsod T P, Flanagan D E, Rife F, Saenz R, Caprio S, Sherwin R S, Tamborlane W V, (2002), "Do sensor glucose levels accurately predict plasma glucose concentrations during hypoglycemia and hyperinsulinemia?", Diabetes Care, vol 25, no 5, pp 889-893.

Moody 1982 Moody G B, Mark R G, (1982), "Development and evaluation of a 2-lead ECG analysis program", Computers in Cardiology, IEEE Computer Society Press, 1982, pp 39-44.

Morio 1989 Morio S, Kawahara S, Okamoto N, Suzuki T, Okamoto T, Harada M, Shimizu A, (1989), "An expert system for early detection of cancer of the breast", Comput Biol Med, vol 19, no 5, pp 295-305.

Mylopoulos 1983 Mylopoulos J, Shibahara T, Tsotsos J K, (1983), "Building knowledge-based systems: the PSN experience", IEEE Comput, vol 16, no 10, pp 83-88.

Ong 1993 Ong J J, Sarma J S, Venkataraman K, Levin S R, Singh B N, (1993), "Circadian rhythmicity of heart rate and QTc interval in diabetic autonomic neuropathy: implications for the mechanism of sudden death", Am Heart J, vol 125, no 3, pp 744-52.

Padrini 1995 Padrini R, Butrous G, Camm J, Malik M, (1995), "Algebraic decomposition of the TU wave morphology patterns", PACE vol 18, pp 2209-2215.

Pan 1985 Pan J, Tompkins W J, (1985), "A Real-Time QRS Detection Algorithm", IEEE Transactions on Biomedical Engineering, vol 32, no 3, pp 230-236.

Patterson 1990 Patterson D, (1990), "Introduction to Artificial Intelligence and Expert Systems", Prentice-Hall.

Paul 1997 Paul K, (1997), "A connectionist approach to ECG interpretation", M C S E thesis, Comp. Sc. and Eng. Dept. Jadavpur University, Calcutta, India.

PD1 Personal Discussions with Dr Simon R Heller.

PD2 Personal Discussion with Dr Nigel D Harris.

Pople 1975 Pople HE, Myers JD, Miller RA, (1975), "DIALOG (INTERNIST): A Model of Diagnostic Logic for Internal Medicine", Fourth International Joint Conference on Artificial Intelligence. pp 848-855, Tbilisi, Georgia, USSR, September 1975.

Porta 1994 Porta A, Lombardi F, Benedetti M, Sandrone G, Baselli G, Malliani A, Cerutti S, (1994), "Reliability of the measurement of RT variability", Computers in Cardiology, pp 217 – 220.

Porta 1998 Porta A, Baselli G, Lombardi F, Cerutti S, Antolini R, Del Greco M, Ravelli F, Nollo G, (1998), "Performance assessment of standard algorithms for dynamic R-T interval measurement: comparison between R-Tapex and R-T(end) approach", Medical and Biological Engineering and Computing, vol 36, pp 35-42.

Puddu 1988 Puddu P E, Jouve R, Mariotti S, (1998), "Evaluation of 10 QT prediction formulas in 881 middle-aged men from the seven countries study: emphasis on the cubic root Fridericia's equation", J Electrocardiol, vol 21, pp219-229.

Rao 1995 Rao G, Guy R H, Glikfeld P, LaCourse W R, Leung L, Tamada J, Potts R O, Azimi N, (1995), "Reverse Iontophoresis: Noninvasive Glucose Monitoring in vivo In Humans", Pharmaceutical Research, vol 12, pp 1322-1326.

Rautaharju 1993 Rautaharju P M, Zhou S H, Wong S, Prineas R, Berenson G S, (1993), "Functional characteristics of QT prediction formulas. The concepts of QTmax and QT rate sensitivity", Comput Biomed Res, vol 26, pp 188-204.

Rebrin 1998 Rebrin K, Steil G M, Van Antwerp W P, Mastrototaro J J, (1998). "Reliability of interstitial glucose measurements by subcutaneous glucose sensors in humans" Diabetologia, vol 41, suppl 1, pp 880.

Rebrin 2000 Rebrin K, Steil G M, (2000), "Can Interstitial Glucose Assessment Replace B lood Glucose Measurements?", Diabetes Technology and Therapeutics, vol 2, no 3, pp 461-472.

Robinson 2004 Robinson R T C E, Harris N D, Ireland R H, Macdonald IA, Heller SR, (2004), "Changes in cardiac repolarization during clinical episodes of nocturnal hypoglycaemia in adults with Type 1 diabetes", Diabetologia, vol 47, pp 312-315.

Rosenblatt 1962 Rosenblatt F, (1962), "Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms", Washington DC, Spartan. Books.

Rumelhart 1986 Rumelhart D E, McClelland J L, and the PDP Research Group (1986), "Parallel Distributed Processing: Explorations in the Microstructure of Cognition", vol 1 and 2, Cambridge MA, MIT Press.

Sarma 1990 — Sarma J S, Venkataraman K, Nicod P, Polikar R, Smith J, Schoenbaum M P, Singh B N, (1990), "Circadian rhythmicity of rate-normalized QT interval in hypothyroidism and its significance for development of class III antiarrhythmic agents", Am J Cardiol, vol 66, no 12, pp 959-63.

Saxena 2002 — Saxena S C, Kumar V, Hamde S T, (2002), "QRS detection using new wavelets", J Med Eng Technol, 2002, vol 26, no 1, pp 7-15.

Sharp 2001 — Sharp P S, Krishna S, Rainbow S, (2001), "Glycated haemoglobin and continuous glucose monitoring", Diabetologia, vol 44, suppl 1, pp 913.

Shibahara 1983 — Shibahara T, Tsotsos J, Mylopoulos J, Covvey H D, (1983), "CAA: A Knowledge Based System using Causal Knowledge to Diagnose Cardiac Rhythm Disorders", International Joint Conference on Artificial Intelligence, p231 – 237, Karlsruhe, Germany, Aug. 1983.

Shin 2002 — Shin J J, Dangui N D, Kanderian S Jr, Gross T M, Mastrototaro J J, (2002), "A new retrospective calibration algorithm (New-RA) for continuous glucose monitoring provides more complete and more accurate CGMS data", Diabetes vol 51, suppl 2, pp 843.

Simon 1997 — Simon B J, Eswaran C, (1997), "An ECG classifier designed using modified decision based neural networks", Computers and Biomedical Research, vol 30, pp 257-272.

Simpson 1990 — Simpson P K, (1990), "Artificial Neural Systems: Foundations, Paradigms, Applications, and Implementations", New York, Pergamon Press.

Srinivasan 2002 — Srinivasan N, Ge D F, Krishnan S M, (2002), "Autoregressive modelling and classification of cardiac arrhythmias", Second Joint EMBS/BMES conference, pp 1405-1406, Oct 23-26, Houston TX, USA.

Steil 2000 — Steil G M, Bernaba B, Saad M, Mastrototaro J, Rebrin K, (2000), "Accurate determination of plasma glucose during hyper- and hypoglycemia with a subcutaneous glucose sensor", Diabetes, vol 49, suppl 1, p510.

Stockman 1976 — Stockman G, Kanal L, Kyle M C, (1976), "Structural pattern recognition of carotid pulse waves using a general waveform parsing system", CACM, vol 19, no 12, pp 688-695.

Taddei 1989 — Taddei A, Biagini A, Distante G, Marchesi C, Mazzei MG, Pisani P, Roggero N, Zeelenberg C, (1989), "An annotated database aimed at performance evaluation of algorithms for ST-T change analysis", Computers in Cardiology vol 16, pp 117-120.

Taddei 1991 — Taddei A, Biagini A, Distante G, Emdin M, Mazzei M G, Pisani P, Roggero N, Varanini M, (1991), "The European ST-T database: development, distribution and use", Computers in Cardiology, pp 177-180.

Taddei 1992 — Taddei A, Distante G, Emdin M, Pisani P, Moody G B, Zeelenberg C, Marchesi C, (1992), "The European ST-T Database: standard for evaluating systems for the analysis of ST-T changes in ambulatory electrocardiography", Eur Heart J, vol 13, pp 1164-1172.

Takagi 1985 — Takagi T, Sugeno M, (1985), "Fuzzy identification of systems and its application to modelling and control", IEEE Trans on Systems, Man and Cybernetics, vol 15, no 1, pp 116-132.

Tamada 1995 — Tamada J A, Bohannon N J V, Potts R O, (1995), "Measurement of Blood Glucose in Diabetic Subjects Using Non-invasive Transdermal Extraction", Nature Medicine, vol 1, pp 1198-1201.

Tattersall 1991 — Tattersall R B, Gill G V, (1991), "Unexplained deaths of type 1 diabetic patients", Diabetic Med, vol 8, pp 49-58.

Tsotsos 1987 — Tsotsos J K, Shibahara T, (1987), "Knowledge organization and its role in temporal and causal signal understanding:the ALVEN and CAA projects", N. Ccercone, G. McCalla (Eds.), The Knowledge Frontier, Springer, Berlin.

Tukey 1977 — Tukey J, (1977), "Exploratory Data Analysis", Addison-Wesley.

Vila 2000 — Vila J A, Gang Y, Rodriguez Presedo J M, Fernandez-Delgado M, Barro S, Malik M, (2000), "A new approach for TU complex characterization", IEEE Trans Biomed Eng, vol 47, no 6, pp 764-772.

Wang 1990 — Wang C H, Tseng S S (1990), "A brain tumor diagnostic system with automatic learning abilities", Proceedings of the Third Annual IEEE Symposium on Computer-Based Medical Systems, pp 313-320.

Wang 1997 — Wang L X, (1997), "A course in fuzzy systems and control", Prentice Hall International Editions.

Wilkes 2001 — Wilkes M, (2001), "A personal revisitation of neural nets".

Available at:

www-lce.eng.cam.ac.uk/publications/mvw/neuralnets.ps last checked 31/10/04

Wohlfart 1987 — Wohlfart B, (1987), A simple model for demonstration of STT-changes in ECG". Eur Heart J vol 8, pp 409:416.

Xiong 1983 — Xiong F L, Lambird B A, Kanal L N, (1983), "An experiment in recognition of electrocardiogram using a structural analysis algorithm", IEEE International Conference on Systems, Man and Cybernetics, pp. 382-387, Bombay and New Delhi, Dec. 1983 - Jan. 1984.

Yang 1994 — Yang T F, Devine B, Macfarlane P W, (1994), "Use of artificial neural networks within deterministic logic for the computer ECG diagnosis of inferior myocardial infarction", J Electrocardiol, vol 27, pp 188–193.

Zavalkoff 2002 Zavalkoff S R, Polychronakos C, (2002), "Evaluation of conventional blood glucose monitoring as an indicator of integrated glucose values using a continuous subcutaneous sensor", Diabetes Care, vol 25, no 9, pp 1603–1606.

URL 1 http://www.shef.ac.uk/~mpce/robi/diabetes-research-gif.html - Last accessed March 2005

URL 2 http://www.gnu.org - Last accessed 20/5/2005

URL 3 http://www.sleepsentry.com/ - Last accessed 12/10/2004

URL 4 http://www.minimed.com - Last accessed 20/5/2005

URL 5 http://www.glucowatch.com - Last accessed 20/5/2005

URL 6 http://www.cygn.com - Last accessed 20/5/2005

URL 7 http://www.numed.co.uk/seca.html - Last accessed 20/5/2005

URL 8 http://www.numed.co.uk/pocketview.html - Last accessed 20/5/2005

URL 9 http://www.emssupply.com/pocketview.htm - Last accessed 20/5/2005

URL 10 http://www.schiller.ch/products/powerslave,id,27,nodeid,5,_country,hq,_language,en.html - Last accessed 20/5/2005

URL 11 David Claus, Computer Vision Reading Group, Oxford. http://www.robots.ox.ac.uk/~dclaus/cameraloc/samples/nearestneighbour.ppt - Last accessed 20/5/2005

URL 12 http://www.akri.org/cognition/machmemmod.htm - Last accessed 20/5/2005

URL 13 http://www.physionet.org - Last accessed 20/5/2005

URL 14 http://www.physionet.org/physiobank - Last accessed 20/5/2005

URL 15 http://www.insulin-pumpers.org.uk/cgmsaccuracy - Last accessed 20/5/2005

URL 16 http://www.menarini.com Last accessed 16/5/2005

Appendix A: Dataset used

The data used in this thesis originated from Type 1 diabetic patients. The demographic details for the patients are given in table A.1.

Table A.1: Summary demographics

patient code	sex	DOB	age @ test	dur^n of diabetes(yrs)	HbA1c (%)
201	m	24/08/1979	21	3	9.2
202	f	17/01/1948	53	26	8.9
203	m	06/01/1980	21	8	8.1
204	f	09/10/1952	48	23	8.5
205	f		32		
207	m	09/12/1965	36	4 months	
208	m	10/07/1973	27	16	7.6
209	m	03/11/1938	62	24	11.3
210	f	05/02/1972	29	21	
211	f	27/03/1981	20	8	
212	f	11/07/1970	30	17	
215	m	14/02/1974	27	26	
216	f	25/01/1946	55	22	
218	m	08/02/1961	40	26	
219	m	15/11/1964	36		
220			22	3	
221	m	18/09/1975	26	2.5	
222	m	09/07/1963	38	18	
223	f	24/08/1983	18	2	
225	m	14/05/1961	40	13	
226	m	17/10/1973	28	5	
227	f	10/10/1959	42	27	
228	f	03/02/1963	39	23	
229	m	20/11/1981	20	3	
230	f	08/01/1962	40	12	
231	m		62	36	
232	m		35	29	
244	m	06/12/1979	23		

"HbA1c" presented in the last column of the table, is a measure of the quality of glucose control. HbA1c lies around 5% for healthy non-diabetic subjects while it is higher for diabetic patients. The higher the value, the poorer the quality of glucose control.

The ECG and glucose profiles for a subset of the patients are presented at the remainder of this appendix. ECG-glucose profiles for patients 202 and 204 have already been presented in Chapter 3. The ECG data for patient 201A (night1 and night2) are given in Figures A.1 and A.2. The ECG traces for each night are superimposed and plotted with different colours. The glucose profiles are presented in Figure A.3.

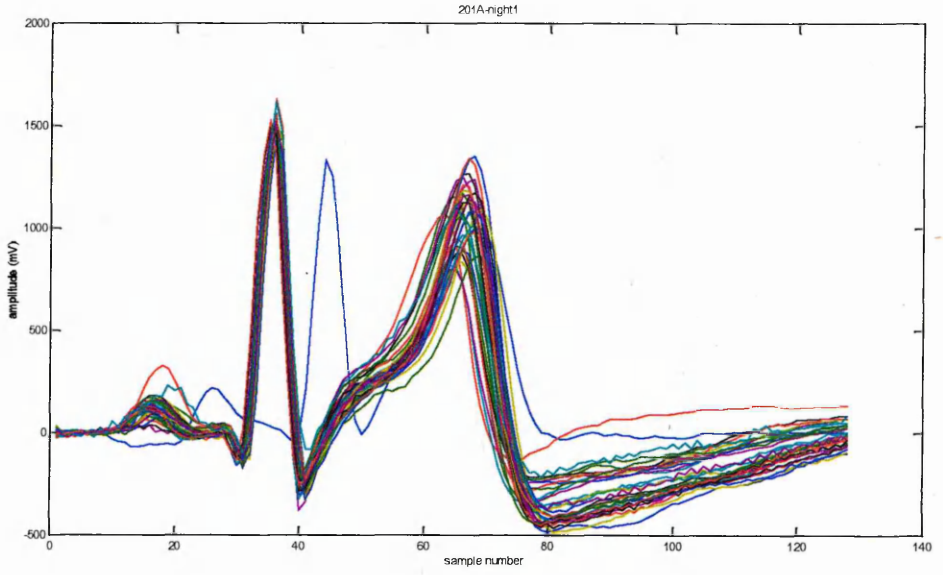


Figure A.1: ECG traces superimposed for patient 201A-night1

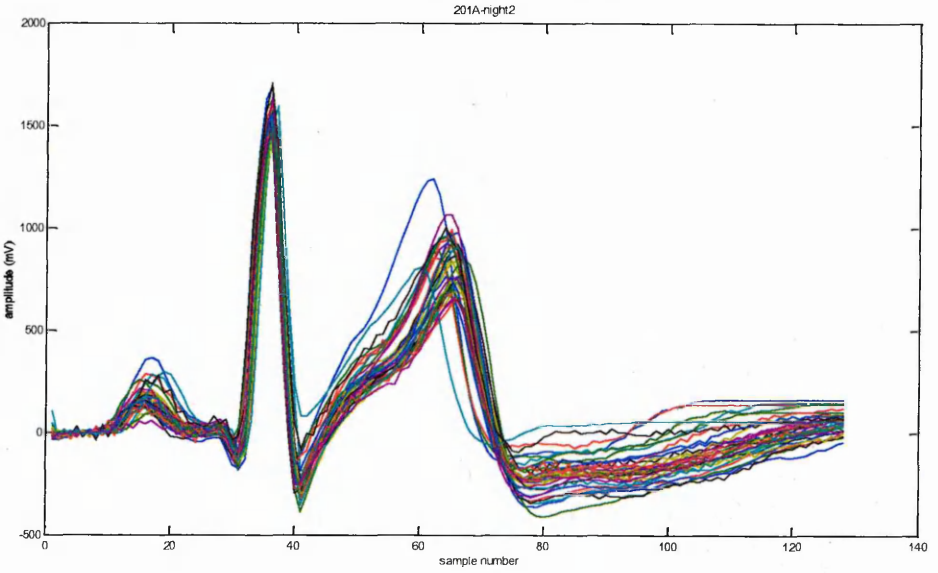


Figure A.2: ECG traces superimposed for patient 201A-night2

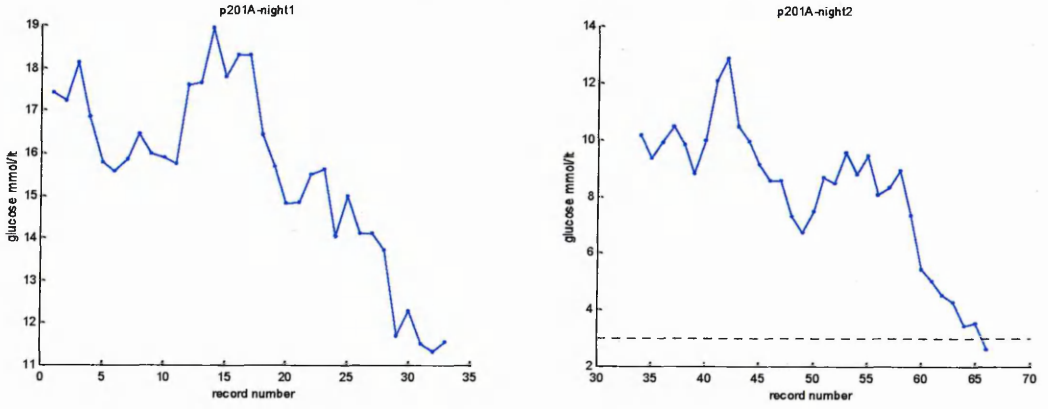


Figure A.3: Glucose profiles for 201A night1 (LHS) and night2 (RHS)

ECG and glucose profiles are given in Figures A.4-A.6 for patient 205.

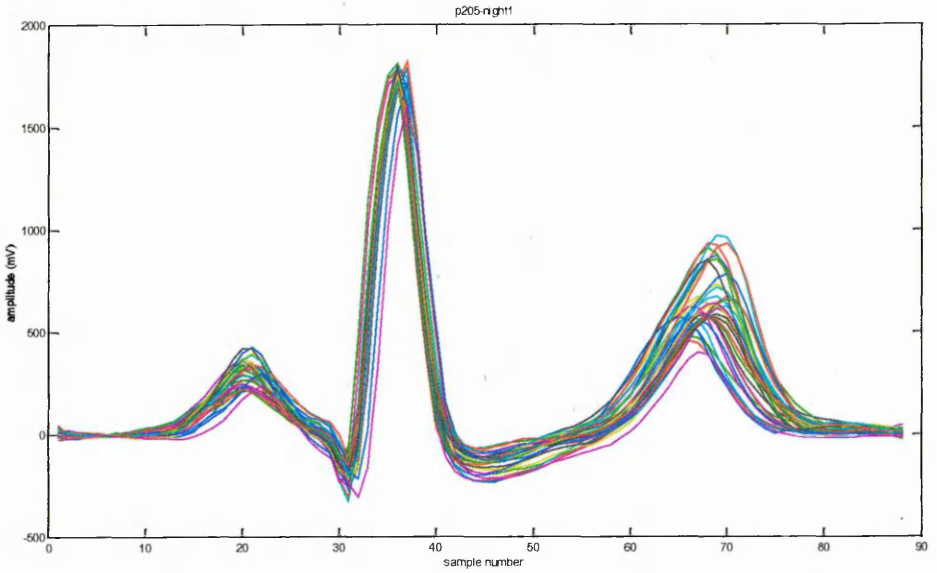


Figure A.4: ECG traces superimposed for patient 205-night1

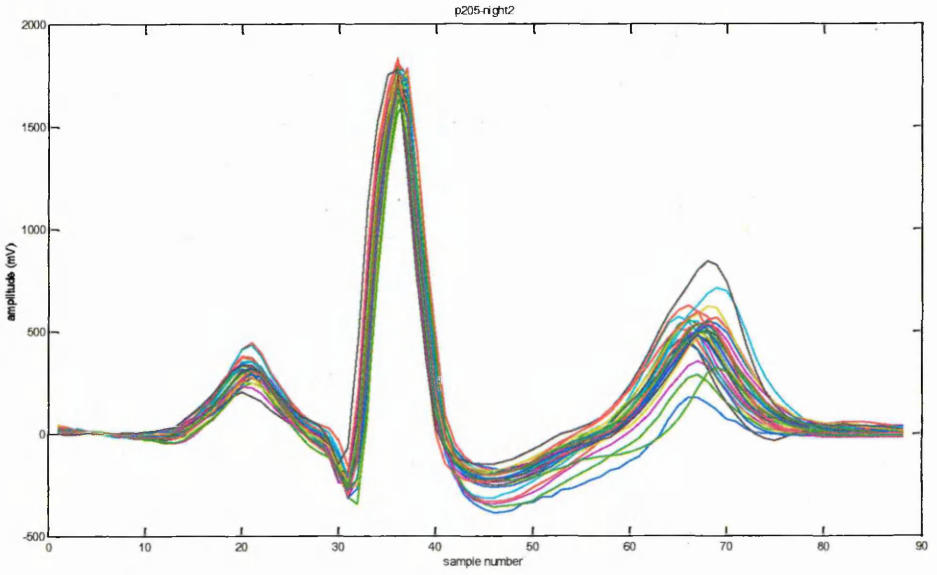


Figure A.5: ECG traces superimposed for patient 205-night2

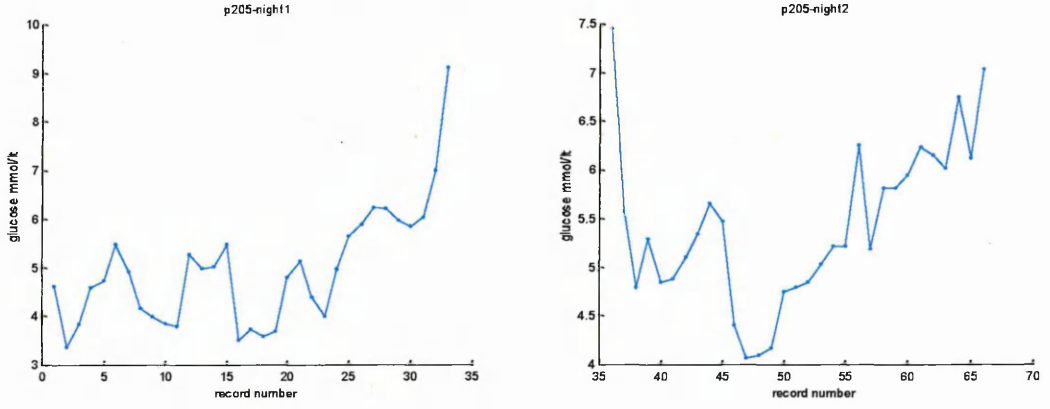


Figure A.6: Glucose profiles for 205 night1 (LHS) and night2 (RHS)

ECG and glucose profiles for patient 207 are given in Figures A.7-A.9.

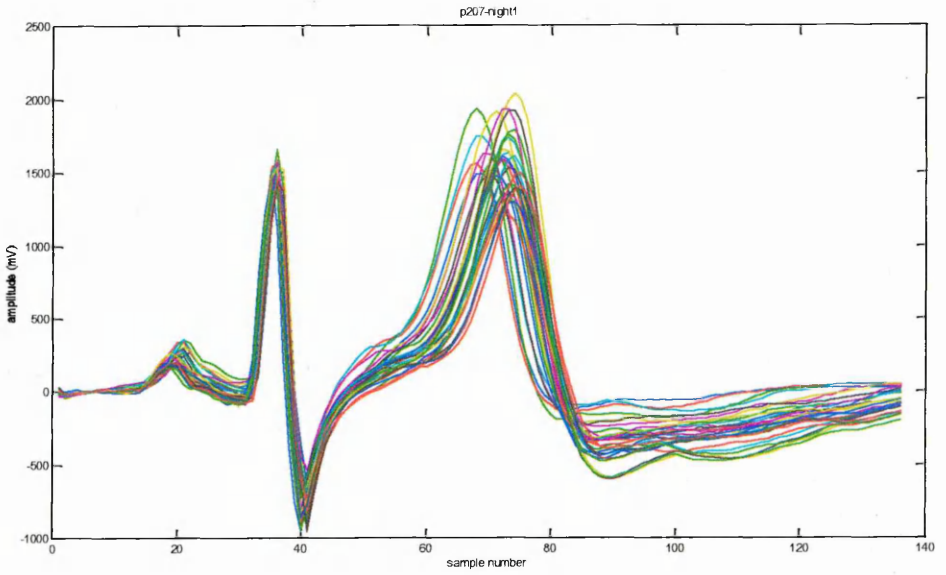


Figure A.7: ECG traces superimposed for patient 207-night1

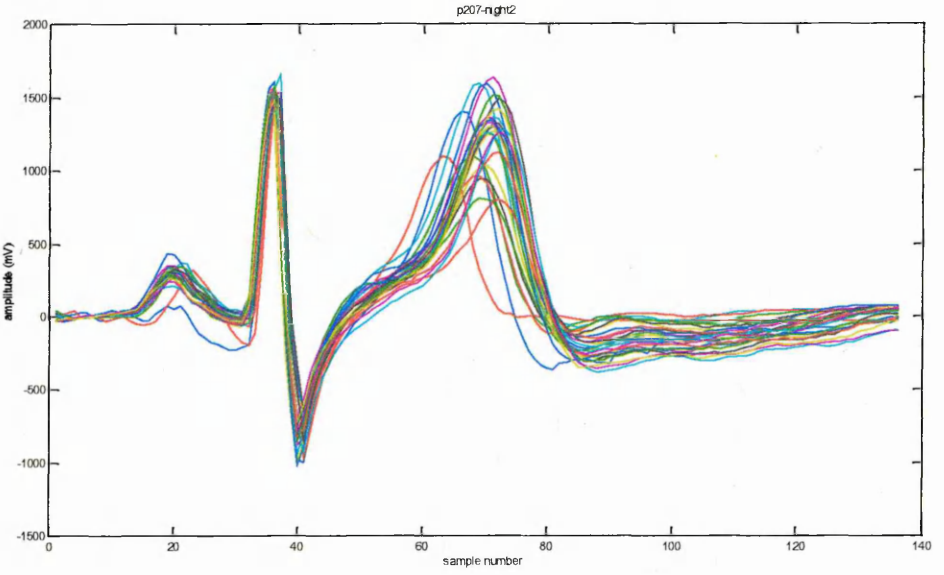


Figure A.8: ECG traces superimposed for patient 207-night2

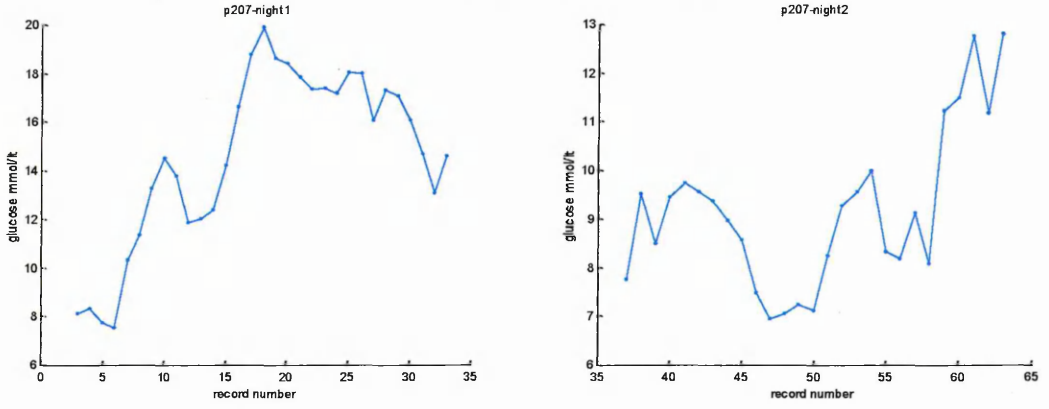


Figure A.9: Glucose profiles for 207 night1 (LHS) and night2 (RHS)

ECG and glucose profiles for patient 209 are given in Figures A.10-A.12.

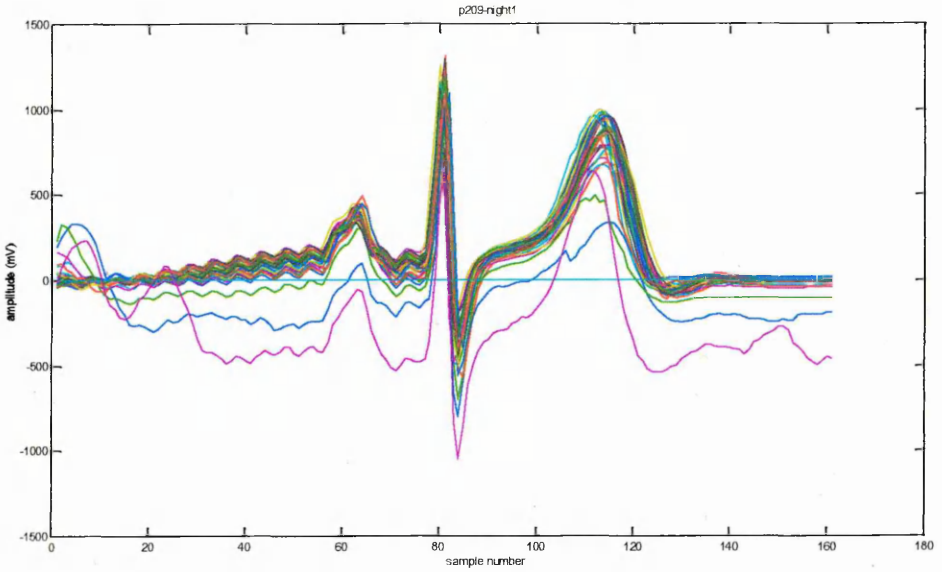


Figure A.10: ECG traces superimposed for patient 209-night1

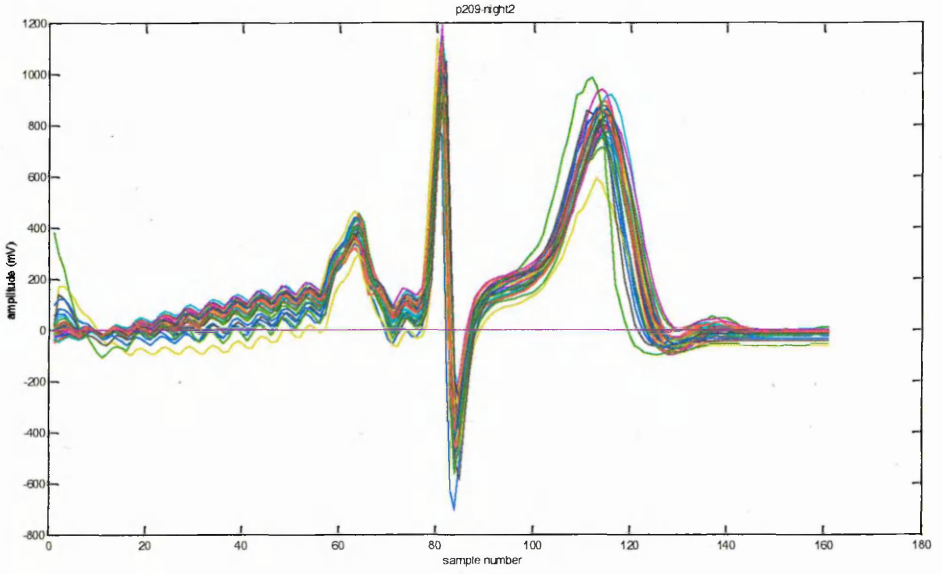


Figure A.11: ECG traces superimposed for patient 209-night2

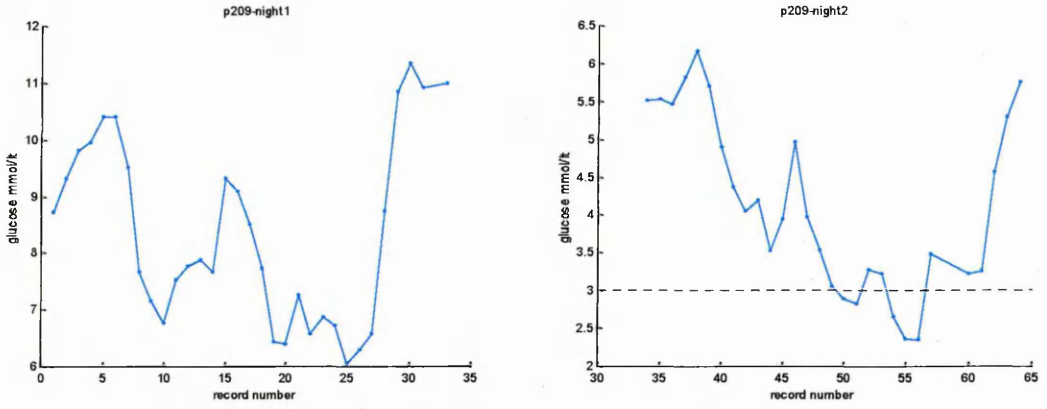


Figure A.12: Glucose profiles for 209 night1 (LHS) and night2 (RHS)

ECG and glucose profiles for patient 212 are given in Figures A.13-A.15.

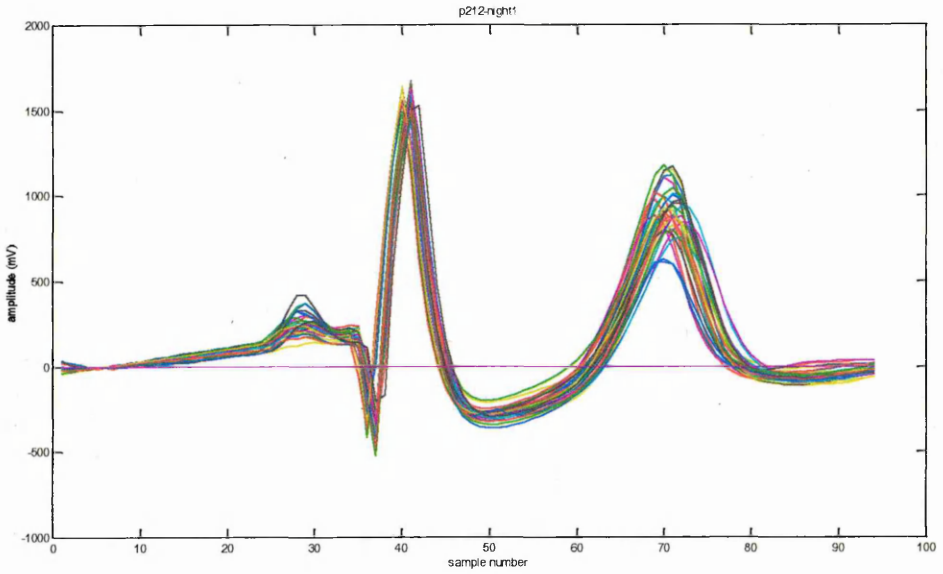


Figure A.13: ECG traces superimposed for patient 212-night1

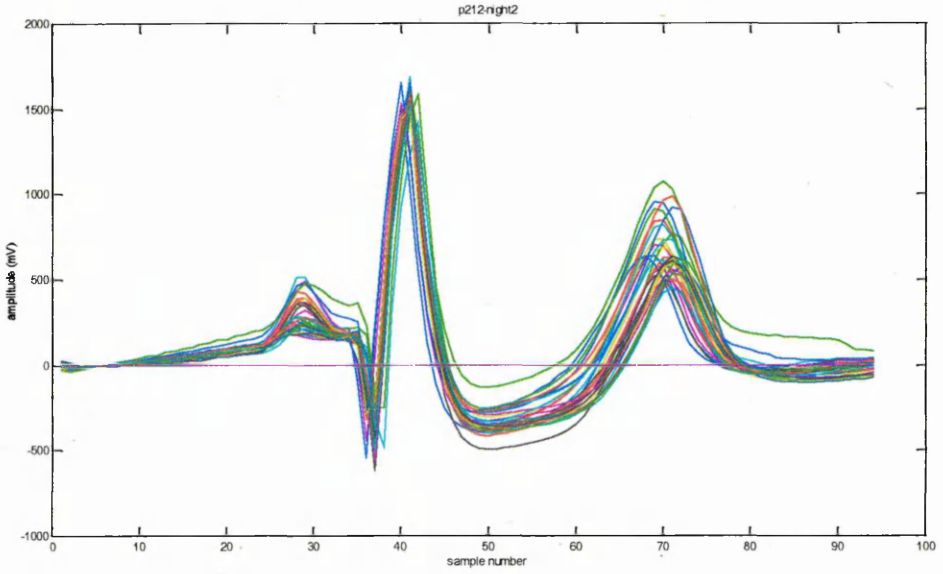


Figure A.14: ECG traces superimposed for patient 212-night2

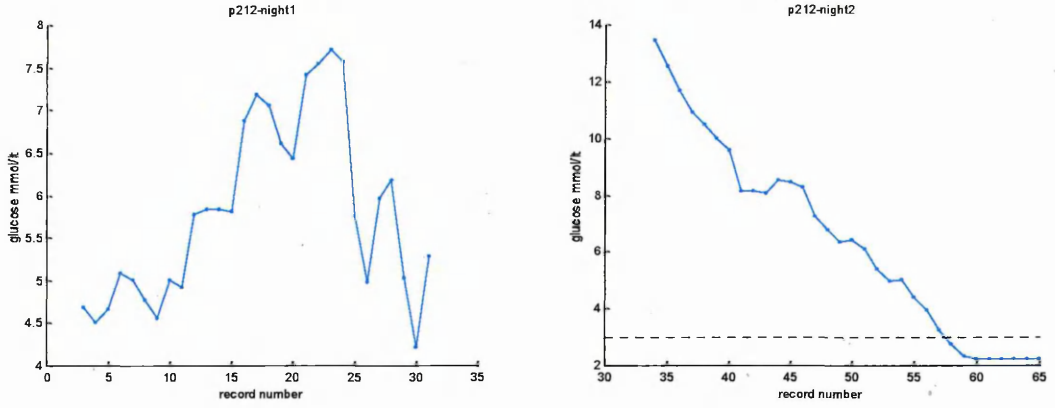


Figure A.15: Glucose profiles for 212 night1 (LHS) and night2 (RHS)

ECG and glucose profiles for patient 215 are given in Figures A.16-A.18.

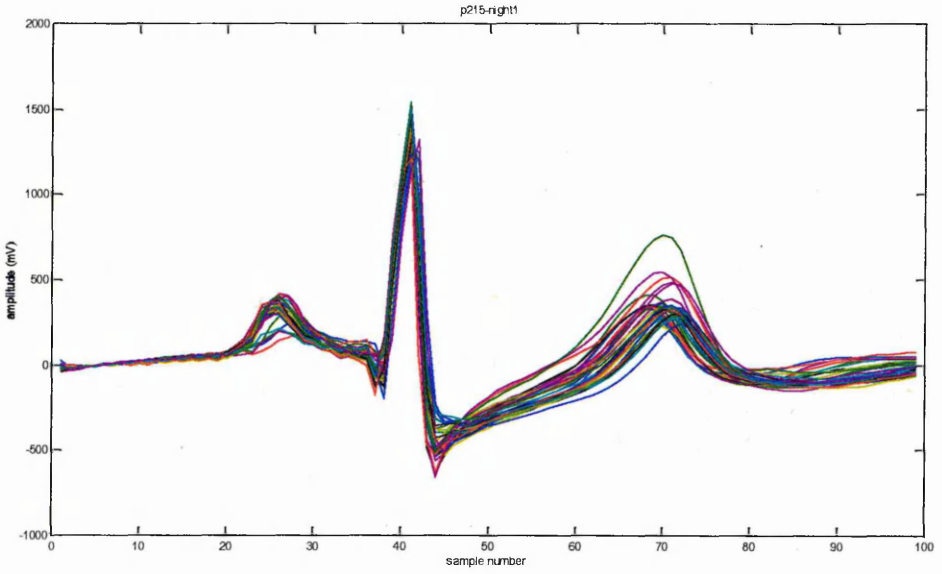


Figure A.16: ECG traces superimposed for patient 215-night1

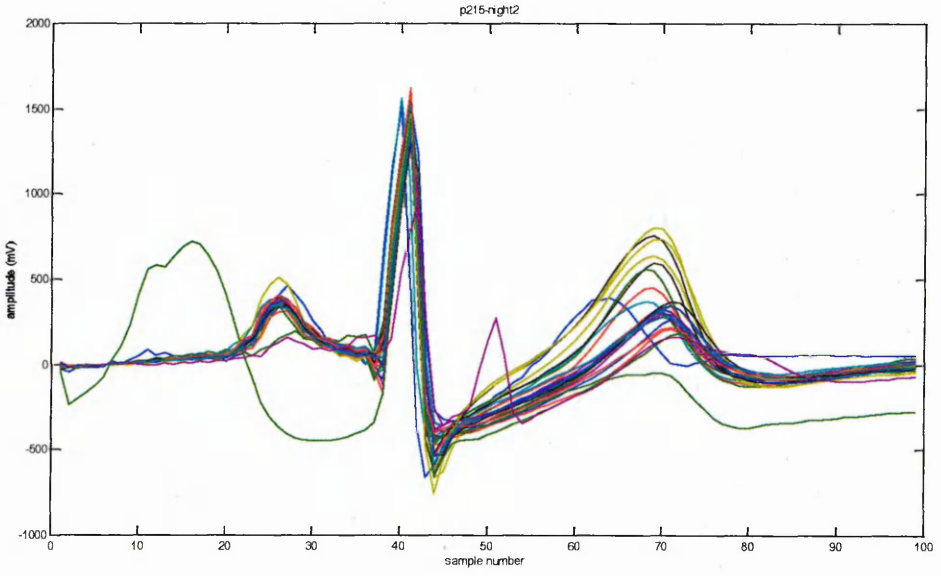


Figure A.17: ECG traces superimposed for patient 215-night2

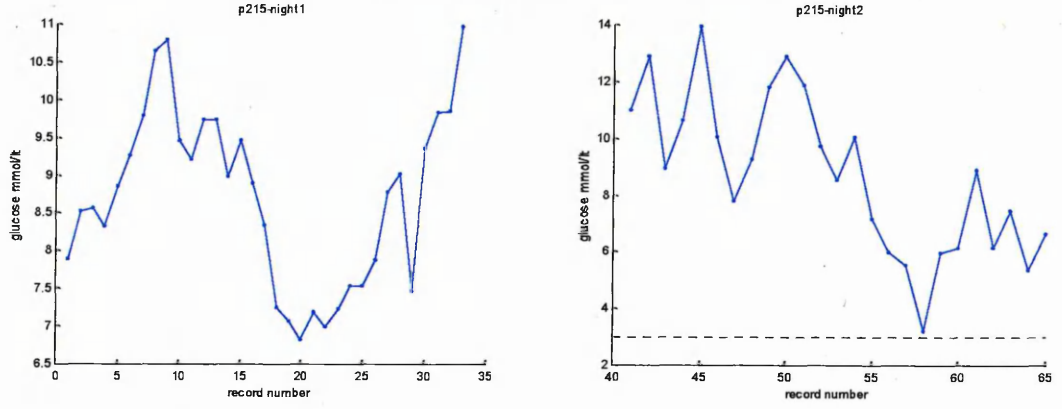


Figure A.18: Glucose profiles for 215 night1 (LHS) and night2 (RHS)

ECG and glucose profiles for patient 227 are given in Figures A.19-A.21.

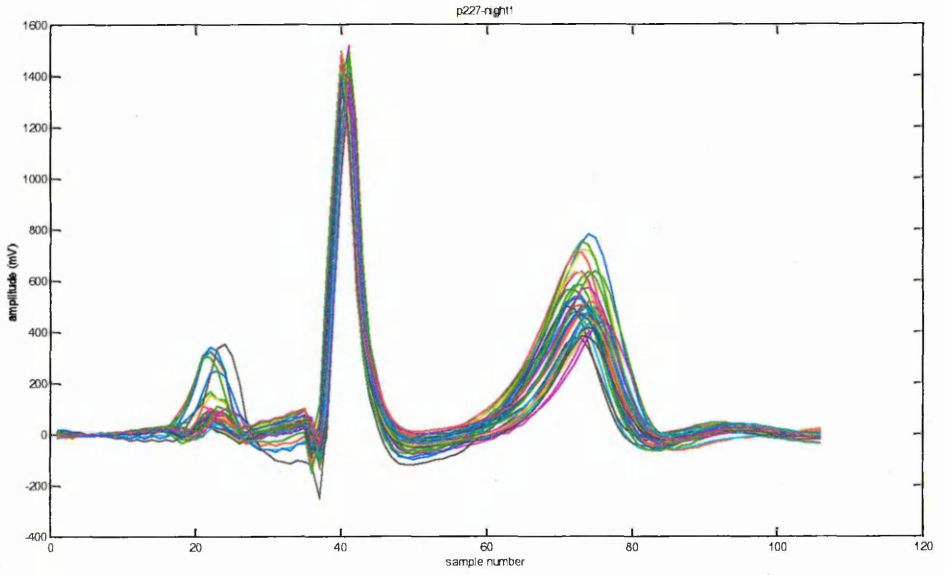


Figure A.19: ECG traces superimposed for patient 227-night1

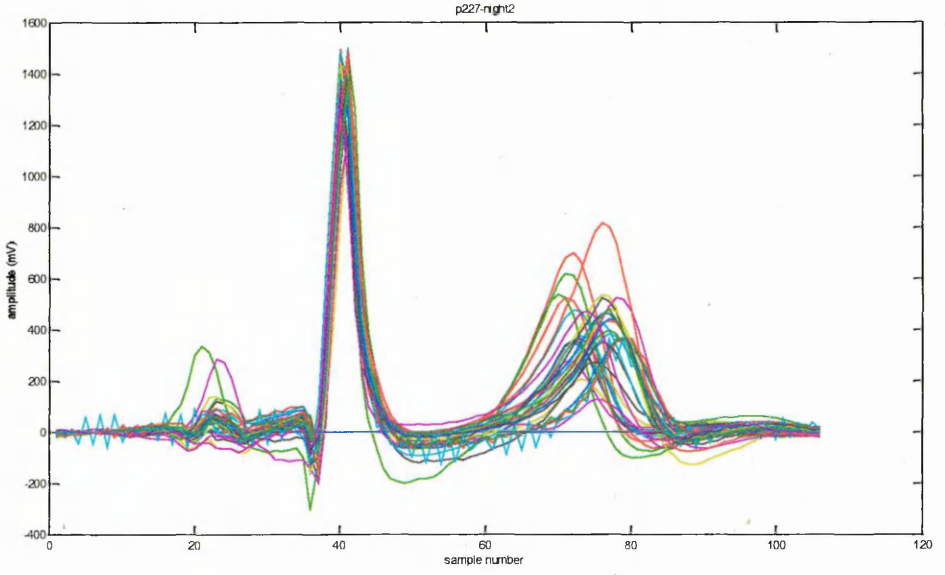


Figure A.20: ECG traces superimposed for patient 227-night2

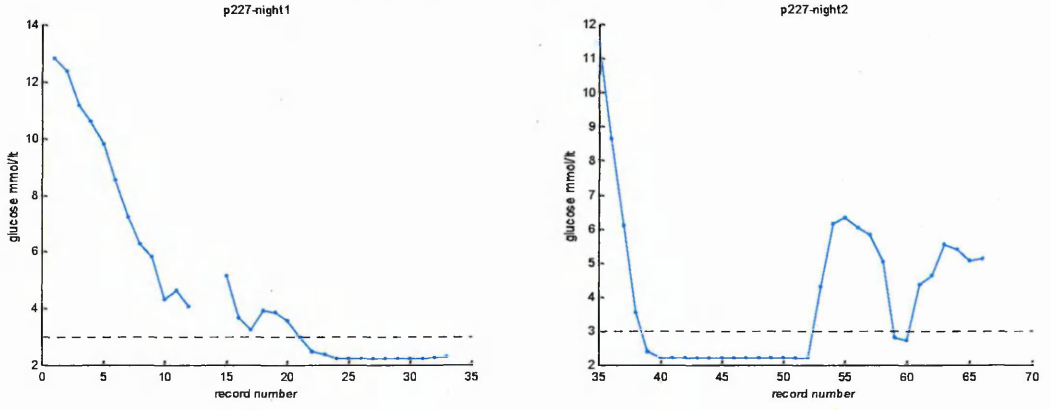


Figure A.21: Glucose profiles for 227 night1 (LHS) and night2 (RHS)

ECG and glucose profiles for patient 244 are given in Figures A.22-A.24.

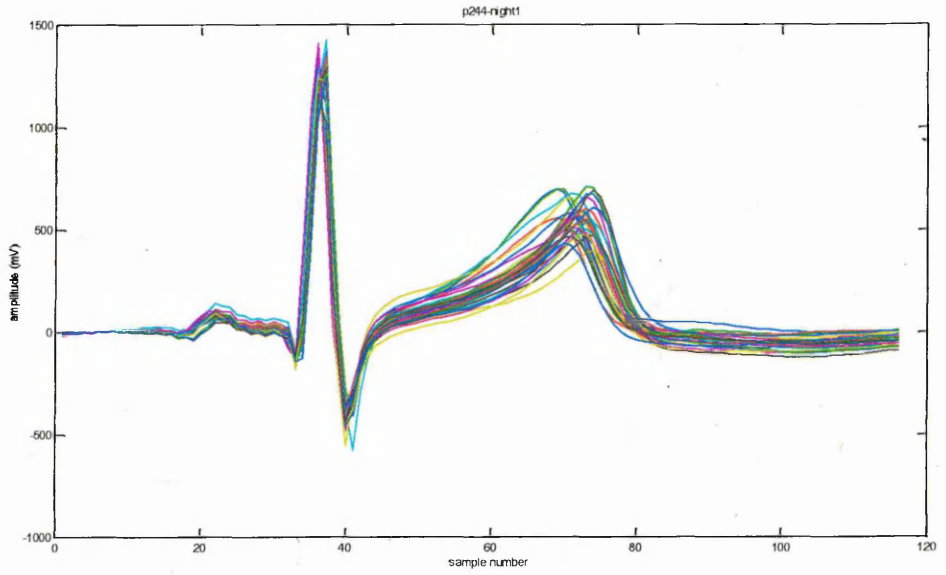


Figure A.22: ECG traces superimposed for patient 244-night1

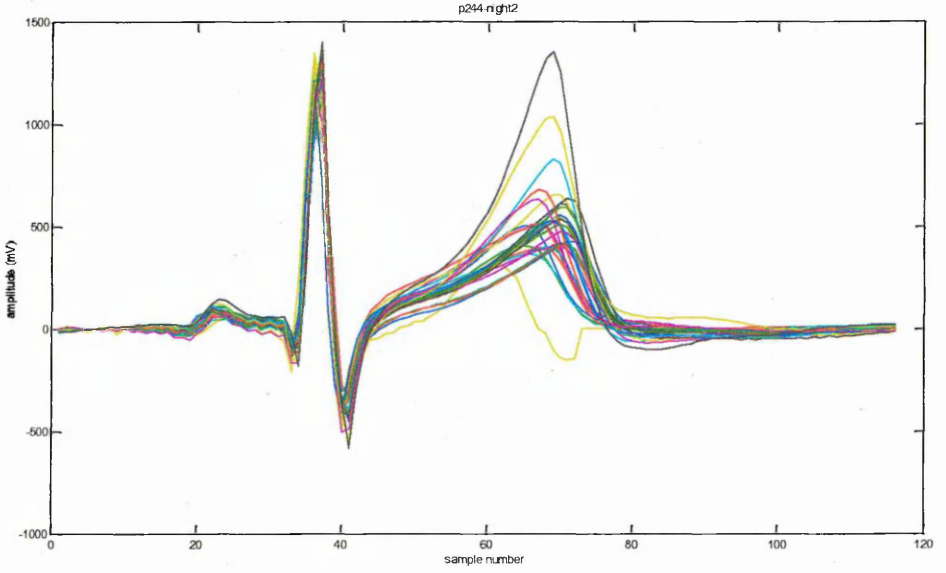


Figure A.23: ECG traces superimposed for patient 244-night2

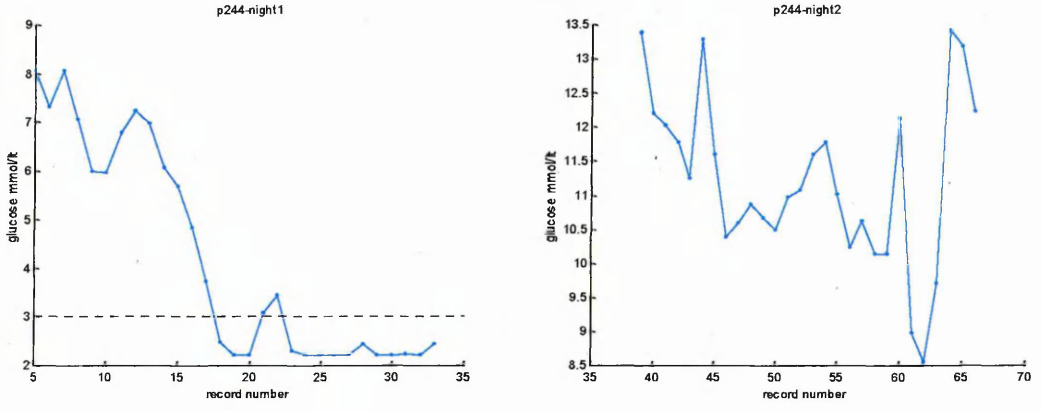


Figure A.24: Glucose profiles for 244 night1 (LHS) and night2 (RHS)

Appendix B: Simple fuzzy logic system example

In order to illustrate the concept of fuzzy set operations and fuzzy rules processing for a Mamdani-type system, a simple example is considered and briefly presented, from the MATLAB fuzzy logic toolbox manual [MathWorks Fuzzy]. This example is based on the basic tipping problem, namely what is the "right" amount to tip the waiting staff in a restaurant. This is formulated as: "Given a number between 0 and 10 that represents the quality of service at a restaurant (where 10 is excellent), what should the tip be?"[47] The fuzzy system to tackle the tipping problem, consisting of two inputs, three rules and one output and is presented in Figure B.1. The two inputs (quality of service, quality of food) also vary in the interval [1 10].

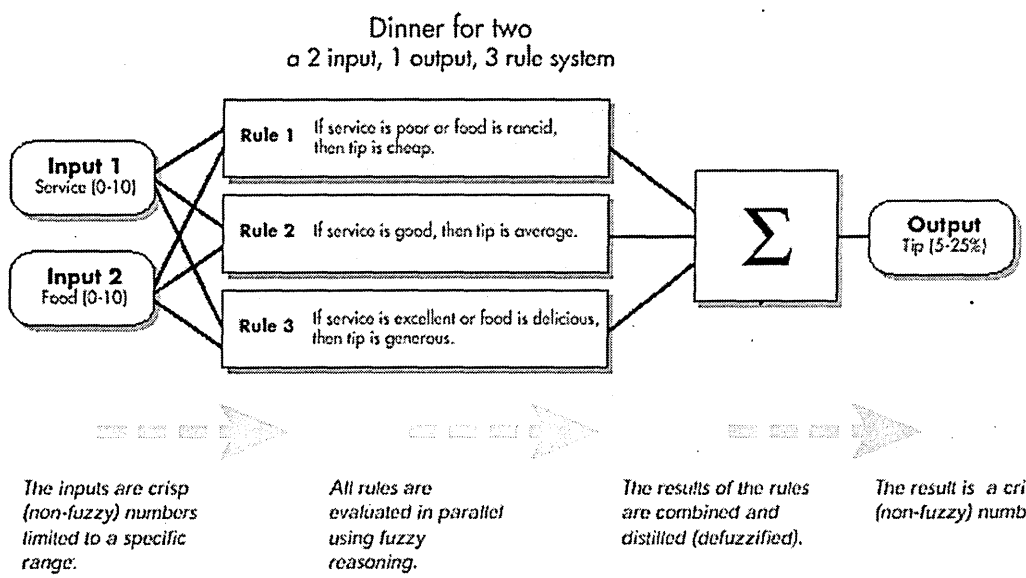


Figure B.1: Fuzzy tipper system [MathWorks Fuzzy tb]

B.1 Fuzzification

Figure B.2 illustrates the fuzzification of the input of the tipper system describing the quality of food. If the food is assessed as 8 (out of 10), this is fuzzified as having a membership of 0.7 in the "delicious" fuzzy set.

---

[47] This problem is based on tipping as it is typically practiced in the United States. An average tip for a meal in the U.S. is 15%, though the actual amount may vary depending on the quality of the service provided.

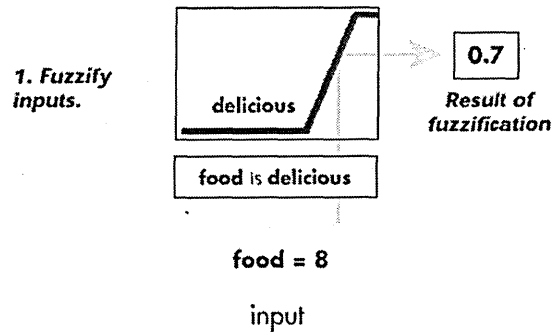


Figure B.2: Fuzzification [MathWorks Fuzzy]

B.2 Combination of fuzzy sets using the "Union" operator

Let's consider only the 3rd rule of the rule base: "if service is excellent or food is delicious then tip is generous". Figure B.3 demonstrates the use of the "Union" operator (OR) to combine two fuzzy sets related to this rule. After the fuzzification process, a set of inputs: service = 3, food = 8, will have memberships of 0 and 0.7 in the "excellent" and "delicious" fuzzy sets respectively which are involved in the above rule.

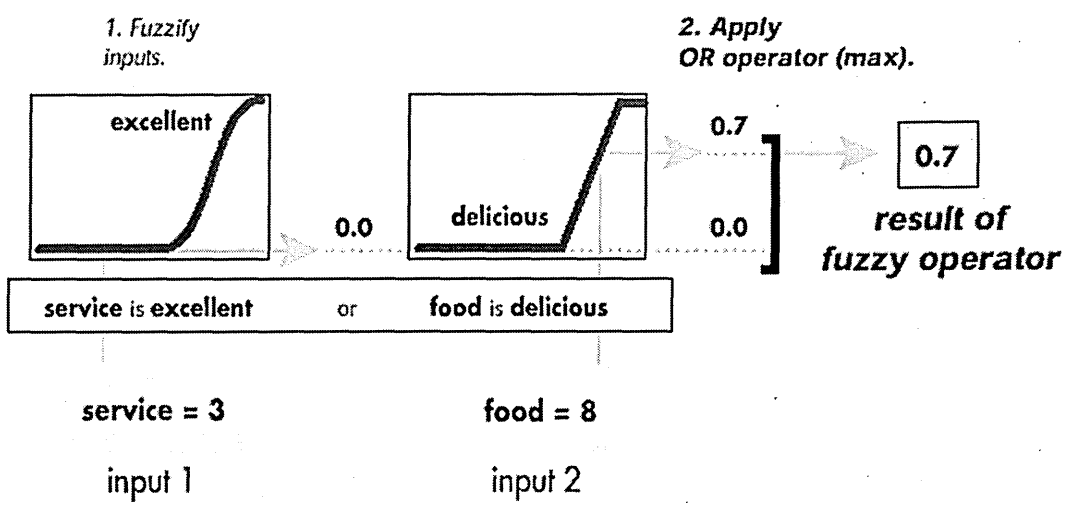


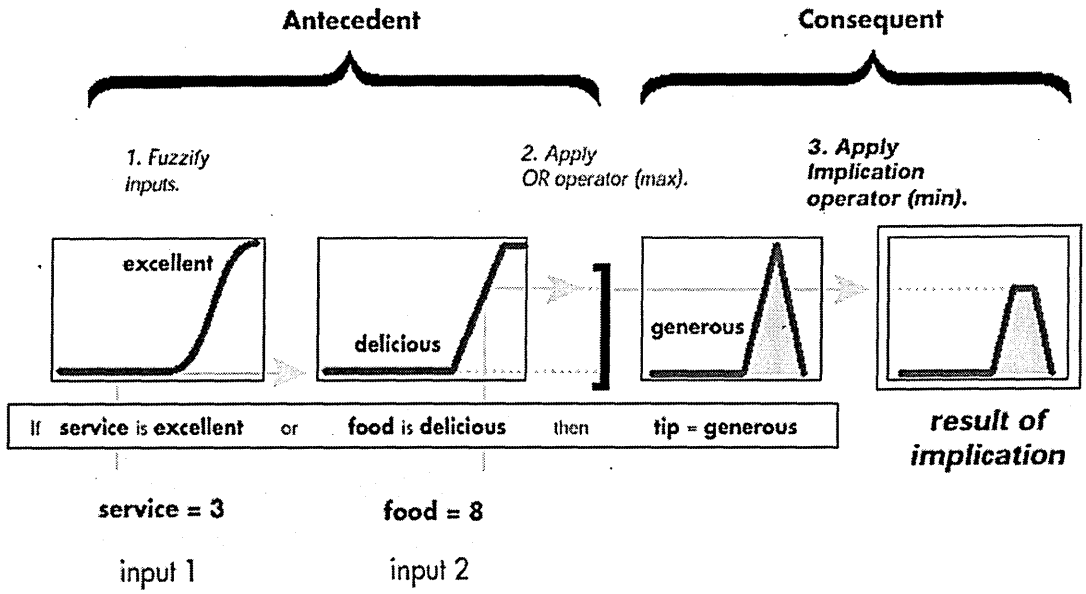
Figure B.3: Demonstration of Union operator (OR) [MathWorks Fuzzy]

The OR operator will combine the membership degrees of the two inputs. A "maximum" operator is used to implement the "fuzzy union" in this system giving an output of 0.7 (max(0, 0.7) = 0.7).

B.3 Implication operator

Once the fuzzy sets in the antecedent of the current rule are combined, the Implication operator is used to determine the firing strength of the rule. In this system, Implication is implemented by a "minimum" operator which is applied on the membership function with linguistic value "generous" as seen in Figure B.4. The result of the Implication

operator is shown in the right most graph of the figure. As it can be seen, the result of the operator is a MF identical to the "generous" MF but truncated at the 0.7 level.



FigureB.4: Demonstration of Implication operator [MathWorks Fuzzy]

B.4 Aggregation and Defuzzification

The process of aggregation for the set of inputs discussed above (service = 3, food = 8) is illustrated in Figure B.5.

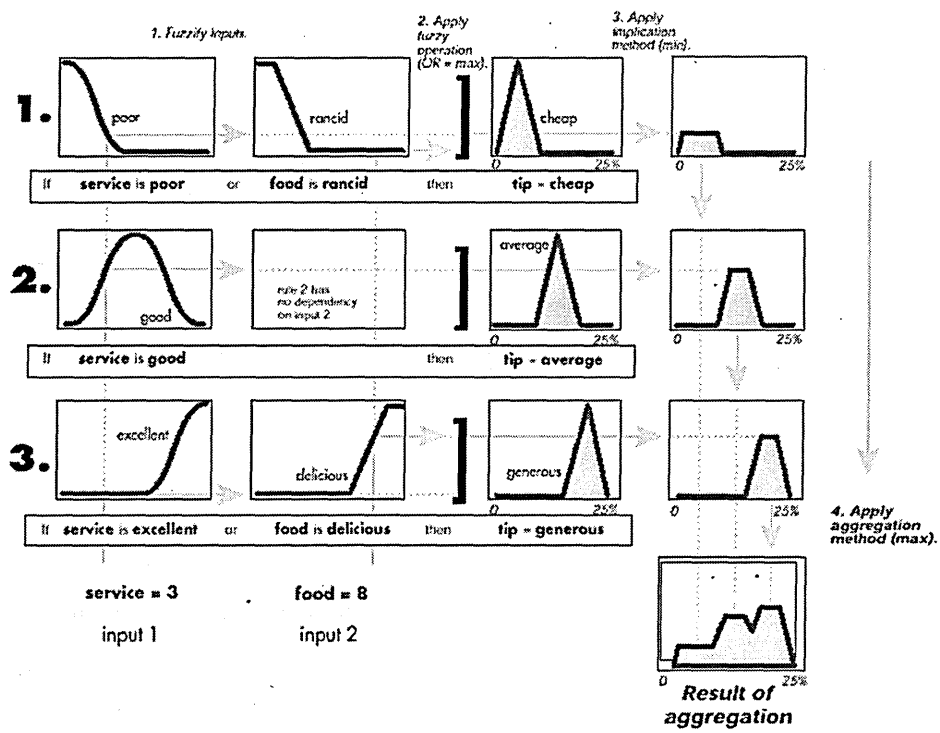


Figure B.5: Aggregation [MathWorks Fuzzy]

The rule considered up to now is the 3rd one on the figure. The use of the intersection and implication operators to determine the firing strengths of all the rules can be seen in the figure. The aggregation process combines all the rules into a single fuzzy set as seen in the bottom-right graph of Figure B.5. The fuzzy set produced by the aggregation is then defuzzified to produce a numerical output. The defuzzification method used in this system is the Centre of Gravity (COG). It calculates the centroid of the resulting fuzzy set as illustrated in Figure B.6.

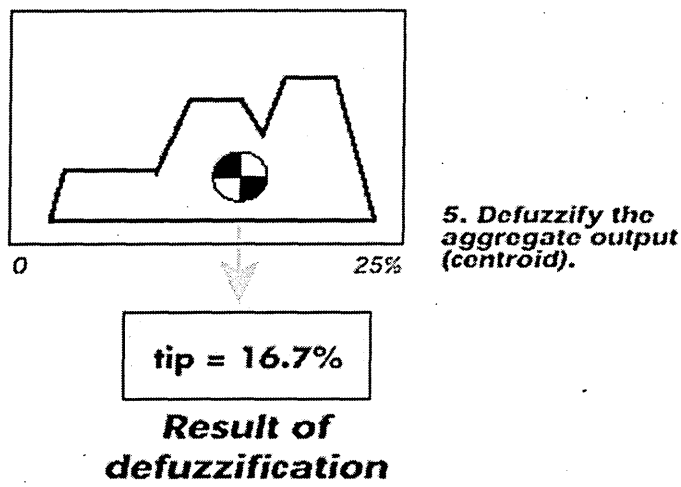


Figure B.6: Defuzzification [MathWorks Fuzzy]

The tipper system infers that when the service is rated as 3 and the food as then the tip should equal 16.7% of the bill.

Appendix C: Extracted ECG features

A small sample of the sets of features extracted is presented in this appendix.

Table C.1: ECG features for 203-night1

%Record	gl	HR	RTmsi	RTcmsi	Tdur	Tdurc	Tampl	Tarea	HAR	SKEW	KURTexc
7	12.24	87.69	261.52	316.17	156.32	188.98	612.97	7450.10	1.11	0.02	-0.67
8	8.81	85.80	246.92	295.28	159.43	190.66	810.41	10125.51	0.78	-0.74	-0.50
9	6.10	84.14	262.01	310.27	164.95	195.34	736.78	9707.78	1.23	0.18	-0.71
10	3.52	76.50	269.62	304.45	171.87	194.06	1041.20	13757.31	0.62	-1.12	-0.30
11	2.20	74.62	265.03	295.57	171.71	191.49	1088.90	14356.80	0.57	-1.21	-0.30
12	2.20	76.24	267.00	300.98	179.58	202.43	1136.23	15293.22	0.77	-0.92	-0.33
13	2.20	78.56	262.11	299.92	176.29	201.72	1166.76	15506.59	0.71	-1.03	-0.29
14	2.20	89.23	250.29	305.22	154.33	188.20	676.04	8302.61	1.00	-0.24	-0.67
15	2.20	83.22	281.43	331.44	186.05	219.11	761.50	11238.28	1.44	0.75	-0.60
16	2.20	69.80	285.50	307.93	171.51	184.98	681.20	9435.47	1.28	0.44	-0.69
17	2.20	72.23	285.26	312.99	175.38	192.43	706.32	10185.58	1.28	0.37	-0.74
18	2.20	66.30	279.56	293.88	148.17	155.76	540.41	6565.77	1.40	0.52	-0.78
19	2.20	64.30	287.22	297.33	152.25	157.61	561.00	6878.94	1.26	0.29	-0.78
20	2.20	65.94	287.23	301.10	156.43	163.99	537.89	6675.08	1.24	0.12	-0.73
21	2.20	71.24	287.84	313.64	176.03	191.81	547.38	7755.17	0.88	-0.50	-0.58
22	2.20	79.47	271.99	313.02	163.42	188.06	577.32	7367.79	0.80	-0.75	-0.55
23	2.20	79.24	282.40	324.54	178.86	205.55	676.14	9389.45	0.94	-0.49	-0.57
24	2.20	75.31	292.19	327.35	194.98	218.44	732.48	11386.07	0.95	-0.37	-0.62
25	2.20	76.93	282.65	320.06	173.12	196.04	655.18	9029.22	1.24	0.28	-0.76
26	2.20	80.28	279.44	323.23	175.08	202.52	691.76	9709.09	1.45	0.55	-0.71
27	2.20	82.19	270.13	316.17	173.74	203.34	744.81	10505.46	1.43	0.45	-0.76
28	2.20	77.85	283.48	322.91	175.16	199.52	921.04	12356.98	0.74	-0.82	-0.50
29	2.20	81.44	277.86	323.72	174.92	203.79	868.46	11832.22	0.60	-1.13	-0.37
30	3.25	77.82	278.88	317.61	181.24	206.41	1323.60	18425.94	0.65	-1.19	-0.23
31	4.37	80.98	276.97	321.76	167.43	194.51	1264.22	16302.14	0.73	-0.97	-0.40
32	7.14	76.78	275.76	311.96	167.13	189.06	1358.07	17037.79	0.54	-1.35	-0.18
33	8.26	82.20	273.99	320.70	177.61	207.89	948.89	12511.21	0.60	-1.27	-0.13

Table C.2: ECG features for 204

%Record	gl	HR	RTmsi	RTcmsi	Tdur	Tdurc	Tampl	Tarea	HAR	SKEW	KURTexc
5	8.48	84.55	291.38	345.89	100.88	119.75	206.62	1536.83	1.49	1.03	-0.33
6	8.48	83.30	294.13	346.57	107.02	126.10	248.57	1937.30	1.48	1.16	-0.15
7	7.09	86.97	297.79	358.54	105.79	127.37	226.08	1737.40	1.18	0.77	-0.44
8	5.08	85.81	297.92	356.28	122.20	146.13	368.19	3304.47	0.86	-0.27	-0.60
9	5.41	86.70	291.59	350.52	122.20	146.89	404.66	3636.03	0.98	-0.11	-0.60
10	5.66	95.56	283.54	357.83	118.78	149.90	362.94	3133.20	0.69	-0.74	-0.55
11	6.14	86.46	292.87	351.57	123.78	148.59	409.12	3766.22	0.72	-0.60	-0.51
12	5.23	86.77	293.65	353.14	106.16	127.66	201.02	1606.02	0.95	0.43	-0.53
13	4.91	87.77	296.20	358.25	121.81	147.33	391.82	3514.68	1.26	0.51	-0.53
14	4.69	84.48	297.87	353.44	125.63	149.07	435.32	3962.74	0.83	-0.32	-0.60
15	4.06	89.98	285.42	349.52	119.86	146.78	391.05	3453.42	0.76	-0.47	-0.56
16	4.09	92.30	277.29	343.92	104.57	129.70	271.56	2161.09	0.78	-0.40	-0.61
17	2.82	88.12	291.84	353.69	92.68	112.32	132.84	912.46	1.17	0.74	-0.45
18	2.20	90.40	295.01	362.11	98.98	121.49	154.70	1153.99	1.26	0.87	-0.33
19	2.20	86.86	290.00	348.94	100.46	120.88	168.75	1275.51	1.40	1.02	-0.32
20	2.20	89.10	282.57	344.34	97.13	118.37	171.53	1229.85	1.07	0.49	-0.61
21	2.20	92.09	290.14	359.45	98.57	122.11	155.88	1159.65	1.58	1.16	-0.25
22	2.20	88.53	291.86	354.53	104.14	126.50	157.15	1185.12	1.36	1.09	-0.36
23	2.20	92.31	294.11	364.81	97.73	121.22	127.94	962.25	1.22	0.88	-0.37
24	2.20	84.16	300.01	355.33	121.76	144.21	432.17	3982.23	1.05	0.17	-0.53
25	2.20	86.82	297.47	357.82	123.14	148.12	446.38	4002.90	0.92	-0.19	-0.60
26	2.20	84.54	294.02	349.00	98.69	117.15	162.38	1193.52	1.26	1.03	-0.30
27	2.20	91.03	294.74	363.06	100.72	124.06	179.40	1262.76	1.16	0.82	-0.32
28	2.20	86.74	288.66	347.07	96.63	116.19	162.56	1184.09	0.94	0.22	-0.62
29	2.20	88.92	292.33	355.87	104.35	127.04	215.67	1766.22	0.87	0.25	-0.57
30	2.76	87.17	290.94	350.69	105.43	127.08	208.13	1608.22	1.34	0.96	-0.40
31	2.20	86.15	302.17	362.07	123.68	148.20	404.52	3757.61	0.72	-0.56	-0.55
32	2.2	90.59	289.88	356.19	123.52	151.77	390.61	3531.06	0.91	-0.19	-0.60
33	2.2	88.88	294.15	358.01	124.84	151.94	413.89	3782.15	1.04	0.14	-0.56

Table C.3: ECG features for 205-night1

%Record	gl	HR	RTmsi	RTcmsi	Tdur	Tdurc	Tampl	Tarea	HAR	SKEW	KURTexc
1	4.60	95.48	295.40	372.65	116.74	147.26	529.25	4449.52	0.95	-0.09	-0.53
2	3.36	87.16	309.21	372.68	126.57	152.56	476.56	4418.54	1.40	0.83	-0.38
3	3.83	91.88	297.21	367.78	114.38	141.54	454.80	3762.01	1.08	0.08	-0.57
4	4.58	91.32	293.01	361.48	118.98	146.78	577.99	5077.83	0.89	-0.39	-0.51
5	4.72	86.64	299.38	359.76	127.83	153.62	619.86	5930.57	0.91	-0.46	-0.48
6	5.47	90.83	294.02	361.76	136.59	168.06	670.66	6721.54	0.69	-0.92	-0.36
7	4.91	85.87	311.29	372.41	124.28	148.68	582.62	5439.50	0.76	-0.55	-0.48
8	4.16	79.93	314.48	362.98	137.18	158.33	870.87	8649.23	0.58	-1.22	-0.16
9	3.99	79.16	309.87	355.92	140.51	161.39	909.40	9151.49	0.70	-1.04	-0.17
10	3.85	80.38	312.35	361.53	143.84	166.49	930.75	9443.58	0.77	-0.96	-0.20
11	3.78	79.03	321.32	368.76	148.61	170.55	968.06	10119.47	0.75	-1.00	-0.22
12	5.26	88.39	305.54	370.85	122.87	149.13	590.01	5300.68	0.99	-0.14	-0.58
13	4.98	83.69	319.96	377.88	129.68	153.15	570.81	5505.92	0.76	-0.57	-0.53
14	5.00	81.06	322.28	374.60	134.64	156.50	587.53	5883.30	0.77	-0.61	-0.48
15	5.47	97.29	288.76	367.70	127.29	162.09	573.76	5483.15	0.84	-0.49	-0.52
16	3.50	75.59	329.63	369.98	143.04	160.55	654.39	7042.03	0.71	-0.91	-0.32
17	3.71	75.55	326.43	366.29	135.84	152.43	672.90	6734.90	0.72	-0.88	-0.36
18	3.57	76.95	324.96	368.01	131.52	148.94	621.05	6151.77	0.66	-0.91	-0.36
19	3.69	78.26	320.54	366.09	128.43	146.68	610.22	5810.33	0.84	-0.57	-0.47
20	4.79	81.29	317.61	369.68	142.62	166.01	730.85	7751.01	0.62	-1.11	-0.26
21	5.12	81.31	320.20	372.75	137.79	160.41	570.33	5761.47	0.70	-0.84	-0.39
22	4.37	79.33	317.52	365.09	140.86	161.97	778.41	8104.57	0.62	-1.09	-0.25
23	3.99	80.30	315.89	365.45	136.39	157.78	853.43	8475.00	0.59	-1.21	-0.13
24	4.96	79.31	315.71	362.97	139.12	159.94	929.29	9259.51	0.60	-1.22	-0.12
25	5.64	81.58	320.22	373.38	130.58	152.26	714.17	6927.32	0.80	-0.60	-0.50
26	5.87	82.34	317.82	372.31	131.95	154.58	637.86	6338.91	0.68	-0.83	-0.39
27	6.22	82.83	319.20	375.05	138.50	162.73	628.00	6436.88	0.65	-0.91	-0.41
28	6.20	82.87	307.27	361.12	151.15	177.64	849.34	8622.01	0.57	-1.32	0.02
29	5.97	88.59	296.95	360.83	117.01	142.17	550.70	4679.16	1.01	-0.17	-0.49
30	5.82	83.51	316.99	373.96	132.71	156.56	567.61	5658.28	0.92	-0.21	-0.54
31	6.01	82.08	321.59	376.14	136.42	159.57	569.32	5744.83	0.98	-0.13	-0.58
32	6.97	83.64	320.49	378.40	146.49	172.96	670.43	7296.42	0.62	-1.04	-0.34
33	9.09	97.86	287.42	367.07	98.84	126.23	402.94	2870.93	1.03	0.03	-0.55

Table C.4: ECG features for 205-night2

%Record	gl	HR	RTmsi	RTcmsi	Tdur	Tdurc	Tampl	Tarea	HAR	SKEW	KURTexc
38	4.79	96.36	296.14	375.29	112.73	142.86	510.52	4123.69	1.51	0.95	-0.33
39	5.29	96.22	290.71	368.14	113.44	143.66	568.67	4695.72	1.12	0.30	-0.55
40	4.84	97.25	292.07	371.84	104.11	132.55	353.90	2640.57	1.01	0.22	-0.45
41	4.87	98.73	285.25	365.91	113.46	145.55	448.40	3753.62	0.85	-0.33	-0.48
42	5.10	96.73	283.74	360.26	106.17	134.81	457.56	3541.17	1.08	-0.15	-0.56
43	5.34	93.39	285.19	355.80	111.99	139.71	438.59	3547.22	0.89	-0.31	-0.54
44	5.65	90.36	301.32	369.78	122.47	150.29	557.27	4950.36	1.00	0.00	-0.50
45	5.47	91.99	294.46	364.61	123.31	152.69	623.41	5659.71	0.81	-0.58	-0.41
46	4.40	90.06	295.43	361.95	121.01	148.26	549.22	4793.45	0.82	-0.33	-0.54
47	4.06	90.77	297.11	365.43	104.37	128.37	490.70	3748.01	1.32	0.58	-0.50
48	4.09	91.41	295.63	364.89	108.89	134.40	466.31	3654.35	1.15	0.29	-0.48
49	4.17	91.58	307.80	380.26	117.71	145.43	478.09	3953.11	1.04	0.16	-0.58
50	4.74	92.21	305.87	379.19	118.06	146.36	474.68	3992.55	0.91	-0.14	-0.50
51	4.79	83.58	310.05	365.93	121.53	143.43	536.77	4740.87	1.20	0.41	-0.52
52	4.84	84.58	314.01	372.82	128.05	152.04	563.57	5190.71	0.78	-0.56	-0.46
53	5.02	88.20	311.94	378.21	118.80	144.04	526.10	4465.38	0.90	-0.26	-0.50
54	5.21	89.12	309.77	377.52	121.35	147.88	493.97	4312.80	1.19	0.41	-0.49
55	5.21	87.45	309.08	373.14	128.51	155.15	498.08	4493.40	0.98	-0.02	-0.54
56	6.25	87.29	315.83	380.96	124.95	150.72	504.93	4450.67	0.92	-0.19	-0.57
57	5.18	88.48	314.45	381.85	131.69	159.92	546.49	5207.43	1.09	0.11	-0.59
58	5.81	91.82	306.94	379.71	89.69	110.96	323.29	2071.29	1.34	0.88	-0.36
59	5.81	89.64	313.46	383.15	127.30	155.60	540.39	5149.17	1.15	0.31	-0.57
60	5.94	84.69	320.64	380.95	134.53	159.84	711.71	7157.57	0.76	-0.73	-0.43
61	6.22	87.47	317.87	383.81	131.55	158.84	529.66	5200.64	0.95	-0.04	-0.57
62	6.14	89.29	316.65	386.27	139.24	169.85	619.88	6518.94	0.80	-0.52	-0.50
63	6.01	81.32	311.70	362.87	139.02	161.84	838.88	8455.97	0.74	-0.89	-0.35
64	6.74	92.07	307.72	381.19	119.42	147.93	501.61	4376.78	1.07	0.18	-0.63
65	6.12	88.90	319.63	389.08	140.10	170.54	494.66	5006.76	0.77	-0.47	-0.49
66	7.03	93.62	303.90	379.60	139.86	174.71	593.84	5880.00	0.71	-0.92	-0.34

Table C.5: ECG features for 209-night1

%Record	gl	HR	RTmsi	RTcmsi	Tdur	Tdurc	Tampl	Tarea	HAR	SKEW	KURTexc
1	8.70	112.12	159.90	218.58	82.88	113.29	945.67	12014.05	0.57	-1.29	-0.17
2	9.32	107.80	165.68	222.07	88.25	118.29	780.62	10770.64	0.56	-1.32	-0.16
3	9.80	98.75	171.78	220.37	93.56	120.03	687.21	9759.87	0.60	-1.25	-0.22
4	9.95	92.69	171.02	212.56	93.78	116.56	674.77	9766.67	0.56	-1.33	-0.13
5	10.39	105.09	168.87	223.50	90.52	119.80	714.50	9754.39	0.73	-1.03	-0.30
6	10.39	85.20	169.55	202.05	90.56	107.92	741.04	10051.31	0.71	-1.04	-0.31
7	9.51	109.15	162.23	218.81	90.46	122.00	919.80	12661.41	0.60	-1.28	-0.16
8	7.65	109.32	164.94	222.64	94.14	127.08	933.79	13386.60	0.66	-1.19	-0.24
9	7.14	101.57	166.18	216.21	92.66	120.56	981.67	13939.19	0.57	-1.34	-0.10
10	6.76	105.15	169.75	224.71	95.94	127.00	959.22	14092.50	0.65	-1.25	-0.15
11	7.52	84.47	167.49	198.73	92.79	110.09	987.66	13766.76	0.60	-1.25	-0.20
12	7.75	104.29	171.66	226.31	96.79	127.61	962.96	14016.55	0.74	-1.04	-0.29
13	7.86	105.71	171.05	227.04	93.22	123.73	974.95	13890.79	0.57	-1.34	-0.11
14	7.65	104.31	172.21	227.07	95.17	125.48	961.05	13902.60	0.60	-1.30	-0.12
15	9.32	107.01	164.96	220.30	87.03	116.23	972.24	12858.67	0.58	-1.33	-0.08
16	9.08	102.58	173.30	226.60	96.40	126.05	883.61	12882.51	0.61	-1.28	-0.13
17	8.49	100.97	169.21	219.50	91.31	118.46	941.60	13133.98	0.71	-1.15	-0.18
18	7.73	108.36	159.41	214.23	95.00	127.67	963.95	13103.95	0.54	-1.47	0.17
19	6.42	107.17	172.12	230.04	101.26	135.34	875.57	13518.78	0.69	-1.15	-0.25
20	6.38	104.76	176.81	233.64	104.44	138.01	861.06	13815.94	0.63	-1.24	-0.21
21	7.25	106.62	178.12	237.44	103.67	138.20	785.22	12449.61	0.55	-1.33	-0.20
22	6.57	103.93	162.65	214.07	69.68	91.71	335.46	3748.97	0.61	-1.38	-0.12
23	6.87	125.54	150.39	217.54	68.94	99.72	497.31	5864.55	0.59	-1.18	-0.39
24	6.70	85.97	167.87	200.95	87.55	104.81	873.83	11297.32	0.50	-1.39	-0.17
25	6.04	106.87	170.73	227.86	99.71	133.07	779.21	11133.09	0.53	-1.41	-0.05
26	6.27	113.14	170.74	234.45	93.42	128.29	797.50	11380.46	0.51	-1.37	-0.16
27	6.57	115.75	156.37	217.19	84.29	117.08	1003.26	12771.99	0.50	-1.45	-0.05
28	8.72	112.11	172.01	235.13	91.01	124.40	874.23	12114.32	0.66	-1.20	-0.17
29	10.84	115.35	161.14	223.43	87.61	121.47	960.37	12825.70	0.64	-1.29	-0.08
30	11.32	108.16	168.86	226.72	92.69	124.45	894.98	12448.80	0.54	-1.41	-0.04
31	10.90	115.04	154.11	213.39	77.42	107.20	836.96	10127.86	0.49	-1.37	-0.18
33	10.98	78.74	143.46	164.35	53.52	61.31	635.47	5014.27	0.99	-0.34	-0.87

Table C.6: ECG features for 209-night2

%Record	gl	HR	RTmsi	RTcmsi	Tdur	Tdurc	Tampl	Tarea	HAR	SKEW	KURTexc
34	5.51	118.88	166.24	234.00	87.73	123.48	824.66	10854.38	0.81	-0.81	-0.41
35	5.53	108.26	170.58	229.13	90.45	121.50	750.54	10329.68	0.77	-0.89	-0.35
36	5.46	109.77	170.70	230.88	96.37	130.34	823.08	11994.24	0.68	-1.08	-0.29
37	5.81	106.31	175.21	233.23	98.92	131.68	759.50	11500.41	0.61	-1.18	-0.25
38	6.16	104.49	174.75	230.61	99.34	131.10	787.87	11946.91	0.61	-1.20	-0.23
39	5.69	108.91	177.07	238.57	99.38	133.90	815.71	12422.21	0.70	-1.12	-0.23
40	4.89	92.27	172.58	214.01	95.49	118.41	800.28	11493.93	0.59	-1.30	-0.11
41	4.36	106.05	170.66	226.89	93.21	123.92	819.37	11613.58	0.73	-1.07	-0.24
42	4.04	71.22	175.45	191.15	95.46	104.00	836.27	12034.21	0.72	-1.03	-0.29
43	4.19	99.22	174.09	223.87	92.29	118.68	839.47	11839.07	0.70	-1.10	-0.23
44	3.52	99.99	178.28	230.15	99.09	127.91	917.93	13671.94	0.61	-1.24	-0.20
45	3.94	105.20	167.72	222.08	92.86	122.95	941.20	13128.18	0.63	-1.24	-0.18
46	4.95	104.41	175.19	231.10	97.90	129.14	863.77	13060.69	0.69	-1.16	-0.21
47	3.96	109.15	165.02	222.57	93.54	126.16	876.81	12339.17	0.49	-1.42	-0.12
48	3.53	108.10	162.29	217.83	107.06	143.70	785.73	11953.75	0.57	-1.49	0.16
49	3.05	103.40	180.17	236.53	105.85	138.95	771.30	12477.60	0.74	-0.99	-0.38
50	2.89	101.00	177.76	230.63	101.20	131.29	793.97	12353.34	0.67	-1.13	-0.29
51	2.82	106.03	175.00	232.64	100.23	133.25	756.56	11671.82	0.79	-0.88	-0.43
52	3.27	104.86	181.25	239.61	104.02	137.51	814.91	13022.51	0.63	-1.22	-0.21
53	3.22	119.72	168.35	237.81	98.16	138.66	592.06	8013.48	0.79	-0.97	-0.25
54	2.64	118.60	156.03	219.36	83.14	116.88	857.47	11059.48	0.91	-0.82	-0.39
55	2.36	106.62	175.29	233.67	100.33	133.75	861.59	13184.17	0.64	-1.19	-0.24
56	2.34	104.67	166.65	220.11	91.07	120.28	712.46	10032.95	0.59	-1.29	-0.20
57	3.48	106.81	166.71	222.43	88.85	118.54	891.96	12262.15	0.51	-1.44	-0.04
60	3.22	108.40	173.27	232.90	93.37	125.51	852.99	12092.58	0.64	-1.17	-0.23
61	3.25	110.63	175.55	238.38	99.61	135.26	841.91	12443.26	0.69	-1.18	-0.17
62	4.56	117.17	164.55	229.95	92.19	128.83	865.94	11986.80	0.68	-1.20	-0.17
63	5.28	100.24	146.80	189.74	79.61	102.90	984.53	12115.43	0.42	-1.55	0.00
64	5.74	109.32	171.90	232.04	98.28	132.66	859.22	12925.17	0.66	-1.17	-0.21

Appendix D: Selected publications

The following papers that are based on material from this thesis are included in this appendix.

1. C Alexakis, HO Nyongesa, R Saatchi, ND Harris, C Davies, C Emery, RH Ireland, SR Heller, "Feature Extraction and Classification Of Electrocardiogram (ECG) Signals Related to Hypoglycaemia", Computers in Cardiology 2003, vol 30, pp 537-540, Thessaloniki, Greece, 21-24 September 2003
2. HO Nyongesa, C Alexakis, R Saatchi, M Rodrigues, ND Harris, SR Heller, "Classification Of SAECG By Auto-Regressive Modelling And Neural Networks", IEEE Africon 04, Gaborone, Botswana, 15 – 17 September 2004
3. C Alexakis, M Rodrigues, R Saatchi, HO Nyongesa, ND Harris, SR Heller, "Detection Of Abnormal Electrocardiogram Traces, Related To Hypoglycaemia, Using A Knowledge-Based Approach", Postgraduate Research Conference in Electronics, Photonics, Communications & Networks, and Computing Science (PREP) 2005, University of Lancaster, 30 March - 1 April 2005

Feature Extraction and Classification of Electrocardiogram (ECG) Signals Related to Hypoglycaemia

C Alexakis[1], HO Nyongesa[1], R Saatchi[1], ND Harris[2], C Davies[3], C Emery[3], RH Ireland[4], SR Heller[3]

[1]Schools of Computing & Engineering, Sheffield Hallam University, Sheffield, UK
[2]University of Bath, Bath, UK, [3]University of Sheffield & [4]Weston Park Hospital, Sheffield, UK

Abstract

Nocturnal hypoglycaemia has been implicated in the sudden deaths of young people with diabetes. Experimental hypoglycaemia has been found to prolong the ventricular repolarisation and to affect the T wave morphology. It is postulated that abnormally low blood glucose could in certain circumstances, be responsible for the development of a fatal cardiac arrhythmia.

We have used automatic extraction of both time-interval and morphological features, from the Electrocardiogram (ECG) to classify ECGs into normal and arrhythmic. Classification was implemented by artificial neural networks (ANN) and Linear Discriminant Analysis (LDA). The ANN gave more accurate results. Average training accuracy of the ANN was 85.07% compared with 70.15% on unseen data.

This study may lead towards the demonstration of the possible relationship between cardiac function and abnormally low blood glucose.

1. Introduction & background

The aim of this work is to detect the onset of nocturnal hypoglycaemia indirectly through analysis of the Electrocardiogram (ECG) of type 1 diabetics. In order to achieve this, ECG feature extraction is performed and the features produced are classified according to their corresponding glucose levels.

Nocturnal hypoglycaemia has been implicated in the sudden deaths of diabetics, especially those of an early age, a syndrome known as "Dead in Bed" [1]. The mechanism and cause of such deaths is still not very clear. The diabetics were well the night before and were found dead in an undisturbed bed the following morning. There was no brain damage, a symptom of profound hypoglycaemia, hence the deaths were caused by a different mechanism. It is suspected that deaths were caused by a fatal cardiac arrhythmia. It has been shown that experimental hypoglycaemia prolongs the ventricular repolarisation (VR) and hence it affects the rythmicity of the heart [2].

The 3-lead ECG was used for the purposes of this research. A typical ECG cycle is presented in figure 1. The T wave corresponds to the ventricular repolarization of the myocardium. During hypoglycaemia, the counter-regulatory responses cause the release of adrenaline and a fall in potassium, which delays repolarisation. These changes may be reflected on the ECG by changes in T wave morphology. If these changes can be automatically identified it may provide a warning of hypoglycaemia or of a potentially pro-arrhythmogenic condition.

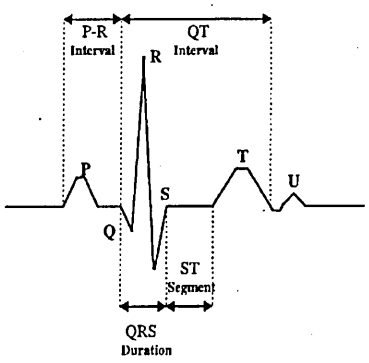


Fig 1: a typical ECG

2. Methods

2.1. Data acquisition

The data used in this study consisted of both the ECG traces and their corresponding blood glucose levels. They were obtained from eleven type 1 diabetic patients, with mean (sd) age 35.9 (14.53), recruited by the Diabetic Clinic of the Royal Hallamshire Hospital in Sheffield. The ECG data were recorded in the patient's own environment by a custom-built system that captures data from the YY' orthogonal lead [3]. One-minute worth of recording was captured every 15 minutes. Blood glucose was recorded by an implantable glucose sensor (MiniMed CGMS) [4] that measures glucose in the trancutaneous tissue every 5 minutes. The above acquisition was carried out for two successive nights and produced a data-set of paired ECG-glucose readings. This data-set was used for offline feature extraction and classification.

2.2. ECG features

An illustration of the time-interval features that can be extracted from an ECG cycle is given in figure 1. The QT was considered in this study since it describes the duration of VR. Correction of QT for heart rate was carried out using Bazett's formula ($QTc = QT/\sqrt{RR}$) [5]. Besides the time interval features, other features describing the amplitude, morphology or area of certain waves were considered.

Five ECG features were used in this study namely: RR, RTc, T wave amplitude (Tampl), T wave skewness (skew) and T wave kurtosis (kurt). These features were extracted using automatic algorithms. The onset and end of the T wave were detected using the tangent method [6,7].

RT is the time interval from the R peak to the end of the T wave. RTc is the corrected version using Bazzett's formula. The RT interval was chosen for this study, instead of the QT, since R point detection is more robust than Q point detection especially in the presence of noise. Moreover the RT interval still describes the process of ventricular repolarisation satisfactorily. RT has been used before [8] but to a lesser extend than the QT.

Skewness is used to evaluate the symmetry of the T wave shape. Kurtosis is used to quantify the degree of peakedness of the T wave shape. Traditionally skewness and kurtosis are used to evaluate the symmetry and peakedness of statistical distributions but in this study they are used for the shape analysis of the T waveform [9].

2.3. Neural network classification

Artificial Neural Networks (ANN) are computational models inspired by the functioning of the human brain. They consist of simple but highly interconnected computing devices, each of which imitates the biological neuron. The ANN "learns" by adapting connections between its computational neurons to match input-output combinations.

The neural network architecture used in this study for classifying ECG traces was the Multi-Layer Perceptron (MLP). Classification was binary, into normal and arrhythmic (corresponding to hypoglycaemia) ECG records. The ANN mapped normal ECG records as negative and arrhythmic ones as positive. A threshold of 2.5 mmol/lt was used to distinguish euglycaemia from hypoglycaemia. ECG traces corresponding to glucose equal or below 2.5 mmol/lt were classed as arrhythmic (hypoglycaemic) while those corresponding to the glucose interval [4 8] mmol/lt were classed as normal (euglycaemic). Records belonging to the interval (2.5 4) were excluded since they belong in the transition region between the normal and the hypoglycaemic class. Hyperglycaemic records (defined as: glucose> 8mmol/lt) were also excluded.

The 5 ECG features produced were combined in two combinations of 4 features namely RR, RTc, Tampl, skew and RTc, Tampl, skew, kurt. Apart from the above features, a third combination was considered. It consisted of a total of 10 ECG features, including the above 5. The extra 5 features were: RT, Tduration, corrected Tduration (using Bazett's formula), area under T and ratio of areas under T on either side of T peak. These 10 features were preprocessed using Principal Component Analysis (PCA) to produce an orthogonalised set of features and reduce the dimensionality of the input vector (i.e. the number of features used). Any feature with less than 2% contribution to the variation in the data set was discarded by the PCA algorithm. PCA typically reduced the 10 initial features into 4 or 5 orthogonalised features. Neural networks were trained using the above three feature combinations and comparisons were made in order to identify the best one.

A classifier was trained for each patient considered in the study. Alternatively a single classifier could have been trained to work on all patients. The second approach was not preferred because of inter-patient variability problems. Such variability is common when dealing with physiological data, making it difficult for the classifiers to generalise on unseen data, across the population of all patients. Some parameters that are typically varying across patients are: age, sex, duration of diabetes, level of glycaemic control, fitness level etc. By allowing a classifier to focus on the dynamics of a single patient the problem of inter-patient variability is overcome and the only problem we are faced with is that of intra-patient variability.

Producing a classifier for each patient means that each classifier only sees data from a single patient. This introduces the problem of small data-sets since the data has to be partitioned per patient. In order to maximise the data available five-fold cross-validation was applied and the results were averaged over 5. Data-sets consisted of a maximum of 66 feature vectors, each vector consisting of four (or more for PCA) features. Since the length of the data-sets was short, the size (number of neurons) of the ANNs was kept small to avoid overfitting. A maximum of 5 neurons was used in the single hidden layer. For the same reason, the number of input ECG features was limited to 4 although more features were available. The preprocessing of the features included removal of outliers (using the mean ± 2sd criterion) and normalisation in the interval [-1 1].

The performance measures used to evaluate the performance of the classifiers were: accuracy, hitrate (sensitivity), false-alarm-rate, true-negatives-ratio (tnratio) and missed-hypos (false-negatives ratio). They are defined as:

- $Accuracy = tp + tn / (tp+tn+fp+fn)$ *(1)*
- $Hitrate = tp / (tp +fn)$ *(2)*

- False-alarm-rate = fp / (fp +tn) (3)
- Tnratio = tn / (tn + fp) (4)
- Missed-hypos = fn / (fn + tp) (5)

where *tp*, *tn*, *fp* and *fn* stand for: true positives, true negatives, false positives and false negatives respectively. Positive refers to hypoglycaemia while negative refers to euglycaemia.

Hitrate describes the number of arrhythmic traces classified correctly while false-alarm-rate describes the number of normal traces that were classified as arrhythmic (i.e. false alarms). Tnratio describes the number of normal traces classified correctly while missed-hypos describes the number of arrhythmic traces classified as normal, i.e. the number of hypoglycaemic events that were missed.

3. Results

Neural network classification results for the 11 subjects and for features RTc, Tampl, skew, kurt are given in table 1. The table contains performance measures for both the training and testing datafiles.

Table 1: ANN classification results (RTc Tampl skew kurt)

	TRAIN					TEST				
patient	accuracy (%)	hitrate (%)	false alarm (%)	tnratio (%)	missed hypos (%)	accuracy (%)	hitrate (%)	false alarm (%)	tnratio (%)	missed hypos (%)
202	89.82	100.00	3.53	89.26	0.00	71.52	73.85	24.29	74.29	24.75
203	93.78	98.46	10.90	89.10	1.54	87.50	90.46	26.67	73.33	9.54
204	77.08	100.00	0.00	77.22	0.00	58.33	62.00	37.00	58.67	34.00
208	88.30	94.86	3.50	90.33	4.80	66.00	71.00	29.00	66.00	26.67
212	83.50	100.00	0.00	79.00	0.00	77.66	85.45	16.67	83.33	12.73
216	79.15	93.89	29.26	70.07	5.88	76.82	85.61	57.00	39.33	14.13
220	83.89	97.78	0.00	83.89	2.22	65.19	70.87	16.67	50.00	27.14
223A	82.19	96.19	0.95	79.76	3.81	69.11	84.00	21.90	68.33	13.33
227	93.17	100.00	4.86	89.33	0.00	62.00	65.86	27.00	68.67	30.17
229	78.21	81.88	6.41	78.21	17.18	64.19	60.00	29.69	65.14	38.33
244	86.67	100.00	0.00	86.67	0.00	73.28	80.60	28.00	58.00	18.60
mean	85.07	96.64	5.40	82.99	3.22	70.15	75.43	29.72	64.10	22.67
std	5.79	5.38	8.63	6.59	5.10	8.36	10.41	10.95	12.17	9.58
min	77.08	81.88	0.00	70.07	0.00	58.33	60.00	16.67	39.33	9.54
max	93.78	100.00	29.26	90.33	17.18	87.50	90.46	57.00	83.33	38.33

To allow comparisons, Linear Discriminant Analysis (LDA) was also used for classification of the ECG records in normal and arrhythmic. LDA works by minimising the Mahalanobis distance [10] which is a multivariate measure of the separation of a data set from a point in space. The same ECG features that were fed into the ANN were used in LDA. Five-fold cross-validation was applied and the results were averaged over 5. Partitioning of the data into training and testing was exactly the same as for the ANN. The classification results for RTc, Tampl, skew, kurt obtained from LDA are tabulated in table 2.

Overall the ANN were superior to the LDA. The weakest point of LDA was the percentage of missed-hypos. This ratio was high even for the training data-set.

For both the ANN and the LDA, the hitrate was greater than the tnratio for both training and test results. This means that both classifiers were better in classifying hypoglycaemic records correctly than in classifying normal records correctly.

Table 2: LDA Classification results (RTc Tampl skew kurt)

	TRAIN					TEST				
patient	accuracy (%)	hitrate (%)	false alarm (%)	tnratio (%)	missed hypos (%)	accuracy (%)	hitrate (%)	false alarm (%)	tnratio (%)	missed hypos (%)
202	83.42	86.99	20.15	79.85	13.01	69.62	72.58	37.14	62.86	27.42
203	91.35	92.05	9.36	90.64	7.95	82.92	86.52	34.67	65.33	13.48
204	70.56	66.11	25.00	75.00	33.89	50.67	26.67	25.33	74.67	73.33
208	73.40	70.17	23.37	76.63	29.83	63.67	46.67	19.33	80.67	53.33
212	100.0	100.00	0.00	100.0	0.00	89.96	92.00	60.00	40.00	8.00
216	77.28	82.28	27.72	72.28	17.72	69.88	71.67	36.67	63.33	28.33
220	68.89	90.83	53.06	46.94	9.17	87.10	89.77	100.00	0.00	10.23
223A	64.90	76.86	47.05	52.95	23.14	56.25	84.76	65.28	34.72	15.24
227	65.95	51.14	19.24	80.76	48.86	44.09	39.33	41.33	58.67	60.67
229	79.68	87.18	27.82	72.18	12.82	36.33	45.00	64.49	35.51	55.00
244	87.78	97.78	22.22	77.78	2.22	83.44	90.68	61.33	38.67	9.32
mean	78.47	81.94	25.00	75.00	18.06	66.72	67.79	49.60	50.40	32.21
std	11.24	14.75	14.92	14.92	14.75	18.24	23.95	23.18	23.18	23.95
min	64.90	51.14	0.00	46.94	0.00	36.33	26.67	19.33	0.00	8.00
max	100.0	100.00	53.06	100.0	48.86	89.96	92.00	100.00	80.67	73.33

For some patients the test figures for false-alarm-rate (for both the ANN and the LDA) were extremely high while the accuracy and hitrate were also high. This can be understood by looking at equations 1-3 in the previous section. If in the data-set there exist very few *tn* compared to the number of *fp* the false-alarm-rate will be high. At the same time the accuracy and hitrate can be high if *tp* is much higher than *fp* and *fn*. If the data-sets were sufficiently large there would not be such a problem.

4. Discussion

Classification of ECG traces was carried out by MLP and LDA. Both are supervised classification methods but the way they work is not the same. The LDA is a linear

statistical classifier while the MLP is non-linear. Both types of classifiers had reasonable performance with the MLP performing better than the LDA. Longer data-sets will be necessary for obtaining a clearer picture of the differences in performance of the two classifiers.

The three feature combinations used had very similar performance when considering the average performance metrics. Looking at individual patients, the three feature combinations did not have systematic performance for the various patients.

For the given data-sets and input features, the performance of LDA cannot, because of its nature, be improved further. However, in the neural network case the performance could be further improved. Many different parameters are involved which have not been explored fully. By tuning the parameters better classification performance could be possible.

5. Conclusion

This paper focused on automatic feature extraction and classification of ECG signals for detection of the delayed ventricular repolarisation, a cardiac arrhythmia that is suspected to be introduced by hypoglycaemia. ECG features were used that describe both the duration and morphology of the relevant ECG components. Classification was carried out using multi-layer perceptrons and statistical classifiers (LDA). The two types of classifiers performed quite closely to each other, with the ANN being more accurate. The ANN can be further improved to achieve even better performance, because of the nature of its architecture being multi-parametric. It is suspected that the optimal neural network recipe has not been found yet.

Future work will focus on improving the ANN classification and also on experimenting with other feature combinations and probably the introduction of new features. Non-linear PCA may be used instead of PCA in order to, more effectively, reduce the dimensionality of the input space. Fuzzy logic will also be considered in order to offer transparency to the classification process.

Regarding data acquisition, data sets from adolescent and prepubescent type 1 diabetic patients will be used in the near future. The incidence of sudden death is highest in young people or those with a short duration of diabetes and these data may show more pronounced changes.

Acknowledgements

Data Collection is supported by MiniMed Inc.

Charilaos Alexakis is supported by a Digital Techonologies Research Centre (DTRC) Studentship from Sheffield Hallam University.

References

[1] Campbell IW. Dead in bed syndrome: a new manifestation of nocturnal hypoglycaemia? Diabetic Medicine 1991; 8: 3-4.

[2] Marques JLB, George E, Peacy SR, Harris ND, MacDonald IAM, Cochrane T, Heller SR. Altered ventricular repolarisation during hypoglycaemia in patients with diabetes. Diabetic Medicine 1997; 14: 648-654.

[3] Harris ND, Ireland RH, Marques JLB, Hudson S, Davies C, Lee S, Robinson RT, Heller SR. Can changes in QT interval be used to Predict the Onset of Hypoglycaemia in Type 1 Diabetes? Computers in Cardiology 2000; 27: 375-378.

[4] MiniMed website: http://www.minimed.com/doctors/md_products_cgms_ov_completepic.shtml

[5] Ahnve S. Correction of the QT interval for heart rate: review of different formulas and the use of Bazett's formula in myocardial infarction. Am Heart J 1985; 109: 568-574.

[6] Ireland RH, Robinson RTCE, Heller SR, Marques JLB, Harris ND. Measurement of high resolution ECG QT interval during controlled euglycaemia and hypoglycaemia. Physiol Meas 2000; 21: 295-303.

[7] McLaughlin NB, Campbell R WF, Murray A. Comparison of automatic QT measurement techniques in the normal 12 lead electrocardiogram Br.Heart J. 1995; 74: 84-89.

[8] Porta A, Baselli G, Lombardi F, Cerutti S, Antolini R, Del Greco M, Ravelli F, Nollo G. Performance assessment of standard algorithms for dynamic R-T interval measurement: comparison between R-Tapex and R-T(end) approach. Medical & Biological Engineering & Computing 1998; 36: 35-42.

[9] Alexakis C, Nyongesa HO, Saatchi R, Harris ND, Davies C, Emery C, Heller SR. Detection Of Hypoglycaemia-Induced Cardiac Arrhythmias Using Neural and Statistical Classifiers. Neural Networks and Expert Systems in Medicine 2003;

[10] MathWorks Inc. Statistics Toolbox for Use with MATLAB 1999;

Address for correspondence:

Charilaos Alexakis
School of Engineering, Sheffield Hallam University
City Campus, Sheffield, England, S1 1WB
Email: c.alexakis@shu.ac.uk

CLASSIFICATION OF SAECG BY AUTO-REGRESSIVE MODELLING AND NEURAL NETWORKS

HO Nyongesa (nyongesa@mopipi.ub.bw),
University of Botswana.

C Alexakis, , R Saatchi, M. Rodrigues
Sheffield Hallam University, UK.

ND Harris,
University of Bath

SR Heller, C Davies, C Emery, RH Ireland
University of Sheffield, UK

Abstract

The paper describes investigations into the classification of signal-averaged electro-cardiogram (SAECG) signals, with regard to detection of the onset of hypoglycaemia in diabetic patients. Firstly, feature extraction is carried out to obtain time-domain features, which are classified by neural networks. Secondly, the SAECG signals are modelled by Autoregressive modelling (AR), and the parameters classified using Linear Discriminant Analysis. The classification performances using both approaches are compared. ECG datasets were obtained from ongoing related research, and consist of paired ECG-Glucose readings from Type-1 diabetic patients. Data was recorded overnight in the patient's own homes.

1 INTRODUCTION

The ECG signal describes the electrical activity of the heart and is among the most widely used physiological signals. In this paper, the study is related to detection of hypoglycaemia in Type-1 diabetes patients. Hypoglycaemia is a condition, among mostly diabetic patients, where the blood glucose drops to abnormal levels. Nocturnal hypoglycaemia has been implicated in the sudden deaths of diabetic patients, especially those of an early age, a syndrome known as "Dead in Bed" [1]. The mechanism and cause of such deaths is still not very clear. The patients were well the night before and were found dead in an undisturbed position the following morning. There was no brain damage, a symptom of hypoglycaemia, suggesting that the deaths were caused by a different mechanism. It is suspected that deaths were caused by a fatal cardiac arrhythmia.

2 DATA

Available datasets were obtained from other diabetes related research. They consist of paired ECG-Glucose readings from 11 Type-1 diabetic patients and have been analysed in the present research to investigate any symptomatic manifestation of hypoglycaemia within the ECG. Data was recorded overnight in the patient's own homes in order to capture spontaneous hypoglycaemic events. One-minute worth of Y lead ECG recording was captured every 15 minutes. Each SAECG cycle was produced by averaging over the 1-minute time period. ECG acquisition was carried out using a custom-made system (Hypoglycaemia Online Monitoring Ensemble) [2] developed in the Royal Hallamshire Hospital in Sheffield. The sampling rate was 125 Hz and was limited by the specification of the palmtop computer used. Glucose measurements were acquired using MiniMed CGMS [3].

3 METHODS

Two approaches were followed in representing the ECG. Individual ECG features, both time domain and morphological, that describe certain components or processes of the cardiac function were used as the first approach. In the second approach, the whole of the relevant segment of the ECG was described via modelling by Auto-Regressive coefficients. The ECG feature extraction results were classified using Artificial Neural Networks (ANN) and Linear Discriminant Analysis (LDA) while the AR modelling results were only fed to the statistical classifiers (LDA). Comparison of the two ECG representation methods and also the two classification approaches are presented.

3.1 ECG Feature Extraction

Five ECG features were used in this study namely: RR, RTc, T wave amplitude (Tampl), T wave skewness (skew) and T wave kurtosis (kurt). These features were extracted using automatic algorithms. The onset and end of the T-wave were detected using the tangent method [4]. RT is the time interval from the R peak to the end of the T wave. RTc is the corrected version using Bazzett's formula [5]. The RT interval was chosen for this study, instead of the QT, since R point detection is more robust than Q point detection especially in the presence of noise. Moreover the RT interval still describes the process of ventricular repolarisation satisfactorily. RT has been used before but to a lesser extend than the QT.

Skewness is used to evaluate the symmetry of the T wave shape. Kurtosis is used to quantify the degree of peakedness of the T wave shape. Traditionally skewness and kurtosis are used to evaluate the symmetry and peakedness of statistical distributions but in this study they are used for the shape analysis of the T waveform.

3.2 ECG Modelling by AR Coefficients

This approach has been used before for detection of certain cardiac arrhythmias. In this study it was used for the detection of the delayed ventricular repolarisation often exhibited during hypoglycaemia.

The general form of an n^{th} order AR model is:

$$y(k+1) = -\sum_{i=0}^{i=n} a_i \cdot y(k-i) + e(k+1) + offset \qquad (1)$$

where e(k) is the noise parameter and "offset" denotes an offset parameter.

The ECG section to the right of the *R* peak and until the end of the trace is used because this is the section affected by hypoglycaemia.

The Least Squares (LS) algorithm was used to find the estimates of the optimal model parameters. It works by minimising the sum of the squares of the model errors and is given by:

$$\hat{B} = (\Phi^T \Phi)^{-1} \Phi^T Y \qquad (2)$$

Vector B contains the model parameters (a_i and offset) and vector Y contains the data-points describing the ECG trace.

Figure 1 illustrates a SAECG trace together with its modelled version using a 2rd order AR model. The whole ECG cycle is plotted (blue) but only the post-R peak section is modelled (dotted).

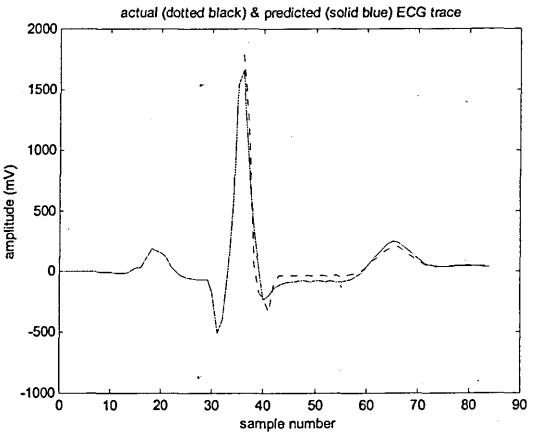


Figure 1: actual (solid) vs modelled (dotted) ECG

The actual ECG modelling was carried out using a 3rd order AR model. The correlation coefficient between actual and modelled ECG trace was for all but one patient greater than 91% with an average of 95% across patients. The model order can be increased so that each ECG trace is more closely modelled but this will produce extra model parameters that the classifiers have to handle and classify. Emphasis is placed on making the model simple and hence keeping the classification task simple.

3.3 Classification

Data from individual patients were classified, one at a time, in order to avoid inter-patient variability problems. Such variability is common when dealing with physiological data, making it difficult for the classifiers to generalise on unseen data, across the population of all patients. Some parameters that are typically varying across patients are: age, sex, duration of diabetes, level of glycaemic control, fitness level etc. By allowing a classifier to focus on the dynamics of a single patient the problem of inter-patient variability is overcome and the only problem we are faced with is that of intra-patient variability.

Partitioning data per patient means that each classifier only sees data from a single patient which introduces the problem of small data-sets. In order to maximise the data available five-fold cross-validation was applied and the results were averaged over 5.

Classification was binary, into normal and arrhythmic (corresponding to hypoglycaemia) ECG records. The classifiers mapped normal ECG records as negative and arrhythmic ones as positive. A threshold of 2.5 mmol/lt was used to distinguish euglycaemia (normal) from hypoglycaemia. ECG traces corresponding to glucose equal or below 2.5 mmol/lt were classed as arrhythmic (hypoglycaemic) while those corresponding to the glucose interval [4 ... 8] mmol/lt were classed as euglycaemic. Records belonging to the interval [2.5 ... 3.5] were excluded since they belong in the transition region between the normal and the hypoglycaemic class. Hyperglycaemic records (defined as: glucose > 8mmol/lt) were also excluded.

The 5 ECG features produced were combined in two combinations of 4 features namely "RR, RTc, Tampl, skew" and "RTc, Tampl, skew, kurt". Neural networks were trained using both feature combinations and comparisons were made in order to identify the best one. The neural network architecture used in this study for classifying ECG traces was the Multi-Layer Perceptron (MLP).

Data-sets consisted of a maximum of 66 feature vectors, each vector consisting of four features. Since the length of the data-sets was short, the size (number of neurons) of the ANNs was kept small to avoid overfitting. A maximum of 5 neurons was used in the single hidden layer. For the same reason, the number of input ECG features was limited to 4 although more features were available. The preprocessing of the features included removal of outliers (using the mean ± 2sd criterion) and normalisation in the interval [-1 1].

The performance measures used to evaluate the performance of the classifiers were: accuracy, hitrate (sensitivity), false-alarm-rate, true-negatives-ratio (tnratio) and missed-hypos (false-negatives ratio). They are defined as:

- Accuracy = tp + tn / (tp+tn+fp+fn) (1)
- Hitrate = tp / (tp +fn) (2)
- False-alarm-rate = fp / (fp +tn) (3)
- Tnratio = tn / (tn + fp) (4)
- Missed-hypos = fn / (fn + tp) (5)

where tp, tn, fp and fn stand for: true positives, true negatives, false positives and false negatives respectively. Positive refers to hypoglycaemia while negative refers to euglycaemia.

Hitrate describes the number of arrhythmic traces classified correctly while false-alarm-rate describes the number of normal traces that were classified as arrhythmic (i.e. false alarms). Tnratio describes the number of normal traces classified correctly while missed-hypos describes the number of arrhythmic traces classified as normal, i.e. the number of hypoglycaemic events that were missed.

LDA was used to classify the two ECG feature combinations and also the coefficients produced by AR modelling. LDA works by minimising the Mahalanobis distance, which is a multivariate measure of the separation of a data set from a point in space.

4 RESULTS

Results from 11 Type-1 diabetic patients along with summary statistics are presented in the following tables.

Table 1: ANN classification results (RTc Tampl skew kurt)

patient	TRAIN					TEST				
	accuracy (%)	hitrate (%)	false alarm (%)	tnratio (%)	missed hypos (%)	accuracy (%)	hitrate (%)	false alarm (%)	tnratio (%)	missed hypos (%)
202	89.82	100.00	3.53	89.26	0.00	71.52	73.85	24.29	74.29	24.75
203	93.78	98.46	10.90	89.10	1.54	87.50	90.46	26.67	73.33	9.54
204	77.08	100.00	0.00	77.22	0.00	58.33	62.00	37.00	58.67	34.00
208	88.30	94.86	3.50	90.33	4.80	66.00	71.00	29.00	66.00	26.67
212	83.50	100.00	0.00	79.00	0.00	77.66	85.45	16.67	83.33	12.73
216	79.15	93.89	29.26	70.07	5.88	76.82	85.61	57.00	39.33	14.13
220	83.89	97.78	0.00	83.89	2.22	65.19	70.87	16.67	50.00	27.14
223A	82.19	96.19	0.95	79.76	3.81	69.11	84.00	21.90	68.33	13.33
227	93.17	100.00	4.86	89.33	0.00	62.00	65.86	27.00	68.67	30.17
229	78.21	81.88	6.41	78.21	17.18	64.19	60.00	29.69	65.14	38.33
244	86.67	100.00	0.00	86.67	0.00	73.28	80.60	28.00	58.00	18.60
mean	85.07	96.64	5.40	82.99	3.22	70.15	75.43	29.72	64.10	22.67
std	5.79	5.38	8.63	6.59	5.10	8.36	10.41	10.95	12.17	9.58
min	77.08	81.88	0.00	70.07	0.00	58.33	60.00	16.67	39.33	9.54
max	93.78	100.00	29.26	90.33	17.18	87.50	90.46	57.00	83.33	38.33

Results are tabulated only for the "RTc, Tampl, skew, kurt" feature combination. The "RR RTc Tampl skew" feature combination gave slightly inferior results which are not included. Table 1 presents the ANN classification results for the "RTc, Tampl, skew, kurt" feature combination while Table 2 contains the LDA classification for

the AR coefficients. ANN studies for this approach are on-going.

5 DISCUSSION

The ANN classification results were superior to the LDA results, which is expected since ANNs are non-linear and more powerful classifiers. Comparing the LDA classification results for the individual ECG features and the AR coefficients it is observed that the AR modelling yielded better results. It is anticipated that classifying the AR coefficients using a neural classifier will enhance even more the classification performance. The results suggest that AR modelling is a better ECG representation technique compared to the morphological ECG features. A summary of the training (known data) and test (unseen data) accuracies are presented in the form of a column graph in figure 2.

Table 2: LDA Classification using 3rd order AR model

Patient	Known data					Unseen data				
	accuracy (%)	hitrate (%)	false alarm (%)	tmratio (%)	missed hypos (%)	accuracy (%)	hitrate (%)	false alarm (%)	tmratio (%)	missed hypos (%)
202	84.52	82.06	13.01	86.99	17.94	76.52	93.33	64.76	35.24	6.67
203	99.17	100.0	1.67	98.33	0.00	88.58	91.23	21.43	78.57	8.77
204	97.78	95.56	0.00	100.0	4.44	66.85	56.67	24.76	75.24	43.33
208	93.98	91.17	3.20	96.80	8.83	85.03	87.50	18.67	81.33	12.50
216	78.01	89.26	33.24	66.76	10.74	71.39	75.20	46.67	53.33	24.80
220	79.44	100.0	41.11	58.89	0.00	89.99	91.70	70.00	30.00	8.30
223A	83.67	86.57	19.24	80.76	13.43	65.20	86.00	64.29	35.71	14.00
227	82.26	82.71	18.19	81.81	17.29	59.75	64.95	55.33	44.67	35.05
229	82.82	73.46	7.82	92.18	26.54	52.86	31.67	44.43	55.57	68.33
244	93.19	88.61	2.22	97.78	11.39	52.77	47.06	10.67	89.33	52.94
mean	87.49	88.94	13.97	86.03	11.06	70.89	72.53	42.10	57.90	27.47
std	7.77	8.33	14.11	14.11	8.33	13.89	21.58	21.73	21.73	21.58
min	78.01	73.46	0.00	58.89	0.00	52.77	31.67	10.67	30.00	6.67
max	99.17	100.0	41.11	100.0	26.54	89.99	93.33	70.00	89.33	68.33

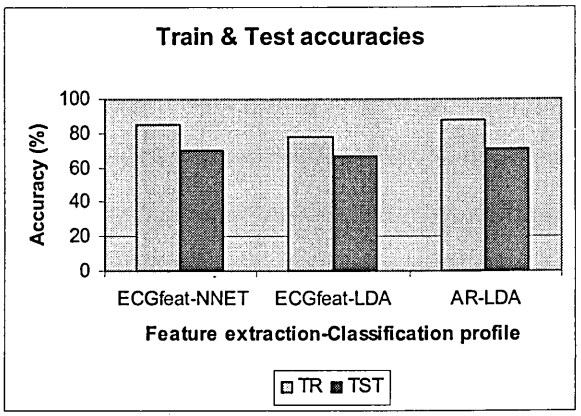


Figure 2: Classification accuracies for all three approaches

6 CONCLUSION

This paper focused on the investigation of two approaches of ECG trace representation and also two ways of classification of represented ECG traces. Previous work had focused mainly on using individual ECG features for detection of hypoglycaemia-induced arrhythmias. Recent work is showing that the use of AutoRegressive coefficients to represent segments of ECG cycles is a promising approach that has the advantage of describing the whole segment of interest as opposed to extracting features from it. Future work will involve the use of more sophisticated classifiers for the classification of AR results.

References

1. Campbell IW. Dead in bed syndrome: a new manifestation of nocturnal hypoglycaemia? Diabetic Medicine 1991; 8: 3-4.
2. Harris ND, Ireland RH, Marques JLB, Hudson S, Davies C, Lee S, Robinson RT, Heller SR. Can changes in QT interval be used to Predict the Onset of Hypoglycaemia in Type 1 Diabetes? Computers in Cardiology 2000; 27: 375-378.
3. MiniMed: www.minimed.com/doctors/
4. Ireland RH, Robinson RTCE, Heller SR, Marques JLB, Harris ND. Measurement of high resolution ECG QT interval during controlled euglycaemia and hypoglycaemia. Physiol. Meas. 2000; 21: 295-303.
5. Ahnve S. Correction of the QT interval for heart rate: review of different formulas and the use of Bazett's formula in myocardial infarction. Am Heart J 1985; 109:

DETECTION OF ABNORMAL ELECTROCARDIOGRAM TRACES, RELATED TO HYPOGLYCAEMIA, USING A KNOWLEDGE-BASED APPROACH

C Alexakis[1], M Rodrigues[1], R Saatchi[1], HO Nyongesa[2], ND Harris[3], SR Heller[4], C Davies[4], C Emery[4], RH Ireland[4]
[1]Schools of Computing & Engineering, Sheffield Hallam University, Sheffield
[2]Department of Computer Science, University of Botswana, Gaborone, Botswana,
[3]School for Health, University of Bath, [4]Department of Medicine, University of Sheffield,
contact email: c.alexakis@shu.ac.uk

Key words to describe the work: Electrocardiogram Signal (ECG), ECG Feature Extraction, Diabetes, Hypoglycaemia, Cardiac Arrhythmia, Dead in Bed Syndrome, Knowledge-Based Monitoring System, Expert System.

Key Results: Successful monitoring on offline data from diabetic patients corresponding to sensitivity and specificity of 100% and 91.30% respectively. Production of a rule-based system to represent knowledge on the relationship between spontaneous nocturnal hypoglycaemia and cardiac arrhythmia.

How does the work advance the state-of-the-art?: By developing a novel automatic hypoglycaemia monitoring method and providing a better understanding of the relationship between nocturnal hypoglycaemia and cardiac arrhythmia. By verifying and strengthening the hypothesis according to which, hypoglycaemia can be manifested on the ECG signal.

Motivation (problems addressed): Mechanism of Dead in Bed syndrome. Demonstration of relationship between spontaneous hypoglycaemia and cardiac function. Production of a non-invasive prototype monitoring system for hypoglycaemia detection and nocturnal death prevention in diabetic patients.

1 Introduction

Hypoglycaemia is the condition, experienced mostly by diabetic patients, according to which there is abnormally low glucose in the blood stream. Severe nocturnal hypoglycaemia has been implicated in the sudden death of diabetic patients, especially those of an early age, commonly known as "Dead in Bed Syndrome" [1]. The mechanism and cause of such deaths is still not very clear. It is suspected that deaths were caused by a fatal cardiac arrhythmia. It has been shown that experimental hypoglycaemia prolongs the ventricular repolarisation and hence affects the rythmicity of the heart [2].

The aim of this research is to detect abnormal Electrocardiogram (ECG) cycles occurring during hypoglycaemia and through this, to detect the onset of nocturnal hypoglycaemia indirectly through analysis of the diabetic's ECG.

2 Methods

2.1 Data Acquisition

The datasets used consisted of paired ECG-glucose readings obtained from Type 1 diabetic patients. Data was recorded overnight at the patient's own home. One-minute of ECG recording was captured every 15 minutes. Data from 19 patients, recorded over 32 nights were used to test the system.

2.2 Knowledge-Based Monitoring System

We have developed a prototype system to interpret the ECG signals. The system is designed to raise alarms if abnormal cardiac events, related to hypoglycaemia are detected. The system comprises an ECG pre-processor, a feature extractor and an Expert System (ES). The knowledge-base for the ES is a set of rules generated from observations of ECG changes under hypoglycaemia, within guidelines provided by clinical experts. The monitoring system is depicted in Figure 1.

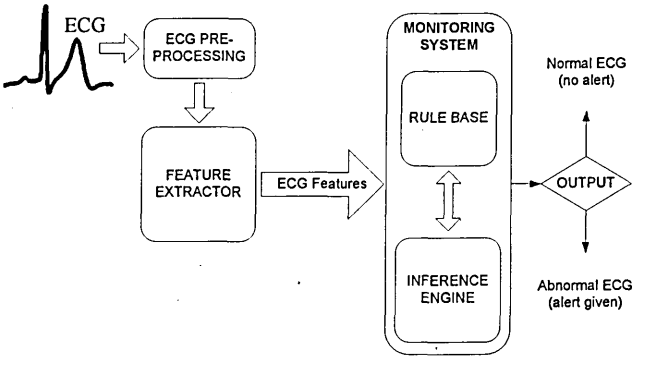


Figure 1: ECG Monitoring System

Monitoring is carried out on offline ECG data simulating an online monitoring scenario. At each sampling instant the ECG is fed to the pre-processing stage where a number of filtering steps are carried out. Next, the filtered ECG is passed to the feature extraction stage where a number of ECG features that describe the morphology and duration of the components of the current ECG cycle are extracted. The ECG features are then fed to the ES that infers, using the rule-base, whether they

correspond to a normal or abnormal ECG cycle. The system is using ECG features from the current, as well as previous ECG cycles to make a decision on whether to raise an alarm or not.

2.3 Assessment of Performance

Two definitions of True-Positives (TP) were used to assess the performance of the system:

i. Each hypoglycaemic night monitored, was assessed as TP if hypoglycaemia was detected at the exact time it occurred during the night.
ii. Each hypoglycaemic night monitored, was assessed as TP if hypoglycaemia was detected within an hour from the time it occurred during the night.

Each hypoglycaemic night monitored was assessed as False-Negative (FN) if hypoglycaemia was not detected, that is no alarm raised. Each euglycaemic (i.e. normal) night monitored correctly was a True-Negative (TN) and each euglycaemic night where a false-alarm was raised was a False-Positive (FP). After performing monitoring on all nights the sensitivity and specificity, over all nights, were calculated by the formulas:

- Sensitivity = TP / (TP + FN) (eq^n 1)
- Specificity = TN / (TN + FP) (eq^n 2)

Using the two different definitions for TP yields two pairs of results for sensitivity and specificity.

3 Results

Out of all nights, the system raised only two false-alarms, for two different nights, which corresponds to a specificity of 91.3%. The sensitivity was 55.56% if alarms were classed as acceptable, when they were produced at the exact time corresponding to the onset of hypoglycaemia (i.e. 1st TP definition). However, allowing alarms to deviate by up to an hour, either early or late, from the hypoglycaemic onset (i.e. 2nd TP definition) increased the sensitivity to 100%.

4 Discussion

The use of a knowledge-based monitoring system for detection of abnormal ECG traces related to hypoglycaemia proved to be a very promising approach. It contributes in strengthening the assumption according to which, the occurrence of hypoglycaemia is manifested on the ECG. In all hypoglycaemic nights, the system raised accurate alerts within 1 hour of the onset of hypoglycaemia. Regarding the euglycaemic nights, there were only two nights where a false-alarm was produced by the system.

In previous approaches we had employed Multi-Layer Perceptrons (MLP) and Linear Discriminant Analysis [3, 4] to perform ECG trace classification related to hypoglycaemia. Using a knowledge-based approach introduced a few advantages. The incorporation of human-expert knowledge allowed the system to focus on the significant ECG changes and ignore the unrelated ones. MLPs were confused by unrelated ECG changes and overcoming this would require very long datasets that were not available. Regardless of the above, the ES made better use of the dataset since all the data could be used to assess performance. In the case of MLP, a portion of the data had to be set aside for training and only the remaining data could be used for assessment of performance.

The ES being transparent also allows the investigation of its internal structure by clinical experts. On the contrary a trained MLP, being a black-box model does not allow easy investigation of its internal structure. A weakness of ES as compared to MLP is that it requires expert knowledge and this knowledge must be successfully coded.

5 Conclusion

This study focused on the design of an Expert System for overnight monitoring of diabetic patients and detection of abnormal ECG traces apparent under hypoglycaemia. Satisfactory performance of the system was achieved. Future work will involve further tuning while acquisition of more data for further investigation will be essential for longer-term continuation of the research.

References

[1] Campbell IW. Dead in bed syndrome: a new manifestation of nocturnal hypoglycaemia? Diabetic Med. 1991;8:3-4

[2] Marques JLB, et al (1997); Altered ventricular repolarisation during hypoglycaemia in patients with diabetes, Diabetic Medicine, 14, pp648-54.

[3] C Alexakis, HO Nyongesa, R Saatchi, ND Harris, C Davies, C Emery, RH Ireland, SR Heller, "Feature Extraction and Classification Of Electrocardiogram (ECG) Signals Related to Hypoglycaemia", Computers in Cardiology 2003, pp 537-540, Thessaloniki, Greece, 21-24 September 2003

[4] HO Nyongesa, C Alexakis, R Saatchi, M Rodrigues, ND Harris, SR Heller, "Classification of SA-ECG by LDA and NN", IEEE Africon 04, Gaborone, Botswana, 15 – 17 September 2004