ARTIFICIAL NEURAL NETWORKS IN COMPUTERISED ELECTROCARDIOGRAPHY

Ten-Fang Yang

MD MSc

Thesis submitted for the degree of Doctor of Philosophy of the University of Glasgow, Scotland.

To :

The Faculty of Medicine

University of Glasgow

The research described in this thesis was carried out in the University of Glasgow, Department of Medical Cardiology, Royal Infirmary, Glasgow, Scotland.

Ten-Fang Yang

Signed:

Date:

TABLE OF CONTENTS

CHAPTER 4

SOFTWARE BASED ARTIFICIAL NEURAL NETWORKS FOR THE DIAGNOSIS OF ATRIAL FIBRILLATION

LISTS OF TABLES

CHAPTER 3

CHAPTER 5

LIST OF FIGURES

ACKNOWLEDGEMENTS

I wish to thank Professor P.W. Macfarlane for his patient and instructive supervision and guidance during the past three years of my research in the field of Computerised Electrocardiography.

I am also indebted to Professor S.M. Cobbe, Professor A.R. Lorimer, Dr I Hutton, Dr A P Rae, and Dr A Rankin of the Department of Medical Cardiology and Consultants of the Department of Cardiac Surgery for their permission to use their patients' clinical data in the present research.

I particularly appreciate the assistance of Mr B. Devine in the field of neurocomputing and in setting up the computer programs.

I am extremely grateful to Mrs J. Watson, Dr S. McLaughlin, Mr S. Latif and the late Mr D. Shoat for their friendship and assistance in computer laboratory work and statistical techniques.

Finally, I wish to thank all the staff and patients in the University of Glasgow Department of Medical Cardiology for the kindness and cooperation shown to me during my stay in Glasgow.

Ten-Fang Yang

DEDICATION

In memory of my grandparents

CHYI-CHUNG YANG & PEN-CHANG YANG

without whose support and encouragement

I would not be a physician today.

To my brother and sisters

RU-ING, ING-FANG, & LING-ING

without whose psychological and financial assistance

it would not have been possible for me to stay abroad for

postgraduate study.

But most of all, to my parents

LI-HSIEN YANG & TSUI-HSIANG CHANG

for whose love and understanding

I can never adequately express my appreciation.

Finally, to my love

TSUO-SI LEE

without whose love and patience

this research could not have been accomplished.

AUTHOR'S DECLARATION

I hereby certify that this thesis has been researched and written entirely by myself and that it has not been submitted previously for any degree in other universities.

Ten-Fang Yang

SUMMARY

This thesis describes a series of studies on the use of artificial intelligence in computerised electrocardiography, the history of which is first of all reviewed. The relatively recent technique of using artificial neural networks, which aim to simulate the human decision making process, is thereafter introduced.

In a preliminary study, the vectorcardiogram derived from the twelve-lead electrocardiogram was studied for the first time in large populations of Caucasians and Chinese. A total of 2058 vectorcardiograms derived from conventional twelve-lead electrocardiograms recorded from 1555 healthy Caucasians and 503 healthy Chinese were analysed to establish the normal limits for Caucasians and Chinese, respectively.

Several conclusions can be drawn from this study:

(1) Age, sex and race dependent variations are present in the derived vectorcardiogram;

(2) In both races, the maximal spatial QRS vector magnitude, as well as the maximal QRS and T vector magnitude in the frontal, horizontal and right sagittal planes, decrease with advancing age in both sexes and are significantly larger in men in all age groups;

(3) In groups younger than 40 years, the magnitude of the maximal spatial QRS vector is greater in Caucasians than in Chinese, while in the groups older than 40 years, it is greater in Chinese than in Caucasians;

(4) This new data indicates that it is necessary to take racial differences into consideration for interpretation of the derived vectorcardiogram.

The first area in which neural networks were studied related to the analysis of cardiac rhythm. In particular, the usefulness of neural networks in separating atrial fibrillation from [sinus rhythm + (supraventricular extrasystoles &/or ventricular extrasystoles)] was investigated. For this purpose, 3080 visually classified ECGs including 2018 with atrial fibrillation and 1062 with [sinus rhythm + (supraventricular extrasystoles &/or ventricular extrasystoles)] were used. Fundamental to this work was the availability of the existing University of Glasgow ECG analysis program which was based on locally developed deterministic logic.

This study was divided into five stages:

(1) Selection of the optimum parameters for input to a neural network;
(2) Determination of the optimum topology of the network;
(3) Assessment of the accuracy of the network;
(4) Combining the deterministic logic result with the output from the neural network;
(5) Modification of the existing logic.

Several conclusions can be drawn:

(1) A neural network can improve the sensitivity but slightly decreases the specificity in detecting atrial fibrillation compared to the existing deterministic program;
(2) Various combinations of the neural network output and the deterministic interpretation are not superior to the use of a neural network alone;

(3) Modification of the existing deterministic logic can produce improved performance compared to the use of either a neural network or the original logic.

The use of neural networks was also intensively studied in the electrocardiographic diagnosis of myocardial infarction. A total of 1269 electrocardiograms [515 from patients with clinically documented myocardial infarction (255 anterior and 260 inferior), 144 from patients with clinically validated left ventricular hypertrophy and 610 from normals] were used to study the usefulness of the neural network approach. This study comprised a series of six experiments. Experiments 1, 2, 3 and 4 concerned the use of neural networks in isolation, while experiments 5 and 6 studied the effects of implanting neural networks into the deterministic program.

The conclusions drawn from experiments 1-4 on the use of neural networks in isolation for the diagnosis of myocardial infarction are that:

(1) There is no significant benefit from using derived vectorcardiographic measurements in addition to electrocardiographic QRS and ST-T parameters as input variables to the neural network;

(2) In the two-group situation (normal versus myocardial infarction), the neural network is superior to deterministic logic for the detection of myocardial infarction;

(3) In the diagnosis of anterior myocardial infarction, neural networks trained in the three-group situation (normal versus myocardial infarction versus left ventricular hypertrophy) perform better than those trained in the two-group situation, whereas in the diagnosis

of inferior myocardial infarction, there is no benefit from training in the three-group situation;

(4) The alteration of the network topology to include three output neurons rather than one does not lead to any improvement in the diagnosis of myocardial infarction;

(5) Neural networks using electrocardiographic QRS and ST-T measurements as input variables are superior to the neural networks using QRS measurements only;

(6) The enhanced sensitivity of detecting myocardial infarction by the neural network is more pronounced in inferior myocardial infarction than in anterior myocardial infarction;

(7) The specificity of diagnosing myocardial infarction by the neural network in the left ventricular hypertrophy cases decreases compared to the original logic no matter what training methodology is used.

Experiments 5 and 6 investigated the incorporation of neural networks into the deterministic logic together with some modification of the logic. Compared to the original logic, this resulted in a significant improvement of sensitivity in diagnosing inferior myocardial infarction (from 69% to 88% on a locally collected test set) and minimal improvement in anterior myocardial infarction (from 76% to 78% on a locally collected test set). A final test was made on a completely independent clinically developed database from the European Union supported project on Common Standards for Quantitative Electrocardiography (the CSE project). The new approach produced a relative 16% improvement in sensitivity in reporting inferior myocardial infarction (from 65.8% using the original program tested in 1992 on the CSE database to 76.2% on the same database) but no significant improvement in diagnosing anterior

myocardial infarction (72.6% versus 72.9% on the CSE database) and also resulted in a higher specificity for the diagnosis of myocardial infarction in left ventricular hypertrophy cases compared to using the neural network alone. The new program was also highly specific with respect to myocardial infarction when tested on the Chinese normals.

This study presents the first report of artificial neural networks being implanted into a deterministic program together with modifications to the logic for the diagnosis of myocardial infarction. This novel technique has the advantage of allowing existing deterministic criteria to be retained and enhanced as necessary with resultant diagnostic output produced in a form acceptable to clinicians.

It was also learned from the present study that it takes less time and experience to develop a well trained neural network to achieve an almost equivalent, or in some cases, a better performance compared to producing a specific section of deterministic logic for the diagnosis of atrial fibrillation or myocardial infarction.

The limitation of the software based neural network approach is closely correlated to the inherent limitation of electrocardiography in the diagnosis of myocardial infarction, whereas in the diagnosis of atrial fibrillation, there is no such problem because the gold standard of the rhythm diagnosis is the ECG wave form itself.

In conclusion, neural networks, if carefully designed and appropriately implanted into a deterministic program along with modifications to the logic, can selectively enhance the accuracy of computer-assisted ECG interpretation.

CHAPTER 1

ELECTROCARDIOLOGY

1.1 INTRODUCTION

Electrocardiology in its broadest sense is the study of the electric field generated by the individual cells of the heart. The term was first proposed by Lempert in 1976 (Lempert, 1976). It is just over one hundred years since the human electrocardiogram (ECG) was first recorded by a physiologist Augustus D. Waller at the St. Mary's Hospital in London U.K. (Burchell, 1987; Waller, 1887) and approximately sixty years since the unipolar electrocardiographic lead was introduced by Frank Wilson in the University of Michigan, U.S.A. (Wilson et al, 1934).

Electrocardiography has come under increasing pressure in recent years with the advent of new techniques such as echocardiography, nuclear perfusion scintigraphy, infarct-avid scan, and nuclear magnetic resonance radiography which undoubtedly provide information that complements the electrocardiogram. On the other hand, the Dutch physician H.J.J. Wellens (1986) regretted the fact that the younger generation of physicians is increasingly unable to interpret electrocardiograms correctly, even though electrocardiography is nowadays part of routine daily medical practice. Fisch (1980) also emphasized the fact that electrocardiography is a non-invasive technique which is relatively inexpensive and simple to use. The electrocardiogram also provides unique information that cannot be obtained by any other investigative

technique, e.g. ST-T changes and their prognostic importance in the various cardiovascular abnormalities.

With the advent of modern computer techniques, the electrocardiogram can be rapidly recorded and interpreted with portable equipment. Therefore, it is now more practical to use the electrocardiogram as an aid for clinical diagnosis. Much of the early work in electrocardiography was carried out in Europe, where today there are still considerable developmental efforts being expended, particularly in the field of computer-assisted electrocardiographic interpretation. On the basis of increasing numbers of electrocardiograms being recorded, there is a need for their automated interpretation. There are also other reasons for introducing automated analysis of ECGs. For instance, it is well known that there is considerable interobserver variation and bias in the interpretation of electrocardiograms. Frequently, this is due to inexperience on the part of physicians, especially when junior hospital doctors have to be involved in this task. On the other hand, cardiologists who are skilled in ECG reporting may differ in their interpretation of electrocardiograms because of using varying criteria to reach their diagnosis. Finally, the computer does not have the human emotional factor and distraction in the interpretation of electrocardiograms. It was hoped that the introduction of computer-assisted ECG interpretation would eliminate these deficiencies and, provided the programming had been accomplished in tandem with skilled and experienced cardiologists, could lead to a more uniform and higher standard of ECG interpretation.

1.2 HISTORY OF ELECTROCARDIOGRAPHY

The first galvanometer had been invented by the mid-nineteenth century and it was also generally agreed and proved around that period that nerves, muscles and so on could be stimulated by artificial electrical generators (Rijlant, 1980). At that time, physiologists were engaged in exploring the discharge from electric eels, the flow of electric current through frogs and the effects of injury. Such research work on the biologically generated electric current was probably initiated by Galvani but was criticised by Volta who thought that only different metals making contact could generate electric current. Volta's work led to the development of batteries (Snellen, 1984), while much of Galvani's research was continued by du Bois-Reymond, who found that the action current in muscle was opposite to the direction of a continuous current. This phenomenon was later demonstrated by his student Hermann to be present only following an injury to the muscle (Rijlant, 1980).

So far as is known, the first recording of cardiac electrical activity was performed in 1856 by Kolliker and Muller's demonstration of bioelectric potentials in the frog's heart (Kolliker & Muller, 1856). They described a negative deflection measured by a galvanometer prior to each contraction. Their experiments further confirmed the earlier research on bioelectric potentials of muscles from guinea pig and frog by du Bois-Reymond.

In 1876, a French physiologist Marey used the capillary electrometer invented by Lippmann to record the electric activity of the frog's heart on a photograph (Marey, 1876). Engelmann as well as Burdon Sanderson and Page were also among the earliest researchers to plot the potential

variation of electrical activity of a tortoise heart in 1878 (Burdon Sanderson & Page, 1878). It is evident that these series of developments finally led to the first known recording of the human electrocardiogram by Waller in 1887 [Fig 1-1].

Waller also investigated the possibility of recording potentials from the limbs of animals as well as from man. In 1889, he published further observations of the electrocardiogram recorded from his dog Jimmie using the capillary electrometer. It was stated by Willem Einthoven in 1912 that Waller first introduced the term "electrocardiogram" into physiological science (Einthoven, 1912). As a consequence of the investigations of A.D. Waller, research work began into the establishment of electrocardiography as a method of clinical examination.

According to Cooper (1986), a French engineer Ader invented the amplifier and developed a highly sensitive, rapidly moving galvanometer that used a small wire instead of a coil to register electrical potential. Dutch physician Willem Einthoven in the University of Leiden had been dissatisfied with the performance of the Lippmann galvanometer and instead he used the Deprez-d'Arsonval galvanometer. He also found that the sensitivity could be improved by replacing the coil with a single fibre (string). He subsequently described his new galvanometer and acknowledged the Ader galvanometer which also used a fine wire stretched between poles of a magnet. Einthoven's wire was 0.002 mm in diameter, approximately one tenth the thickness of that used by Ader (Cooper, 1986).

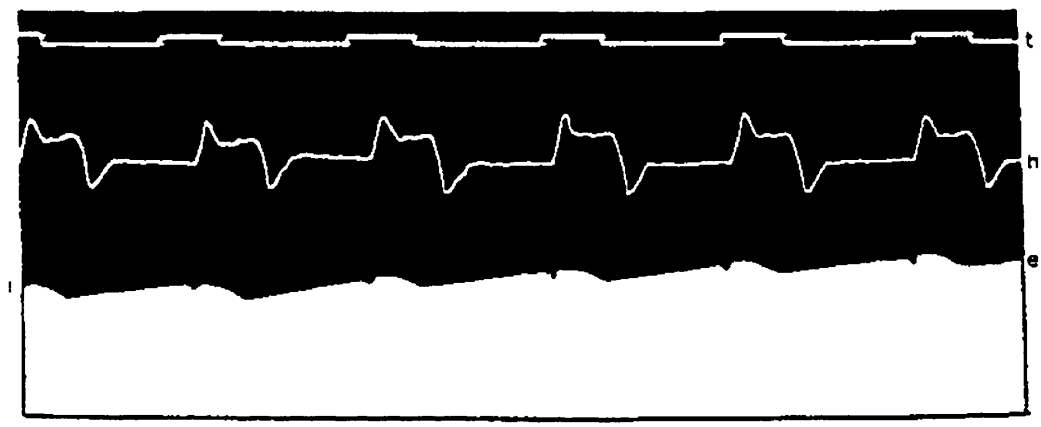


Fig. 1-1 The first human electrocardiogram recorded by Waller in 1887. t=time in seconds, e=electrocardiogram, h=chest wall movement. (After Waller 1887)

Einthoven's major achievement was to design a device that was sensitive enough to record cardiac electrical potentials from the surface of the body. He also introduced the concept known as "Einthoven's triangle" which is now classic, whereby the body was represented in electrical terms by an equilateral triangle, from which the mean QRS axis can be calculated (Einthoven, Fahr & de Waart, 1913). Einthoven also first used the terminology of P, Q, R, S, T to describe the deflection of the electrocardiogram [Fig 1-2] and in 1906 developed a method of so called "Telecardiography" for transmitting the electrocardiogram over telephone lines (Burchell, 1987). At that time, leads I, II, and III had been introduced and a variety of different electrocardiographic abnormalities had been demonstrated by Einthoven (1912).

The three limb leads, denoted I, II, and III can be represented as follows:

$$\begin{aligned} \text{I} &= E_L - E_R \\ \text{II} &= E_F - E_R \\ \text{III} &= E_F - E_L \end{aligned}$$

where E_L, E_R and E_F denote the potential at the left and right arms and left leg, respectively. It follows that

$$\text{II} = \text{I} + \text{III}$$

This is known as Einthoven's Law.

In 1907, electrocardiography was first introduced to clinics by Kraus and Nicolai at the Berlin Charite Hospital (Snellen, 1984). Einthoven's early work demonstrated to the medical profession that the electrocardiograph

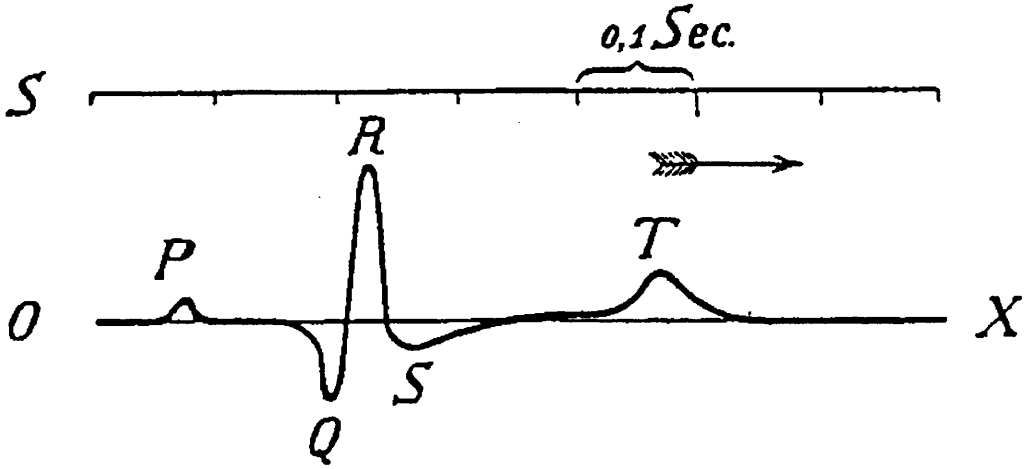


Fig.1-2 Definition of the deflections of the electrocardiogram by Einthoven (1903). S=time in 0.1 second, Lower tracing shows P,Q,R,S, & T deflection of the ECG.

was of practical, as well as theoretical importance. Einthoven also produced records of ventricular extrasystoles, ventricular bigeminy, and atrial fibrillation, although it was Hering and Lewis who examined atrial fibrillation in more detail (Snellen, 1984). In 1924, Einthoven was awarded the Nobel prize for his contribution to electrocardiography.

By the middle of the first decade of the twentieth century, polygraphic methods had been used to investigate the irregularity of the heart beat, the Einthoven galvanometer had also been developed, and the electrocardiogram had been investigated more extensively with respect to clinical correlation. Sir Thomas Lewis in London had used bipolar chest leads extensively in the studies of cardiac rhythmic disorders (Hollman, 1981; Snellen, 1981). He also contributed to the knowledge of the spread of electrical excitation through the heart. His measurements of epicardial activation established the hypothesis on the excitation of the myocardium.

According to Macfarlane and Lawrie (1989), the English edition of the book "Arhythmia of the Heart" written by the Dutch physician Wenckebach was published in 1904. In his book, Wenckebach acknowledged the Scottish physician Mackenzie's study of the pulse which was published in 1902. All of Mackenzie's recordings were made with a polygraph, which allowed two channels of pressure tracing to be recorded. He recorded the arterial and jugular pulse simultaneously. In 1908, Mackenzie's book entitled "Diseases of the Heart" was published.

In 1914, Frank Wilson, at the University of Michigan, obtained a string galvanometer and became deeply involved with electrocardiography. During the 1920s, Wilson and his team undertook many studies

correlating electrocardiographic findings (essentially limb leads I, II, and III) with abnormalities such as ventricular hypertrophy and bundle branch block. Nowadays, their concept of "Ventricular Gradient" is still used to explain certain phenomena. For example, Abildskov (1987) suggested that inequality of the "ventricular gradient" in different areas of the myocardium might be responsible for the generation of ventricular arrhythmias.

Nevertheless, Wilson's major contribution is acknowledged to be his "central terminal" (Wilson, Johnston, MacLeod & Barker 1934) with which unipolar chest leads can be recorded. In summary, however, the concept allows the potential variation at a single point on the chest to be recorded with respect to a relatively constant reference potential obtained by averaging the potentials of right and left arms and the left leg. This configuration records what is known as a "unipolar" lead. If the potential at the Wilson central terminal is denoted by E_{WCT} then

$$E_{WCT} = 1/3\,(E_R + E_L + E_F)$$

The next stage in the evolution of the unipolar lead was for Wilson's team to specify six precordial positions for the exploring electrodes, which were slightly different from the six praecordial leads V1-V6 used nowadays (Kossmann & Johnston, 1935). However, in a series of publications from 1938 to 1943 the Cardiac Society of Great Britain and the American Heart Association agreed on six praecordial electrode positions.

In 1942, Goldberger introduced the "augmented unipolar limb lead" to electrocardiography. He removed the Wilson central terminal connection

from the limb on which the exploring lead was placed and this augmented the potential recorded by 50% (Goldberger, 1942). Mathematically, the relationship between the modified unipolar right arm lead (modified VR), which is modified by the removal of the right arm connection to the Wilson central terminal, and the resultant potential of the modified Goldberger's Terminal (E_{GT}) for this lead is as follows:

$$
\begin{aligned}
\text{VR} &= E_R - E_{WCT} \\
\text{modified VR} &= E_R - E_{GT} \\
&= E_R - 1/2\,(E_L + E_F) \\
&= 3/2\,E_R - 1/2\,(E_L + E_F + E_R) \\
&= 3/2\,[\,E_R - 1/3\,(E_L + E_F + E_R)] \\
&= 3/2\,[\,E_R - E_{WCT}] \\
&= 3/2\,\text{VR}
\end{aligned}
$$

The modified VR became known as aVR. Similar circuitry was introduced to record modified VL and VF, i.e. aVL and aVF.

Therefore, the development of conventional 12-lead electrocardiography was completed. There were three limb leads I, II, and III from Einthoven; three augmented unipolar limb leads aVR, aVL, and aVF from Goldberger's modification of Wilson's central terminal; and six praecordial leads V_1-V_6 arising out of Wilson's central terminal.

The development of electrocardiography continued in many different directions and progressed through the research of several gifted electrocardiographers and electrophysiologists. Improvements in the techniques of measurement, recording, interpretation and modelling as well as the elaboration of various theories have all contributed to the

widening and refining of electrocardiology as its stands today.

1.3 HISTORY OF VECTORCARDIOGRAPHY

A vector is an entity with magnitude and direction. The major difference between the conventional 12-lead electrocardiogram and the vectorcardiogram (VCG) is the method of display. In the vectorcardiogram the cardiac electric potential is represented by a single dipole, the strength and spatial orientation of which at each moment are depicted by a spatial vector. The changing direction and magnitude of the instantaneous vectors during each cardiac cycle are displayed as loops formed by joining the tips of the vectors. The spatial vector loops are viewed in three mutually perpendicular planes: horizontal, sagittal and frontal. The phasic and directional changes of the electric forces generated by the heart are more clearly displayed in the vectorcardiogram, which is a still picture of one cardiac cycle.

The concept of a vector force was invoked initially by Waller (1887) at the beginning of human electrocardiography. He produced an isopotential map which suggested that the electromotive force of the heart could be represented by a single dipole (Waller, 1887). Later in 1913, Einthoven and colleagues introduced the concept of measuring the mean electrical axis of the heart which was represented by a vector (Einthoven et al, 1913) [Fig 1-3].

By building on Waller's and Einthoven's vectorial presentation of excitation, the experimental and mathematical work was accomplished for characterising the spatial spread of activation on the basis of the dipole model. In 1916, Lewis published his work about the concept of a

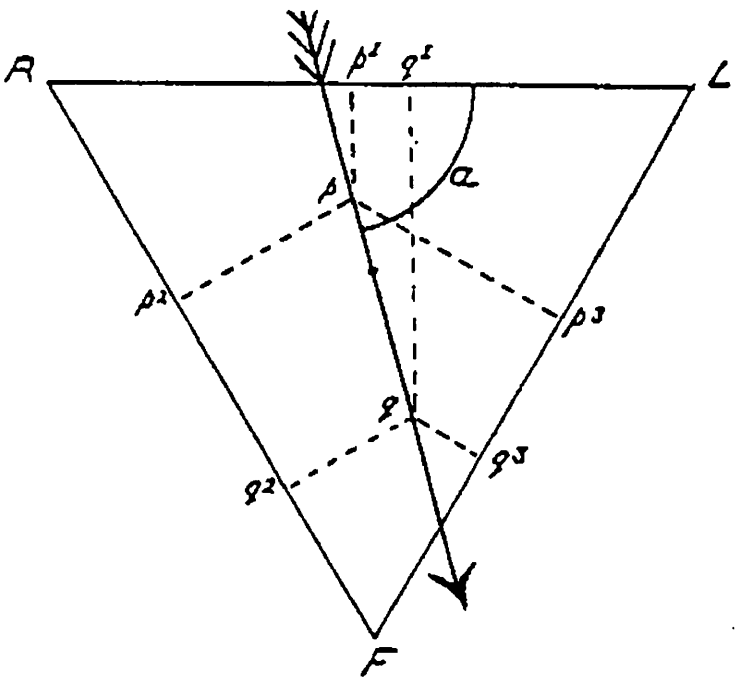


Fig. 1-3 Vector concept of measuring the mean electrical axis of the heart according to Einthoven, Fahr & de Waart (1913).

vectorcardiogram (Lewis, 1916). Preliminary theoretical contributions were on the two component model of Schutz and the "Momentanvektor" by Koch in 1936 (Snellen, 1984). In 1960, Geselowitz expanded this to a "multiple-dipole" theory (Geselowitz, 1960).

According to Burch (1985) and Snellen (1984), the first publication describing a method for manually deriving a "Vectorcardiogram" from standard limb leads was written by Williams in 1914 (Williams, 1914). In this case, however, an amplitude was associated with each vector direction, and if the tips of the vectors had been joined in the correct sequence an approximate figure-of-eight configuration would have been seen. In fact, this was later done by Mann, and the loop [Fig 1-4] thus obtained was called a "Monocardiogram" (Mann, 1920). Mann subsequently also invented a monocardiograph in 1925 which used a cathode ray oscilloscope to display the vector loop (Mann, 1938).

The advent of the cathode ray oscilloscope radically changed the approach to displaying vector loops. The advantage of the cathode ray oscilloscope was that two separate leads could be applied to opposite pairs of plates in order to deflect the electron beam in proportion to the strength of the signal on each axis (Mann, 1938). According to Burch (1985), Schellong of Germany in 1936, Wilson's team in the United States in 1937, and Hollmann W. and Hollmann H.E. of Germany in 1937, independently developed systems for displaying vector loops around the same period. Schellong was the first among them to publish the vector loops recorded with the cathode ray oscilloscope (Schellong, 1936). In 1936, Rijlant also used a cathode ray oscilloscope to display the scalar electrocardiogram (Rijlant, 1936). Although early efforts at

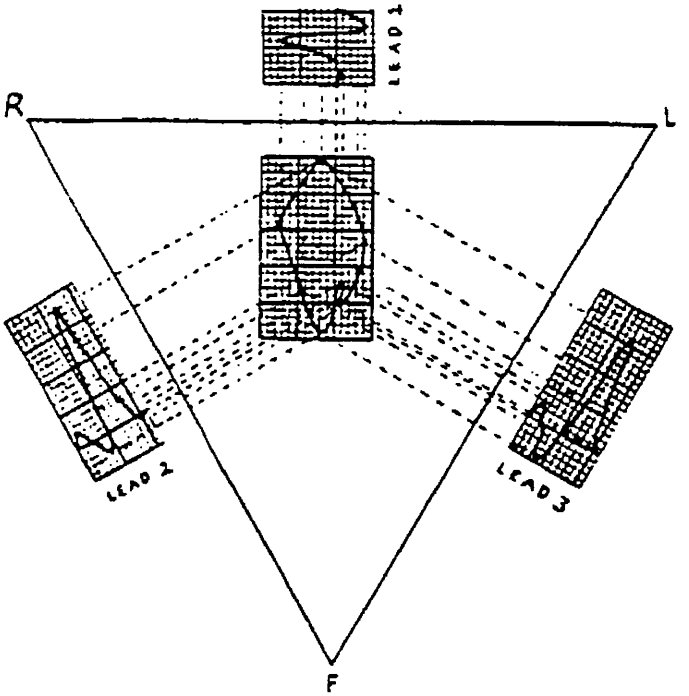


Fig.1-4 Monocardiogram recorded by Mann (1920).

constructing vector loops were in the frontal plane, the use of the oscilloscope also allowed other planes to be viewed with relative ease.

The concept of a three dimensional vector loop in space gradually grew, and a number of lead systems were introduced to derive the components of the resultant cardiac vectors, e.g. McFee and Parungao's axial lead system (McFee & Parungao, 1961) as well as Frank's corrected orthogonal lead system (Frank, 1956).

A perfect lead system for vectorcardiography would consist of three leads with the following characteristics (Chou & Helm, 1967):

[1] The three leads would be mutually perpendicular and each would be parallel to one of the rectilinear coordinate axes of the body. These axes are the horizontal or X axis (right-left axis), the longitudinal or Y axis (superior-inferior or head-foot axis), and the sagittal or Z axis (posteroanterior axis).
[2] The three leads would have equal strength.
[3] The lead vectors of the three leads would not only be of equal amplitude but would also retain the same magnitude and the same direction for all points where cardiac electromotive forces are generated.

Vectorcardiographic leads which purportedly meet condition [1] are referred to as "orthogonal leads". If, in addition, such leads also reasonably meet conditions [2] and [3], they are termed "corrected orthogonal leads", such as those of McFee and Parungao's axial system, and of the Frank system.

Frank (1954) made use of a tank model of a human torso in which an artificial generator was placed in order to study the effects of different

leads on the surface of the body. He created an image surface which effectively delineated an imaginary torso where lines joining two points bore a true theoretical relationship to the strength of current flow between the two corresponding points on the actual torso. In 1956, Frank (1956) described a seven electrode orthogonal lead system; five electrodes are placed on the trunk and the remaining two are located on the back of the neck and on the left leg. Four of the five thoracic leads are placed at the level of the intersection of the fifth intercostal space with the parasternal lines, i.e. where the plane formed by this transverse level intersects the left and right mid-axillary lines, the sternal line, and the vertebral line. The remaining one thoracic lead is placed at the same level on the left side of the chest at an angle of 45° with respect to the centre of the thorax, which is determined by the lines joining the other four points. The Frank lead system is a corrected orthogonal lead system.

Schmitt and Simonson also made many contributions to theoretical and practical studies in electrocardiography and introduced the SVEC III (Stereo vector electrocardiography) system (Simonson, Nakagawa & Schmitt, 1957). Wilson and Johnston introduced the term "vectorcardiogram" to describe the vector loops projected on the three mutually perpendicular planes (Wilson & Johnston, 1938).

1.4 DERIVED VECTORCARDIOGRAPHY

1.4.1 Introduction

Vectorcardiography has been recognised as a useful clinical investigational method for the study of the spatial and temporal

relationships of cardiac potentials (Chou, 1986). Vectorcardiography has been claimed to be superior to electrocardiography for the diagnosis of certain cardiovascular abnormalities, e.g. right ventricular hypertrophy (Chou, 1986), conduction disorders, and inferior myocardial infarction (Brohet, 1991; Starr et al, 1974; Hurd et al, 1981). The vectorcardiogram has also been claimed to have special value in paediatric cardiology, especially for the quantitative assessment of ventricular hypertrophies and accurate diagnosis of ventricular conduction disturbances (Brohet, 1991).

The clinical value of the vectorcardiogram is related to the main characteristics of the method which displays the cardiac potentials in the form of planar loops [Fig 1-5] in which all instantaneous vectors can be clearly defined and identified in terms of magnitude and orientation. In addition, the spatial and temporal relationships of the cardiac potentials can also be clearly demonstrated. In comparison with the conventional 12-lead electrocardiogram, the vectorcardiogram has been claimed to have the advantage of more precise delineation of the whole depolarisation and repolarisation processes and more prominent phase changes (Chou, 1986). Although the diagnostic superiority of the vectorcardiogram has been claimed for decades, the conventional 12-lead electrocardiogram is still the most widely available non-invasive investigation performed in daily cardiological practice.

The reasons why the vectorcardiogram has not succeeded in replacing the 12-lead electrocardiogram are as follows :

[i] The conventional vectorcardiogram recording techniques are cumbersome and time-consuming.

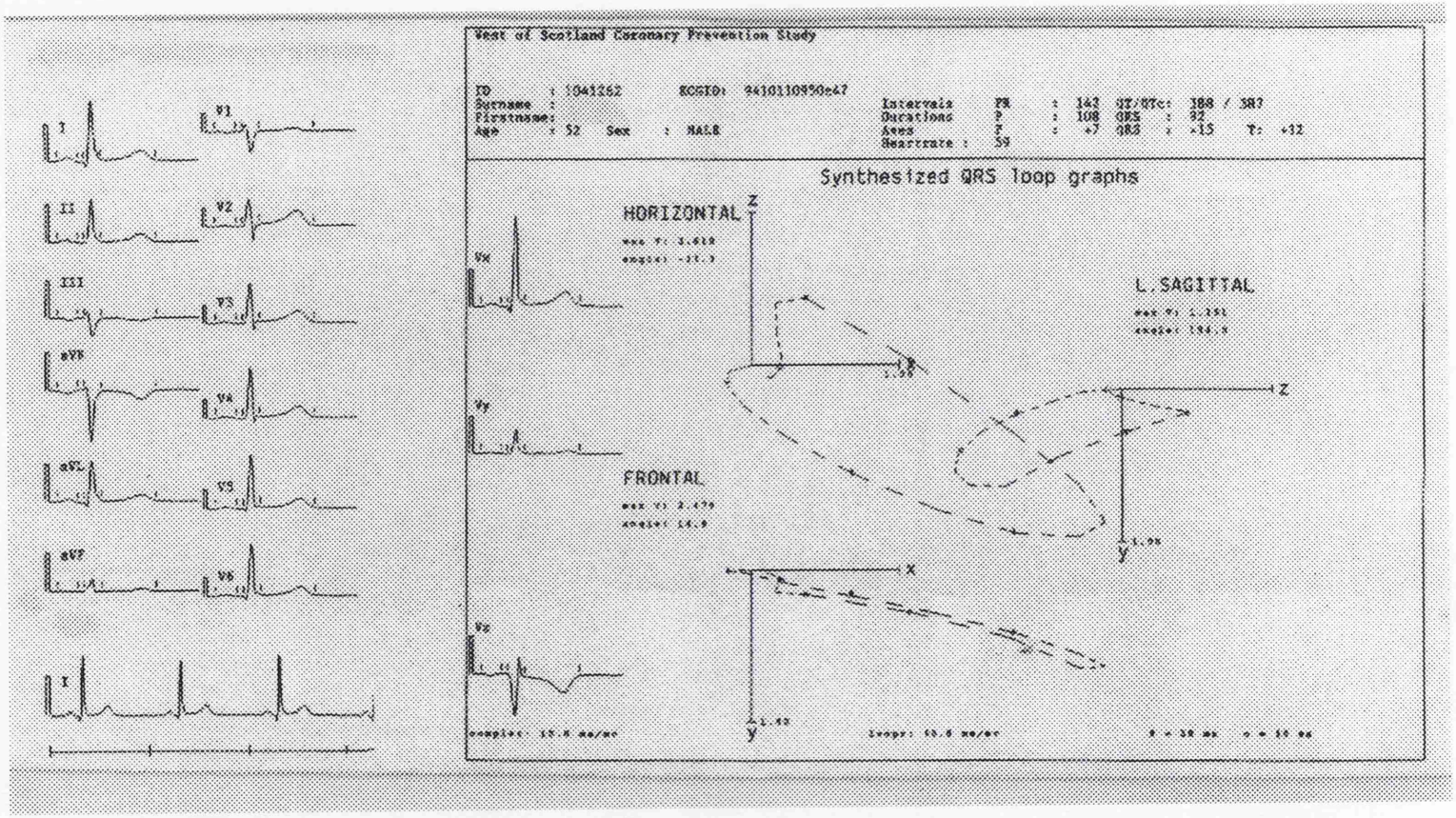


Fig. 1-5 QRS loops from a derived (synthesized) VCG and the corresponding 12-lead ECG.

[ii] The recording of a vectorcardiogram using a corrected orthogonal lead system requires a completely different set of electrodes compared to those used for recording the 12-lead electrocardiogram.

[iii] The interpretation of a vectorcardiogram is a skill that many physicians have not acquired.

Several attempts were made to simplify the procedure of recording, notably the use of the hybrid lead system by Macfarlane (1979), whereby two additional electrodes V_{5R} and V_{neck} were used to record both the 12 and 3-orthogonal lead electrocardiogram simultaneously. This system has been used routinely for over 15 years in Glasgow Royal Infirmary for ECG interpretation. However, this approach and others were never adopted commercially because of the need for extra hardware. Furthermore, cardiologists were generally not convinced that additional benefits could be obtained from the vectorcardiogram. Nevertheless, its proponents have claimed that the vectorcardiogram might still provide complementary, if not supplementary, information to that of the conventional 12-lead electrocardiogram (Brohet, 1991; Edenbrandt & Pahlm, 1988b; Edenbrandt et al, 1990; Kors et al, 1992).

Much effort has been spent since the early 1960s on research concerning computer-assisted diagnostic classification of the conventional 12-lead electrocardiograms (Kors, Talmon & van Bemmel, 1986; Macfarlane, 1990a; Pipberger et al, 1960) and paediatric orthogonal electrocardiograms (Brohet et al, 1984). Research on reconstruction of the vectorcardiographic leads from conventional 12 electrocardiographic leads by computing a linear combination of the electrocardiographic leads, in which the coefficients are arranged in a reconstruction matrix, has also been investigated (Bjerle & Arvedson,

1986; Edenbrandt & Pahlm, 1988a; Kors et al, 1990; Wolf et al, 1976). The approach is relatively new so that research in this field is still active and ranges from standardisation to methodological aspects of the diagnostic classification. This vectorcardiogram reconstruction approach has the advantage that due to utilisation of modern computer techniques, there is no need to use additional leads other than the conventional scalar 12 leads to derive the three orthogonal leads and to produce the vectorcardiographic loops simultaneously.

1.4.2 Methods of Reconstructing the Vectorcardiogram

The synthesis of a particular vectorcardiographic lead can be achieved by a linear combination of the various electrocardiographic leads (I, II, & all praecordial leads) with the coefficients (a_i, b_i, - - - h_i; i= 1, 2, 3) arranged in a reconstruction matrix, i.e.

$$
\begin{aligned}
X= \quad & a_1 I + b_1 II + c_1 V_1 + d_1 V_2 + e_1 V_3 + \\
& f_1 V_4 + g_1 V_5 + h_1 V_6. \\
Y= \quad & a_2 I + b_2 II + c_2 V_1 + d_2 V_2 + e_2 V_3 + \\
& f_2 V_4 + g_2 V_5 + h_2 V_6. \\
Z= \quad & a_3 I + b_3 II + c_3 V_1 + d_3 V_2 + e_3 V_3 + \\
& f_3 V_4 + g_3 V_5 + h_3 V_6.
\end{aligned}
$$

There are basically three approaches for determination of the reconstruction coefficients.

(A) A single-lead approach

A vectorcardiographic lead is represented by only a single electrocardiographic lead that resembles it best when suitably scaled. This

is the simplest form of reconstruction. In this approach, all the input leads have their coefficients set to zero, except one in each of the three "Quasi-Orthogonal" electrocardiographic leads.

For example, an uncorrected orthogonal system might consist of

$$X = h_1 V_6$$
$$Y = b_2 II$$
$$Z = d_3 V_2$$

This method has been described by Kors et al (1986) in the signal-processing part of their ECG interpretation program. From a set of simultaneously recorded electrocardiograms and vectorcardiograms from the same individual, the correlation coefficients between the electrocardiographic and vectorcardiographic leads were computed. The electrocardiographic lead which showed the highest median correlation was chosen for each vectorcardiographic lead. The amplitude adjustment of these electrocardiographic leads was selected as simply as possible without compromising the signal-analysis performance. As a result, the leads X, Y, and Z were approximated by leads V_6, II, and -0.5 V_2 respectively, i.e.

$$X = V_6$$
$$Y = II$$
$$Z = -0.5\ V_2$$

Bjerle and Arvedson (1986) reported another set of coefficients for leads X, Y, and Z, namely,

$$X = 1.06\ V_6$$
$$Y = 1.88\ aVF$$
$$Z = -0.532\ V_2 + 0.043\ V_6$$

(B) A model-based approach

This method was first adopted by Dower et al for the derivation of the 12-lead electrocardiogram from the Frank vectorcardiogram (Dower, 1968; Dower, Machado & Osborne, 1980). Each electrocardiographic lead was represented as a linear combination of the X, Y, and Z components of the vectorcardiogram with the coefficients derived from Frank's torso model (Frank, 1956). In recent years, Edenbrandt and Pahlm (1988 a & b) at the University of Lund, Sweden have computed an "inverse" matrix from Dower's matrix (Dower, 1968; Dower, Machado & Osborne, 1980) for the derivation of the vectorcardiogram from the 12-lead electrocardiogram.

The equations are as follows:

$$
\begin{aligned}
X= &-0.172V1 -0.074V2 +0.122V3 +0.231V4 \\
&+0.239V5+0.194V6 +0.156I -0.010II \\
Y= &\ 0.057V1 -0.019V2 -0.106V3 -0.022V4 \\
&+0.041V5 +0.048V6 -0.227I +0.887II \\
Z= &-0.229V1 -0.310V2 -0.246V3 -0.063V4 \\
&+0.055V5 +0.108V6 +0.022I +0.102II
\end{aligned}
$$

In this case, lead Z is directed positively to the back of the thorax.

(C) A statistical approach

In this approach, the reconstruction matrix is derived by a statistical regression technique. Burger et al (1962) first described the use of this technique in order to transform vectorcardiograms from one lead system to another (Burger, van Brummelen & van Herpen, 1962). The reconstruction coefficients, based on a learning set of simultaneously

recorded electrocardiograms and vectorcardiograms, are calculated by minimizing the sum of the squared differences between the target lead and its reconstruction lead. Wolf et al (1976) derived vectorcardiograms from electrocardiograms by this approach but the reconstruction matrix itself was not published.

A multivariate regression obtained by using the BMDP statistical package was recently described by Kors et al (1990b) for deriving the various reconstruction matrices for the different segments of the P-QRS-T complex. A visual comparison of the reconstruction results for a subset of the learning population, using different reconstruction matrices for P, QRS and T revealed only very minor differences. Therefore, the reconstruction matrix from the regression on the QRS complex was adopted and subsequently applied to the whole PQRST complex.

The equations of Kors et al (1990b) are as follows:

$$
\begin{aligned}
X= &-0.13V1 +0.05V2 -0.01V3 +0.14V4 \\
&+0.06V5 +0.54V6 +0.38I -0.07II \\
Y= &+0.06V1 -0.02V2 -0.05V3 +0.06V4 \\
&-0.17V5 +0.13V6 -0.07I +0.93II \\
Z= &-0.43V1 -0.06V2 -0.14V3 -0.20V4 \\
&-0.11V5 +0.31V6 +0.11I -0.23II
\end{aligned}
$$

1.4.3 The Advantages of Derived Vectorcardiography

Derived vectorcardiography has the following advantages compared to other vectorcardiographic systems:

1. The use of the same set of standard electrodes as those for recording the 12-lead electrocardiogram.

2. The availability of complementary diagnostic information compared to the 12-lead electrocardiogram.
3. The simplicity of calculation of XYZ leads and display of vector loops by modern computer techniques.
4. The possibility of incorporating vectorcardiographic criteria into a conventional 12-lead electrocardiographic analysis program.

Recently, the combination of the computerised electrocardiogram and vectorcardiogram interpretation programs has been claimed to perform significantly better than each program separately (Kors et al, 1992). Thus, it was thought that the performance of the automated Glasgow electrocardiographic analysis program might be further improved by combining both electrocardiographic and derived vectorcardiographic criteria.

1.5 HISTORY OF COMPUTERISED ECG

In 1987, Drazen et al (1988) suggested that there were 50 million electrocardiograms reported annually by computer in the United States. In addition, there were thought to be around 3 million electrocardiograms interpreted annually by computer in Japan and perhaps 8 million per year in Europe (Macfarlane, 1990a).

In the 1960s, there were only two research groups in the U.S.A, both in Washington, D.C., investigating the use of a large digital computer for the analysis of electrocardiograms. Pipberger's group favored the use of simultaneously recorded three orthogonal XYZ leads (Pipberger et al, 1960), while in Caceres' group, the conventional 12-leads were recorded separately on analogue tape using a relatively high-fidelity method and

then processed sequentially (Caceres, Steinberg & Abraham, 1962). The first development was the construction of a device which allowed the electrical signals generated by the electrocardiograph to be converted into digital data which can be handled by a computer. This was accomplished through the construction of an analogue-to-digital converter (Pipberger et al, 1960). Methods for digital signal filtering were also introduced in those early days aimed principally at the removal of 60Hz mains interference.

The advantage of simultaneously recorded three orthogonal XYZ leads was that identical cardiac cycles for all leads were being processed at the same time. This is not so for the sequentially recorded 12-leads. Analysis of only one lead at a time was indeed particularly difficult and caused problems in leads where the projection of the cardiac potentials resulted in a low amplitude QRS complex.

In order to analyse the 12-lead ECG accurately, Macfarlane (Macfarlane, 1971b; Macfarlane & Lawrie, 1974) suggested that the leads should be recorded in groups of three simultaneously. These were "I, aVF, V1"; "aVL, II, V4"; "III, V3, V6"; "V2, aVR, V5". Certainly by this method, the cardiac cycles measured in one group of three leads were still different from those in another. However, this represented a significant step in improving the accuracy of the wave recognition process. This method was subsequently adopted by Bonner's group in IBM (Bonner et al, 1972), although leads were recorded in groups I, II, III; aVR, aVL, aVF; V1, V2, V3; V4, V5, V6.

There were other major differences in approach at the beginning of research into computerised ECG interpretation and these still persist,

e.g. the technique of interpretation. Pipberger's group used statistical techniques involving "prior probabilities". This approach was different from the "deterministic" approach which others used. The "deterministic" approach essentially used a diagnostic tree, with rule-based diagnostic criteria. Most multivariate statistical techniques require certain advance knowledge of the likelihood of a particular interpretation. These "prior probabilities" were dependent upon the population screened. Today, there are still controversies about whether the computer is interpreting the actual electrocardiogram or is biased by prior probabilities to such an extent that the actual waveform is of secondary importance to the clinical classification (Macfarlane, 1990a). Willems et al (1986) have applied this statistical approach to the electrocardiographic interpretation using the known clinical condition and age of the patient. Combinations of the "deterministic logic" and "prior probability" techniques have been investigated where at certain points in the logic of a rule-based program, a decision is made on the basis of statistical probabilities (Zywietz et al, 1977). This area of research is still in progress.

The major technological advance of the 1970s was the advent of the microprocessor on a single chip. This development led to revolutionary changes in the approach to ECG analysis and the arrival of automated ECG analysis at the bedside. In 1977, Macfarlane's group adopted a completely new system with simultaneous acquisition of all eight independent leads of the 12-lead ECG and three orthogonal XYZ leads using a microprocessor-controlled electrocardiograph (Watts & Shoat, 1987) with digital transmission to a central departmental computer for analysis in Glasgow Royal Infirmary (Macfarlane et al, 1990 b).

In the 1980s, the most dramatic advances were in the field of miniaturization of circuitry allowing a complete 12-lead electrocardiographic analysis to be provided within a relatively small electrocardiograph producing either a single-channel or a multi-channel output together with the interpretation. The advent of the thermal writer also allowed flexibility for the user to print whatever is required. Digital transmission of electrocardiograms has become the method of choice, particularly where high-speed networks are available and many telecommunication systems worldwide are now using digital exchanges. The development of ECG management systems has also made serial comparison of electrocardiograms possible. The more recent availability of optical discs has led to advances in the storage of large numbers of electrocardiograms on a small disc.

Interpretation of complex cardiac arrhythmias by a computer system remains a problematic area. There are two types of rhythm analysis techniques. In one type, all leads (whether they are XYZ or 12 leads) are recorded simultaneously for up to ten seconds and the program has the the possibility of using a number of these leads for analysis in reaching the diagnosis. In the other type, the 12-lead electrocardiogram may have been recorded in four groups of three leads with varying duration fom 2.5 seconds to 5 seconds. This complicates the analysis when trying to determine the presence of ectopic beats, particularly if they occur at a time when the leads are switching from one group to another. Through the years, there is no doubt that rhythm analysis by computer has been improved, but there is probably a plateau level above which it is unlikely that computer techniques can proceed.

1.6 The Common Standards for Quantitative Electrocardiography (CSE) Project

ECG computer processing can fundamentally be categorised into three principal stages (Willems et al, 1990):

(1) Acquisition, transmission and storage of digitised ECG data.
(2) Pattern recognition and measurement.
(3) Diagnostic classification.

In each of these stages, evaluation is mandatory. Therefore, a project entitled "Common Standards for Quantitative Electrocardiography" was established and subsequently directed by Willems since 1978 (Macfarlane, 1990a).

The major objectives of the CSE project were formulated as follows (Willems et al, 1990):

(1) Standardisation of ECG measurement procedures in quantitative terms; comparative studies of measurements performed by different programs; drawing of guidelines, definitions and standards for measurement.
(2) Assessment of the performance of diagnostic classification of computer programs and algorithmic documentation of their operation.
(3) Establishment of modest ECG libraries to reach these goals.

The major landmarks have been the set of recommendations for defining ECG waves processed by computer techniques, the creation of two databases of electrocardiograms, one for assessing the accuracy of measurement (Willems et al, 1990) and the other for assessing the

diagnostic accuracy of computer programs (Willems et al, 1990), and for the comparative assessment of cardiologists and programs (Willems et al, 1991).

More recently, two separate projects have been established, one to consider the testing of electrocardiographs which is supported through the program for Conformance Testing Services (CTS) of the EEC (Zywietz & Willems, 1993) and the other to establish a Standardised Communications Protocol (SCP) for ECG data transmission and storage which was supported through the Advanced Informatics in Medicine (AIM) program of the EEC (Zywietz & Willems, 1993). A European pre-normative standard (prENV 1064) has now been established for SCP.

1.7 CURRENT USE OF COMPUTERISED ECG

Nowadays, cardiologists are submerged in the advancing waves of technology (i.e. echocardiography, nuclear scintigraphy, and magnetic resonance imaging, etc). Electrocardiography has suffered understandable neglect. Nevertheless, the startling fact remains that the single most often used, most cost-effective and most appropriate diagnostic investigation in cardiology today is the electrocardiogram. Ironically, it is also, despite its relative simplicity, the most frequently misinterpreted, often with calamitous results (Macfarlane & Lawrie, 1974). The utilisation of computers for ECG analysis can not only save time but perhaps can improve the accuracy of interpretation. It is also acknowledged that computer-assisted ECG interpretations are not always as accurate as those of skilled cardiologists, who have many years of experience on which to base their ECG reporting. On the other hand, cardiologists often disagree over an ECG interpretation. Furthermore,

their interpretations of the same ECG on separate occasions frequently differ, with a median repeatability of only 81.2% according to the CSE studies (Willems et al, 1991). This underlines the problem of choosing a yardstick with which to judge the computer.

There is still room for improvement in computerised electrocardiography. Because artificial neural networks are not rule-based and can be applied in areas where rules are difficult to formalise or are ill defined, they are theoretically particularly suitable for application to visual pattern recognition problems. Therefore it was thought worthwhile to investigate whether the utilisation of artificial neural networks alone or in combination with deterministic logic for the analysis of the ECG would bring enhanced benefit.

CHAPTER 2

ARTIFICIAL NEURAL NETWORKS

2.1 HISTORY OF DEVELOPMENT OF ARTIFICIAL INTELLIGENCE

By the early 1950's, there were several major developments in formal logic and computational theory. Theorists had appreciated the enormous power of abstract systems of symbols that can undergo rule-governed transformations. It was thought that if these abstract systems could be automated, then their abstract computational power would seem to be displayed in a real physical system. This insight led to two theoretical and computational developments (Aleksander & Burnett, 1987).

The first development according to Churchland & Churchland (1990) was "Church's thesis", which states that every effectively computable function is recursively computable. "Effectively computable" means that there is a "rote" procedure for determining the output of the function for a given input in finite time. "Recursively computable" means more specifically that there is a finite set of operations that can be applied to a given input, and subsequently applied over and over again to the successive results of such applications, to yield the output of this function in finite time. The notion of a rote procedure is non-formal and intuitive (Churchland & Churchland, 1990).

The second important development was mathematician Alan M. Turing's demonstration that any recursively computable function can be computed

in finite time by a maximally simple type of symbol-manipulating machine that has come to be called a "universal Turing machine" (Turing, 1950).

A "universal machine" is basically a machine which can replicate the behaviour of all machines, both machines that actually exist and machines that do not exist, but whose behaviour can be specified formally and in detail. The essence of the universality of the computer, therefore, does not lie in its real ability to perform functions in the physical world, but in its capacity to reproduce the logical characteristics of other systems. Therefore, as far as artificial intelligence is concerned, the crucial point is that a computer can simulate the operations of any other mechanisms that process information.

Turing's classic paper "Computing Machinery and Intelligence (1950)" in which the Turing test was outlined, appeared barely two years after the world's first computer ran the world's first stored program in a laboratory at Manchester University U.K. on 21th June 1948. Given the right program, a large enough memory and sufficient time, these two developments (Church's Thesis and Turing's Demonstration) make the standard digital computer able to compute any rule-governed input-output function. These results imply that a suitably programmed symbol-manipulating machine should be able to pass the Turing test for conscious intelligence. This also formed the fundamental research direction of "Classical" or "Program-writing" Artificial Intelligence (Aleksander & Burnett, 1987).

Classical Artificial Intelligence has two characteristics as shown below:

(1) The physical material of any symbol-manipulating machine is not related to the function it computes. That is fixed by the program of the machine.
(2) The engineering details of any machine's functional architecture are also irrelevant, since different architectures running quite different programs can still be computing the same input-output function. It was said (Aleksander & Burnett, 1987) that the idiosyncratic way in which artificial intelligence computes has nothing to do with its function.

This completes the rationale and theoretical background for "Classical Artificial Intelligence".

2.2 INTRODUCTION TO ARTIFICIAL NEURAL NETWORKS

For centuries, mankind has sought to understand and duplicate the processes that constitute intelligence. The advent of modern computer technology, with its ability to perform many complex tasks far better and faster than the human brain, has led many to predict that this quest would soon be completed.

Modern computers are based entirely on one or more central processor units that are capable of only simple arithmetic and logic operations on binary numbers. These operations can be combined to perform new complex tasks. All such work handled by computers including calculations, word processing and robot control, is done by translating tasks into binary operations. Information is stored in groups of storage elements termed "registers", which are separate from the processing unit.

During the last two decades, the computing industries have made astounding advances by making the binary processing units (transistors) smaller and faster, by devising better software environments to facilitate the translation of complex tasks into binary operations, and by developing new applications for the technology.

These advances in technology, although impressive, have failed to produce devices that can approach the capability of human intelligence in pattern recognition, innovation, and creativity. There is considerable debate about whether these capabilities will ever be duplicated by the conventional binary computers (Searle, 1990; Churchland & Churchland, 1990). Other researchers have started to explore alternative architectures that hold more promise of success.

The artificial neural network approach, also referred to as connectionism or parallel distributed processing, adopts a "brain metamorphor" of information processing (Branscombe, 1990). The artificial neural network is one such architecture that simulates biological nervous systems. Its architecture, function, and use differ fundamentally from those of conventional computers. The fundamental processing units, also called neurons, of an artificial neural network have similar properties to those of biological neurons. Neural networks are more like computing memories where the operations are association of similarity. The neural network sums positive [excitatory] and negative [inhibitory] inputs to produce a single output, which in turn synapses with one or more similar processing units in sequence. Hence, in contrast to conventional computer systems, artificial neural networks have multiple processing units functioning in tandem that serve simultaneously as both memory and processing units. In this approach, information processing occurs

through interactions involving large numbers of simulated neurons as shown in Figure 2-1.

Because only numerically valued activation passes from neuron to neuron in a neural network, neural networks are often classified as a sub-symbolic level of computation. The input, output, and internal state of a neural network can all be characterised by patterns of activation across its nodes.

These artificial network architectures were explored in the early years of computer development but were abandoned during the 1960's. Although cybemeticians continued to analyse the behaviour of neural networks on paper, and were soon able to simulate it on computers, the attack had more or less gone by the end of 1960's. The obstacles encountered were both fundamental and irremediable. On the practical front, the major problem was undoubtedly the fact that it proved impossible to devise an artificial equivalent of the neuron which was both reliable and easy to manufacture in quantity.

In recent years, this situation changed following the introduction of a "back-propagation" algorithm (Rumelhart, Hinton & Williams, 1986) for the training of such systems, which has led to enthusiastically renewed research in this field. Back-propagation is a generalised delta (Widrow & Hoff, 1960) learning rule developed simultaneously by Rumelhart et al (1986), Parker (1985), and Le Cun (1985). The back-propagation mechanism is a powerful, general learning algorithm employing a gradient or steepest descent heuristic that enables an artificial neural network to self-organise in ways that improve its performance over time (Jones & Hoskins, 1987). Among connections pointing from a given

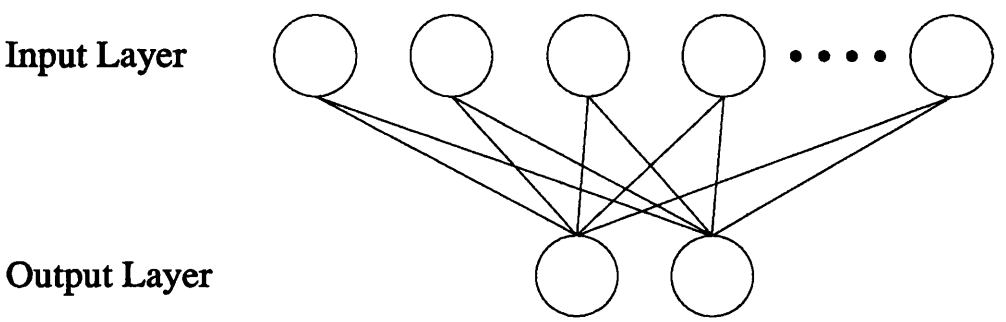


FIG 2-1. Basic Structure of a two-layer Neural Network

input neuron, the larger modifications will involve those connections that point to an output neuron with larger associated deltas (i.e. discrepancies between actual and target activation levels). Back-Propagation is a learning mechanism where an error signal is fed back through the network altering weights as it goes, to prevent the same error from happening again.

2.3 DESCRIPTION OF A NEURAL NETWORK

An artificial neural network is best described by layers of neurons (nodes), which share a functional feature, interconnected by synapses [links], which have certain connection weights or strengths that can be excitatory [positive weights] or inhibitory [negative weights]. The basic component of an artificial neural network is the neuron. The function of a neuron is to compute a numerical value for onward transmission to the next layer. A two-layer neural network with only input and output layers is shown in Fig 2-1.

The state S_i [activation level] of the ith neuron is given by the inner product of the input X_j applied to it [where Xj is the output of the jth neuron] and the strength [weight] of the link W_{ij} between neuron i and neuron j.

This can be expressed by the following equation:

$$S_i = \sum_{j=1}^{n} W_{ij} X_j$$

where n represents the number of input connections to neuron i.

Such a state S_i is processed by a transfer [activation] function F, which produces the output signal Y_i of the neuron as follows [Fig 2-2]:

$$Y_i = F(S_i)$$

$$Y_i = F\left[\sum_{j=1}^{n} W_{ij} X_j\right]$$

The transfer [activation] function F can be linear or non-linear, and is used to determine the outputs of a network as a function of its input variables.

The most commonly implemented non-linear transfer function is the sigmoid function, i.e.

$$F(S_i) = \frac{1}{\left[1+e^{-S_i}\right]}$$

where S_i is the sum of the weighted inputs to the neuron, and e is the exponential function. The other non-linear transfer functions include hard-limiter or step functions and threshold logic elements [Fig 2-3]. Note that the output $F(S_i)$ satisfies: $0 < F(S_i) < 1$.

More complex nodes may include temporal integration or other types of time dependencies and more complex mathematical operations than summation. Neural network models are specified by the topology of the network, characteristics of the nodes, and the training or learning rules. These rules specify an initial set of weights and indicate how weights

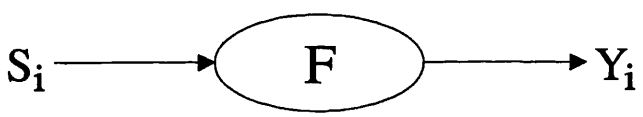


FIG 2-2. Transfer Function

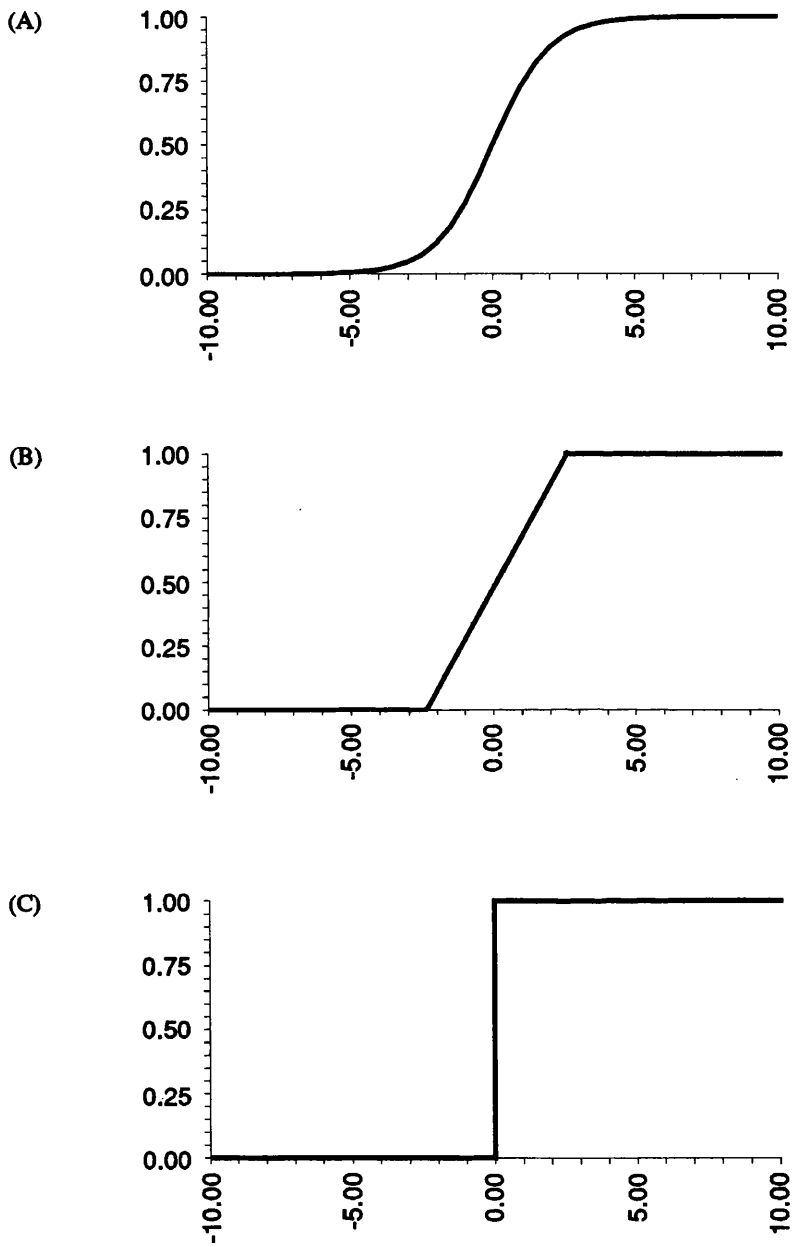


FIG 2-3. Transfer Functions

(A) Sigmoid function

(B) Linear threshold function

(C) Hard limiter or step function

should be adapted during training to improve performance (Lippmann, 1987).

2.4 The McCulloch Pitts Neuron

In 1943, W.S. McCulloch & W. Pitts (1943) of the University of Illinois described their original research on artificial neural networks in which they proposed a model of the neuron's logic function from which it was already apparent that the processing circuitry at the heart of a binary calculator could be seen in logical as well as arithmetical terms.

According to Aleksander & Burnett (1987), this had in fact been pointed out in 1938 by Claude Shannon in a paper entitled " A Symbolic Analysis of Relay and Switching Circuits". The essential point is that the state of a switch can be seen as representing a "yes" or "no", a "true" or "false" just as readily as it can be thought of as representing a "1" or a "0". Thus the logic gates which are used to add, multiply, or apply other mathematical operations to binary digits, can also be thought of as implementing logical functions such as "and", "or", "not", etc.

Even more importantly, Shannon showed that a circuit which is arranged so that the output of one logic operation is fed back to provide one of the inputs to the next operation would be capable of implementing the "if" function.

The formal model of the neuron had been described as a threshold logic unit, which later became known as the "McCulloch-Pitts neuron" [or M-P neuron] (1943). The M-P neuron is characterised by a finite number of excitatory or inhibitory inputs, a threshold level and an output. The

inputs to and outputs from the M-P neuron can assume binary values 0 and 1.

The output S_i of the M-P neuron can be described as a function F of its inputs, that can be expressed as follows:

$$S_i = F\left[\sum_{j=1}^{n} W_{ij}X_j - T\right]$$

where W_{ij} is the weight attached to input X_j, and T is a threshold value. This means that the neuron can be "excited" if the total excitation which it receives reaches or exceeds the threshold value.

Hebb (Hornik, Stinchombe & White, 1989) also worked with neural networks and developed methods allowing neural networks to learn, whereby the weight on an activated connection between two neurons was proportional to the correlation of activity in the two neurons, reflecting the association of units that typically take part in the same path.

2.5 THE PERCEPTRON

In 1958, Rosenblatt (1958) showed how a network of M-P neurons with adjustable weights could be trained to classify certain sets of patterns. These networks were termed "Perceptrons". Initially, the weight settings are arbitrary, so that any stimulation of the network produces an arbitrary response. To obtain the desired response, the weights are adjusted in a procedure known as "training" or "learning". Rosenblatt popularised the notion of the perceptron, a two layer neural network,

where the second layer consisted of a single unit with a threshold activation function, producing binary output. This could classify inputs and learn weights automatically.

This network can be described as follow:

$$S_i = F\left[\sum_{j=1}^{n} W_{ij}X_j - T\right]$$

If $S_i = +1$, then select Class A,

If $S_i = -1$, then select Class B.

where the single node computes a weighted sum of the input elements X_j, subtracts a threshold T and passes the result ($\Sigma W_{ij} X_j - T$) through a hard limiting non-linear function F such that the output S_i is either +1 or -1. The decision rule is to respond with class A if the output is +1 and with class B if the output is -1 [Fig 2-4].

M. Minsky and S Papert (1969) of Massachusetts Institute of Technology published a book "Perceptrons: the principles of computational geometry" which suggested that an artificial neural network system, like the Perceptron, which could be trained to recognise simple images such as letters of the alphabet and which was designed by F. Rosenblatt (1958) of Cornell University, had severe limitations in solving the "Exclusive-OR [XOR]" problem, because the linear function of the perceptron could only form a single decision boundary. Their publication had a devastating effect on the research into artificial neural networks in the early 1970's. The "Exclusive-OR" function has an output which is logically true when the two inputs are opposite, and differs from the

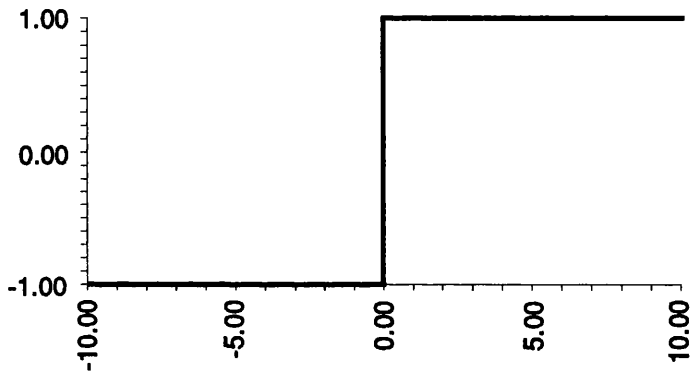


FIG 2-4. Hard limiter non-linear function for the perceptron

"Inclusive-OR" function which is logically true when either input is true, i.e.

XOR		OR	
INPUT	OUTPUT	INPUT	OUTPUT
1 1	0	1 1	1
1 0	1	1 0	1
0 1	1	0 1	1
0 0	0	0 0	0

2.6 MULTI-LAYER PERCEPTRONS

The XOR problem was solved by adding an extra "hidden" layer of neurons between the input layer and output layer of the network. However, the addition of this extra layer caused problems as the correlation between the outputs from the hidden layer and the outputs from the network were not known. Multiple layer networks are feed-forward networks with one or more layers of nodes between the input and output nodes. In the standard feed forward neural network, neurons are arranged in input, hidden, and output layers with interconnections each of which is assigned a weighting factor [Fig 2-5]. Multi-layer networks overcome many of the limitations of single-layer perceptrons, but were generally not used in the past because effective training algorithms were not available. This has recently changed with the development of new training algorithms. In 1986, Rumelhart et al (1986) developed an effective training algorithm (back-propagation delta rules) to adjust the weights of these hidden neurons (vide infra).

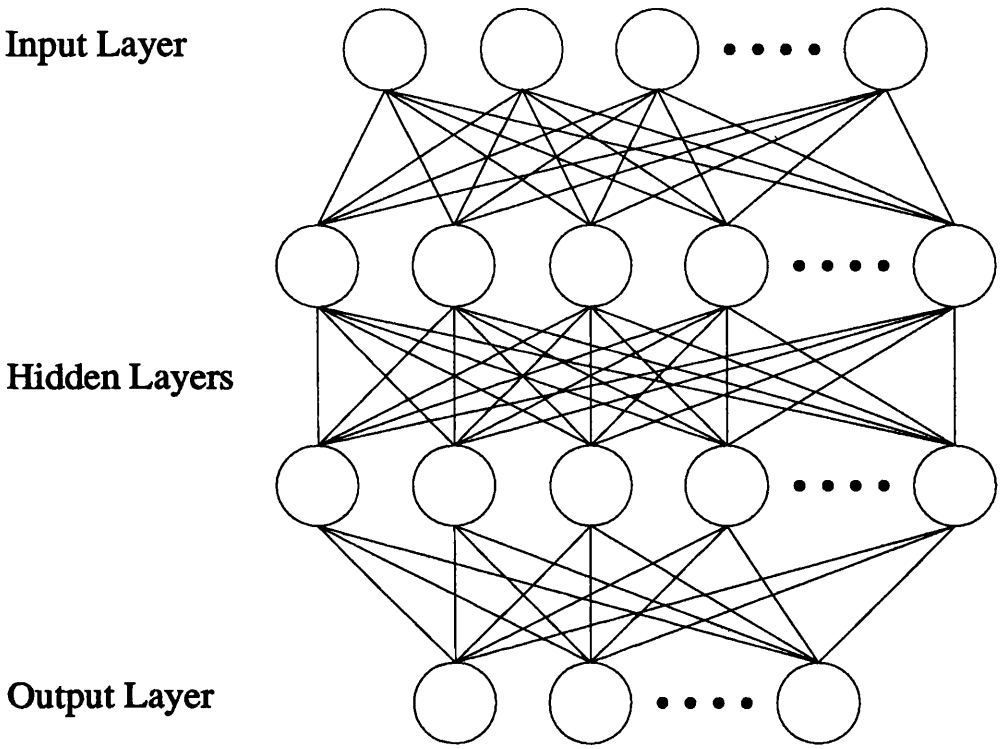


FIG 2-5. Basic Structure of a Multi-Layer Neural Network

2.7 THE ARCHITECTURE OF NEURAL NETWORKS

There are two kinds of basic neural network connections: "feed-forward" and "feed-back" connections (Lippmann, 1987). It might be inferred that a feed-forward network simulates learned behaviour, and a feed-back network simulates instinctive behaviour.

2.7.1 Feed-forward Network

In a "feed-forward" connection network, neurons in any layer can be connected only to neurons in those layers above, i.e. neurons only take their inputs from the previous layer and only send their outputs to the next layer. Assuming the nth layer is not the output layer, the output from any neuron in the nth layer may only be connected to neurons in layer (n+1) and above. For the same neuron, the connections can be spread across many layers. Neurons in a given layer do not connect to each other. The main advantage of the "feed-forward network" is that outputs from the networks can be computed more quickly by utilising a single forward pass. There is no time delay while the neurons interact with themselves and settle into a stable state.

2.7.2 Feed-back Network

In a "feed-back" connection network, the output from a neuron can be connected to any other neuron in any layer of the network, even itself. Since a neuron is allowed to be connected to any other neuron, feed-back networks usually have only one layer. The internal neuron outputs must be computed iteratively until they reach a stable state to reach the output of a feed-back connection network. There is no way to predict

how long this will take. Feed-back networks are not trained, but are constructed. The design of the architecture of the network is usually determined by the problem to be solved.

2.8 LEARNING ALGORITHMS OF THE NEURAL NETWORK

The learning algorithm is the mechanism used to adjust the weights of the neurons in the network so that the association or correlations between the input and output can be learned. The learning algorithms can be categorised as either "supervised" or "non-supervised" (Lippmann, 1987; Soucek, 1989).

2.8.1 Supervised Learning Algorithm

The operation of a supervised learning algorithm requires a set of inputs and desired network outputs in order to train the network to correlate these patterns. This algorithm is applied iteratively to the network to adjust the weights of the connections in order to reduce the overall network errors until all the correlations between the input patterns and output patterns are correctly learned or trained. In this algorithm, some kind of external influence is present during training to tell the network whether or not its output was correct.

2.8.2 Non-Supervised Learning Algorithm

A non-supervised learning algorithm does not function on the basis of introducing the true result to the network. Instead, this algorithm is based on competitive learning. Each neuron competes for the control of an

active input. Neurons receiving the largest input achieve their maximum value while all other neurons are forced to their minimum. These few neurons will try to learn all the associations without allowing the majority of the neurons to participate in the training. Every neuron then learns by shifting weight from its inactive to its active inputs (Soucek, 1989). Therefore, the capacity and generalisation abilities of this algorithm are not very high.

2.8.3 Back-Propagation Learning Algorithm

The most frequently used supervised learning algorithm is the "back-propagation" learning algorithm (Jones & Hoskins, 1987; Kolen & Pollack, 1990). This learning algorithm is an iterative gradient algorithm designed to minimise the mean square error between the actual output and the desired output of each neuron in a multiple layer feed-forward network. The algorithm operates on the basis that inputs are presented to the network and the resulting outputs computed.

The rate and time for convergence of the back-propagation learning algorithm is influenced by several factors. For instance, the initial assignment of weights to connections in feed-forward networks which employ the back-propagation learning algorithm has been shown to have a significant impact on the time taken for convergence (Kolen & Pollack, 1990). The initially assigned weights introduce either correct or incorrect associations between certain input and output patterns. Therefore, the "back-propagation" learning algorithm must not only learn the associations between the newly implanted inputs and output patterns, but also "unlearn" any previously incorrect associations.

It has been suggested that the network weights should be initialised to small random values between -0.5 and +0.5 to achieve best results. In addition, the utilisation of limited precision hardware can introduce errors into the connection weights themselves (Stevenson, Winter & Widrow, 1990). The rate of convergence of the "back-propagation" learning algorithm has been shown to be influenced by the selection of the learning rate parameters (Lippmann & Hoskin, 1987; Kolen & Pollack, 1990).

The choice of values for the learning rate parameters (gain and tolerance) can have the greatest effect on convergence. Large values can make the networks fail to converge because the algorithm continuously overshoots the target by over-correcting the weight errors. Likewise, very small values can result in the same problem because the weights are changing too little. The convergence factor (gain) used in the back-propagation learning algorithm of the network determines how fast the back-propagation converges on the target output value. A large gain may cause the back-propagation to overshoot the target continuously, thus causing wild fluctuation in the network output. However, a gain which is too small may cause a network to take longer to converge. The time for convergence of the back-propagation learning algorithm is ultimately determined by the difficulty of the problem to be solved.

Training tolerance of the network is the value by which the actual output from the network is allowed to differ from the target output, though still being close enough to be regarded as correct. This is a real number between 0.0 and 1.0.

A normalisation procedure in the training process is used to scale the data from the training set. A maximum value and a minimum value for each input parameter are determined. Subsequently the training data are scaled according to such extrema in order to obtain input values ranging from 0 to 1 for each parameter.

2.9 CHARACTERISTICS OF THE NEURAL NETWORK

The learning and processing capability of an artificial neural network is determined by the architecture of the neuronal interconnections and the training algorithm applied. The behaviour of the network is determined by the weights (signal modulation characteristics) of the interneuronal connections which are established during the training process. Rather than being programmed, the artificial neural network is trained by presenting sets of inputs together with the corresponding well classified outputs that the trainer expects the networks to associate with the inputs.

The differences between the artificial neural network and the conventional rule-based deterministic logic approach have been listed as follows (Jones & Hoskin, 1987):

1. The knowledge of a neural network lies in the inter-layer and/or inter-neuron connections and their initial weight assignments. In contrast, much of the knowledge of an expert system lies in the rules and logic.

2. A neural network is driven by the activation that passes from neurons to other neurons in the other layers. In contrast, an expert system is driven by symbols generated as a consequence of rule-firing.

The design or topology of an artificial neural network is to determine the number of layers in the network and the number of neurons in each layer. The number of neurons in the input layer is a reflection of the number of measurements used and is subject to the bias induced from the selection, while the number of neurons in the hidden layer and number of hidden layers are determined by trial and error during the initial development of a neural network.

However, the more neurons used, the greater the capacity of the network to learn and store associations, but not necessarily to yield a better result. With too many hidden neurons, a neural network can simply memorise the correct response to each pattern in its training set instead of learning a general solution. By limiting the size of the hidden layers, the neural network is forced to develop appropriate feature detectors to classify large sets of input patterns efficiently. These general purpose feature detectors are more likely to be relevant to novel inputs, so the neural network performs better when the size of the hidden layer is reduced. This can not only increase the rate at which the computer can simulate the network (i.e. reduce the time needed for the training), but also improve the performance of the network (Touretzky & Pomerleau, 1989).

The hidden layer is composed of neurons that are connected to neurons in other layers, but do not interact with the environment directly. The hidden layers enable the creation of an internal representation of the problem and the correlations found in the training set.

Superficially, artificial neural networks seem to be similar to other statistically derived algorithms in that a set of data is input and the

answer is output. However, the relationship between the input and the output is neither defined nor quantified explicitly in an artificial neural network. Hence the neural network does not identify the relative contribution of each input to the output. The pattern recognition process is similar to the way in which a child learns to distinguish certain differences between two animals, e.g. cats and dogs. After being fed with examples of each pattern along with the desired output, the neural networks correctly learn to differentiate each pattern.

Artificial neural networks have the capabilities of learning to recognise and identify patterns based on past experience without explicitly identifying the basis on which this task is performed. Unlike the human brain, artificial neural networks are not susceptible to bias toward recent or unusual events and do not suffer from emotional influence, fatigue, and distraction. Although it is true that a neural network does not have the capability to create or to perform intuitive reasoning that could prevent obvious mistakes, its clinical application can still be evaluated like any other new technology.

The advantages and disadvantages of artificial neural networks can be summarised as follows:

Advantages:

(1) Artificial neural networks perform well with incomplete data input.
(2) There is no longer a requirement for lengthy program development and maintenance times.
(3) There is no "debugging" for incorrect responses.
(4) There is no need for criteria. Essentially, the neural networks will form their own internal representations (criteria).
(5) The processing of neurons can be carried out in parallel.

Disadvantages:

(1) The reasoning process by which a neural network comes to its decision cannot be known.
(2) The performance of a neural network is dictated by the composition of the training set.
(3) Deficiencies or inadequacies in neural networks can only be noticed through actual use.
(4) There are no rules for determining the optimal number of hidden layers and the neurons within them.
(5) It is possible for a neural network to become highly tuned to the training data.

2.10 CURRENT CLINICAL APPLICATIONS

In recent years, artificial neural networks have been successfully applied in clinical diagnosis especially for pattern recognition, e.g. Radiological Images (Boone, Gross & Greco-Hunt, 1990), Quantitative Cytology (Dytch & Wied, 1990), Electromyography (Spitzer et al, 1990). Neural networks have also been applied as an aid for clinical decision making in acute myocardial infarction (Baxt, 1991a & 1991b) and in dermatological differential diagnosis (Yoon et al, 1989). In computerised electrocardiography, applications have been focused mainly on ECG wave form recognition (Casaleggio, Morando & Ridella, 1991) and diagnostic classification e.g. the differentiation of right and left ventricular hypertrophy (Bortolan, Degani & Willems, 1991), the localisation of myocardial infarction (Bortolan, Degani & Willems, 1992), the classification of electrocardiographic ST-T segments (Edenbrandt, Devine & Macfarlane, 1992), the recognition of left ventricular strain (Devine & Macfarlane, 1993), detection of atrial fibrillation (Yang, Devine & Macfarlane, 1993b) and myocardial infarction (Yang, Devine & Macfarlane, 1994d). Artificial neural networks have also been used for

digital Holter ECG data compression (Iwata, Nagasaka & Suzumura, 1990). Certain results from these studies are claimed to be comparable to those of experts (Baxt, 1991, Edenbrandt, Devine & Macfarlane, 1992) and slightly better than deterministic logic with respect to sensitivity of diagnosing atrial fibrillation (Yang, Devine & Macfarlane, 1993c).

Therefore, it was thought that the application of artificial neural networks in computer-assisted electrocardiographic interpretation might be of value in improving the performance of diagnostic software by introducing objective and unbiased techniques.

CHAPTER 3

NORMAL LIMITS OF THE DERIVED VECTORCARDIOGRAM IN CAUCASIANS AND CHINESE

3.1 INTRODUCTION

As derived vectorcardiography is a relatively new technique for synthesizing the vectorcardiogram from the conventional 12-lead electrocardiogram (Edenbrandt & Pahlm, 1988a), no large scale study of normal limits in either Caucasians or Chinese has as yet been published. Therefore, the present study was undertaken to establish age, sex and race dependent normal limits of the derived vectorcardiogram in Caucasians and Chinese.

In the conventional 12-lead ECG, significant differences between normal Chinese and Caucasians exist in QRS and T amplitudes, as well as in the transitional zone location, for corresponding age and sex groups (Macfarlane, Chen & Chiang 1988; Macfarlane & Lawrie, 1989), necessitating separate normal limits for men and women in different races. Since the derived VCG is synthesized from the standard 12-lead ECG and represents different displays of the same information, logically the effects of age, sex, and race should also be expected in the derived VCG.

Computer assisted interpretations based on combining 12-lead ECG and VCG or derived VCG diagnoses of the CSE data base (1220 ECGs) have

been shown by Kors et al (1992) to be statistically significantly better than each diagnosis separately. The total accuracy of ECG + VCG and ECG + derived VCG were 74.2% and 73.6%, respectively, compared to 69.8% and 70.2% for ECG and VCG, respectively. These authors also showed that it is not necessary to record a separate Frank VCG, since the derived VCG performed as well as the original Frank VCG, so it was thought that, in future, the use of age, sex, and race related derived VCG criteria based on well developed normal limits might enhance the diagnostic accuracy of the Glasgow Program.

The derived VCG, however, has recently been claimed to be different in vector loop configuration, though not in diagnostic content, compared to the Frank VCG as evaluated by statistical techniques (Kors et al, 1990b; Rubel, Benhadid & Fayn, 1992). These studies also confirmed that certain variations exist between the derived VCG and Frank VCG measurements. Therefore, it was felt necessary to establish the normal limits of the derived VCG parameters.

Most diagnostic value of the derived VCG comes from the configuration of the vector loop and not the scalar parameters of the orthogonal leads, and hence this study concentrates more on the quantitative derived VCG loop parameters.

3.2 MATERIALS AND METHODS

3.2.1 MATERIALS

A total of 1555 Caucasians (884 men and 671 women), aged between 16 and 64, were recruited for this study between March 1981 and May

1991 by the Department of Medical Cardiology in Glasgow Royal Infirmary. Most were apparently healthy employees of Glasgow District Council and Strathclyde Regional Council, Scotland. The cohort also included 30 individuals with atypical chest pain investigated by coronary arteriography which demonstrated normal coronary arteries.

A total of 503 Chinese individuals (248 men and 255 women), aged between 18 and 81, were recruited between July 1985 and December 1985 by the Department of Cardiology in Taipei Veterans General Hospital. The subjects were all residents of Taipei, Taiwan. All individuals agreed to participate in the project voluntarily, and those who were younger than 60 years old were selected from the general population. None was referred from a physician's office. On the other hand, volunteers aged 60 and older were recruited from the surgical wards of the Veterans General Hospital in Taipei, Taiwan.

All individuals involved in the present study agreed to participate in the project voluntarily, and other than 30 individuals with atypical chest pain, were selected on the basis of there being a complete absence of any history or physical signs suggestive of cardiovascular disease or any other abnormalities known to affect the cardiovascular system, such as diabetes mellitus, endocrine diseases, and chronic obstructive pulmonary diseases, etc. More than 30% of the Caucasians also had a Chest X ray and/or Echocardiographic examination performed. For individuals younger than 60 years old, it was thought unnecessary to record a chest radiograph or to perform blood biochemistry in keeping with the recommendation of Startt/Selvester (1986).

The age and sex distributions of both races are shown in Table 3-1.

AGE		<30	30-39	40-49	50-59	60+	Total
Caucasian	M	242	217	210	177	38	884
	F	304	131	97	120	19	671
Chinese	M	55	54	47	49	43	248
	F	49	61	53	48	44	255

Table 3-1 Age and sex distribution of the population studied. (M=male, F=female.)

Certain individuals with ECG patterns of right bundle branch block, poor R wave progression over praecordial leads (Zema et al, 1980), or minor intraventricular conduction defects were included in this study, if their medical history, physical examination, chest X-ray, and echocardiographic investigation were all normal. Zema & Kligfield (1979 a & 1979b) have claimed that ECGs demonstrating poor R wave progression and reversed R wave progression comprise approximately 8% and 2% respectively of routine adult hospital tracings. It is also well known that the incidence of the ECG right bundle branch block pattern in an apparently healthy population is 0.5% in men and 0.2% in women (Ostrander et al, 1965), while the true incidence of the other normal variants is unknown. There were no individuals with an ECG pattern of either left bundle branch block or pre-excitation syndrome in this study.

3.2.2 METHODS

For the purpose of the present study, the 12-lead ECGs were recorded in Taiwan on a Siemens-Elema MINGOREC 4 electrocardiograph, which incorporated an analogue-to-digital converter operating at 500 samples per second. A 1-mV calibration pulse generator is integral to the MINGOREC 4, and hence all measurements were referenced to that

signal. All leads were recorded simultaneously and were transferred to a digital cassette which is part of the MINGOREC 4. Subsequently the cassettes were sent to Glasgow Royal Infirmary for analysis by well established software (Macfarlane et al 1990) which incorporated the inverse Dower equations (Edenbrandt & Pahlm 1988a) for the derivation of the three orthogonal X, Y, Z, leads from the 12 conventional electrocardiographic leads, so that the derived vectorcardiographic loops could be printed out together with selected vectorcardiographic measurements. The Caucasian ECGs were recorded using a locally developed electrocardiograph (Watts & Shoat, 1987), which also had a sampling rate of 500 samples/second.

The Glasgow program calculates a median beat from which all amplitudes and durations can be measured but also from which vectorcardiographic measurements and loops can be produced (Macfarlane et al, 1990b). The program accepts all beats in an 8 or 10 second recording and from these the median beat is synthesized. All electrocardiographic and vectorcardiographic measurements were recorded following the recommendations of the American Heart Association's committee on electrocardiography (1975) and of the Common Standards for Quantitative Electrocardiography working party (1985).

Measurements were obtained of the amplitudes and durations of the components of the three scalar orthogonal leads X, Y, Z (Lead Z in the present study is directed positively anteriorly) as well as of the various vectorcardiographic parameters of the frontal, horizontal, and right sagittal planes. These included the directions of inscription of the QRS

vector loops, the magnitude of the initial 20 and 30 millisecond QRS vectors, and the orientation of the initial 20 millisecond QRS vector.

The magnitude of the maximal QRS and T vectors in space and their projection onto the frontal, horizontal, and right sagittal planes, were determined. In this case, the X_M, Y_M, Z_M, values which constitute the maximal spatial vector with magnitude $(X_M^2 + Y_M^2 + Z_M^2)^{1/2}$ are known and the planar projections are calculated therefrom, e.g. the projection of the maximum vector on to the frontal plane, has

$$\text{magnitude} = (X_M^2 + Y_M^2)^{1/2}$$
$$\text{orientation} = \text{Tan}^{-1} (Y_M/X_M)$$

On the other hand, the maximum frontal plane vector magnitude is calculated from the maximum of all

$$(x_i^2 + y_i^2)^{1/2}$$

where (x_i , y_i) are pairs of simultaneous measurements in X and Y respectively. If X_F, Y_F represent the coordinates when the frontal plane vector has its maximum value then

$$\text{magnitude of maximal frontal plane vector} = (X_F^2 + Y_F^2)^{1/2}$$
$$\text{orientation of maximal frontal plane vector} = \text{Tan}^{-1} (Y_F/X_F)$$

Similar reasoning holds for the (X, Z) horizontal and (Y, Z) sagittal planes.

A bite in the vectorcardiogram has been defined in several ways, such as (i) a scallop or bulge of the QRS loop (Selvester et al, 1968), (ii) a deformity of the QRS loop (Selvester, Palmersheim & Pearson, 1971), (iii) a displacement from a smoothly transcribed loop (Selvester et al, 1968), (iv) an indentation in a smooth vectorcardiographic loop (Zoneraich & Zoneraich, 1977), or (v) a sudden brief introflection of the QRS loop profile otherwise normally smooth and convex (Vitolo et al, 1982).

A vectorcardiographic bite is considered to be present if a sector of the vector loop is found to rotate in the opposite way to the main body of the vector loop, e.g. a counterclockwise inscribed sector in an otherwise clockwise inscribed loop. Duration and amplitude are the most common measurements for the quantification of a vectorcardiographic bite. The amplitude of the bite is the largest perpendicular distance from the line through the starting point and end point of the bite (Fig 3-1). Bite amplitude and duration were measured using software kindly provided by the Department of Clinical Physiology in Lund, Sweden (Edenbrandt et al, 1989).

The various parameters were computed automatically with results being output to a data file which was subsequently interfaced to the BMDP Statistical Package for analysis of the data stratified by age sex, and race. In both populations, it was used to derive mean, standard deviations and 96-percentile ranges, which are approximately equivalent to the mean ± 2 standard deviations in a normally distributed population (Simonson, 1961). The descriptive program P7D was used to list the means and standard deviations of the various derived VCG parameters so that 96 percentile ranges of normal limits could be determined by excluding two

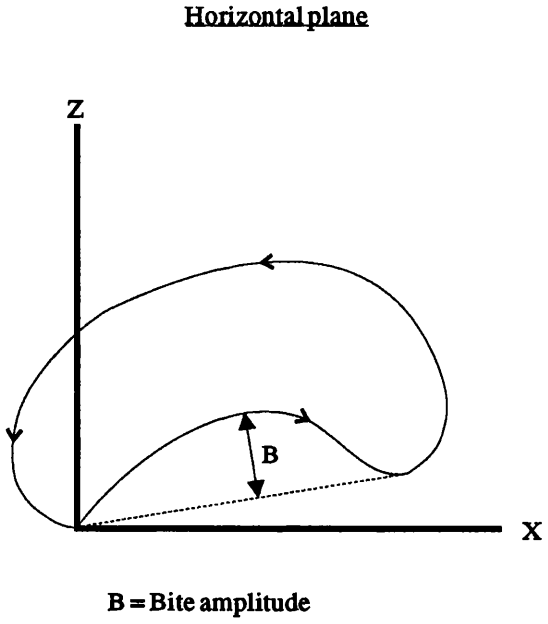


Fig. 3-1 Bite amplitude is defined as the largest perpendicular distance from a line connecting the starting and ending points to the sampling points inside the bite.

percent of values at each extreme of the distribution. The program P6D was used to produce bivariate scatter plots of various derived VCG parameters against age and to provide a correlation between the age of each patient and the parameter under study, as well as a significance value. The two-sample t-test program P3D was used to determine the significance of the differences of derived VCG parameters among various age groups and between the two sexes and races.

3.3 RESULTS

The results of the present study are shown for Caucasians and Chinese separately in Figures 3-2 to 3-10 and Tables 3-2 to 3-29. Tables 3-6 to 3-29 are presented in Appendix 1. Most of the Tables and Figures are self explanatory. Two sets of values are given for each parameter in the tables. On the upper row are shown the mean results with their standard deviations. Since the distribution of ECG data was found not to be normal (Simonson, 1961), 96-percentile ranges were also determined for each measurement. These 96-percentile ranges are shown on the lower row. Such a range is more appropriate than that obtained from a mean ± 2 standard deviations when a normal distribution is present (Simonson, 1961).

The ranges of the maximal QRS and T vector angles in the three planes were presented with due attention being paid to the modal distribution. Only the 96-percentile ranges of the maximal QRS and T vector angles are presented. The ranges of the initial 20 milliseconds vector angle are also presented in the same form. Their means are also shown in Tables 3-14 & 3-15 as their distributions were in a reasonably narrower range.

3.3.1 Results of Vectorcardiographic Measurements

1. Direction of the inscription of the QRS vector loops in frontal, horizontal, and right sagittal planes [Fig 3-2] (Tables 3-2 & 3-3)

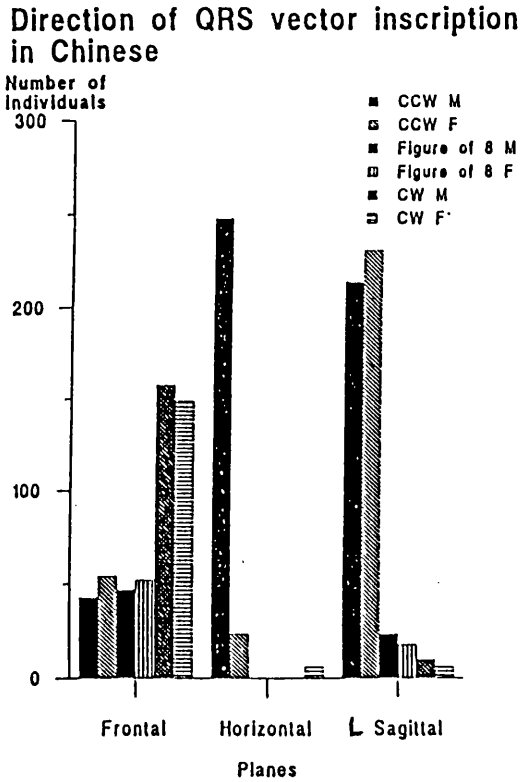
Both in Caucasians and Chinese, the dominant direction of inscription in the frontal plane, horizontal plane, and right sagittal plane is clockwise, counterclockwise, and clockwise respectively. In Chinese, there was no clockwise or figure of 8 inscription for men in the horizontal plane. In Caucasians, there were all varieties of inscriptions in all planes. The trend of the distribution of the directions of inscription was the same in both races.

Planes		CCW	Figure of 8	CW
Frontal	M	203 [22.9%]	169 [19.1%]	512 [58.0%]
	F	146 [21.8%]	170 [25.3%]	355 [52.9%]
Horizontal	M	867 [98.1%]	8 [0.9%]	9 [1.0%]
	F	657 [97.9%]	8 [1.2%]	6 [0.9%]
R Sagittal	M	32 [3.6%]	74 [8.4%]	778 [88.0%]
	F	11 [1.6%]	26 [3.9%]	634 [94.5%]

Table 3-2. Direction of inscription of the QRS vector loop in Caucasians. (CW:Clockwise, CCW: Counterclockwise)

Planes		CCW	Figure of 8	CW
Frontal	M	43(17.3%)	47(18.9%)	158(63.7%)
	F	54(21.1%)	52(20.3%)	149(58.4%)
Horizontal	M	248(100%)	0	0
	F	249(97.6%)	0	6(2.4%)
R Sagittal	M	10(4%)	24(9.7%)	214(86.3%)
	F	6 (2.4%)	18(7%)	231(90.6%)

Table 3-3 Direction of inscription of the QRS vector loop in Chinese. (CW:Clockwise, CCW: Counterclockwise)



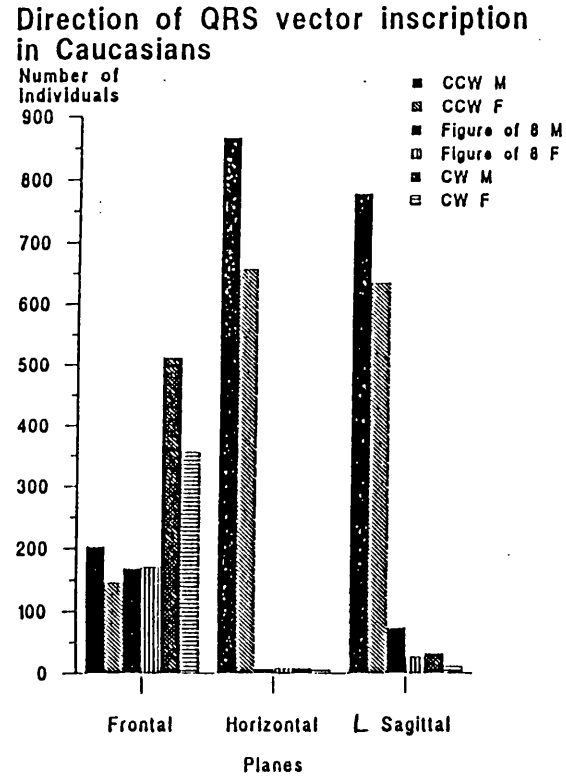


Fig.3-2 Direction of the inscription of the QRS vector loops in frontal, horizontal and left sagittal planes in Caucasians (left panel) and Chinese (right panel).

2. Magnitude of the maximal spatial QRS vector [Fig 3-3] (Tables 3-4 & 3-5)

The magnitude of the maximal spatial QRS vector in men is greater than that of women in each age group both in Caucasians ($p<0.001$) and Chinese ($p<0.001$). The trend of diminution of the maximal spatial QRS vector magnitude over the total age range was significant for both sexes not only in Chinese ($p<0.001$ in men and $p=0.015$ in women) but also in Caucasians ($p<0.001$). The magnitude of maximal spatial QRS vector is greater in Caucasians than in Chinese in the groups aged less than forty, yet in the groups aged forty or more, there was no such relationship; on the contrary, in the older groups, the maximal spatial vector magnitude of Chinese was greater than that of Caucasians. The differences between Caucasians and Chinese are statistically significant in all age groups ($p=0.0003$).

	M		F	
AGE	M±SD	Range	M±SD	Range
<30	2.39±0.62	1.28-3.89	1.76±0.46	0.82-2.79
30-39	2.07±0.58	1.05-3.42	1.75±0.47	0.94-2.74
40-49	1.79±0.49	0.94-2.90	1.46±0.41	0.79-2.26
50-59	1.65±0.45	1.00-2.69	1.46±0.37	0.78-2.22
60+	1.57±0.43	0.96-2.34	1.37±0.51	0.82-2.32

Table 3-4. Magnitude of maximal spatial QRS vector in Caucasians(mV). [M=male, F=female]

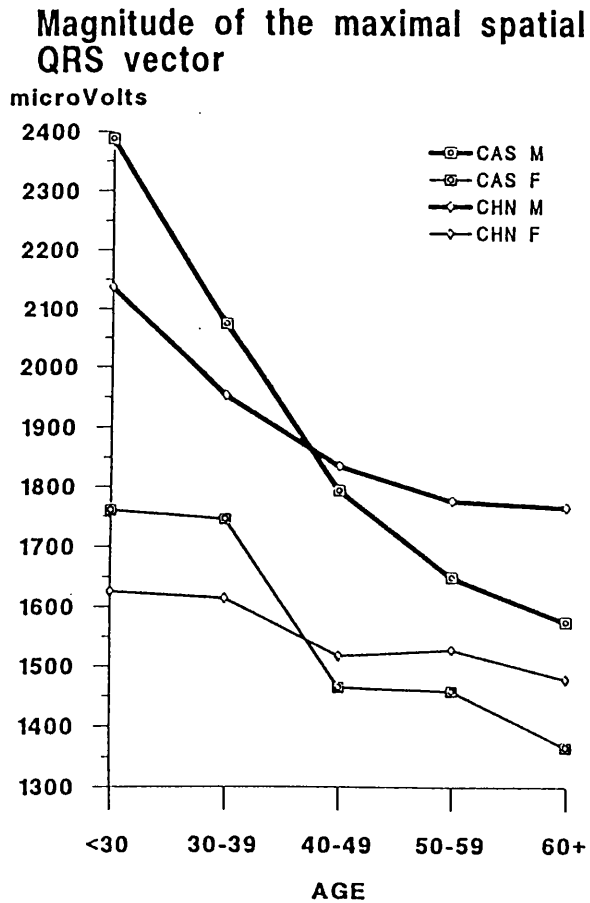


Fig.3-3 Relationship between age and mean magnitude of the maximal spatial QRS vector in Caucasians and Chinese.

	M		F	
AGE	M±SD	Range	M±SD	Range
<30	2.14±0.59	1.23-3.27	1.63±0.41	0.93-2.34
30-39	1.95±0.51	1.24-2.90	1.61±0.34	0.92-2.37
40-49	1.83±0.48	0.96-2.72	1.52±0.40	0.95-2.43
50-59	1.78±0.48	1.03-2.98	1.53±0.37	0.97-2.29
60+	1.77±0.51	1.05-3.01	1.48±0.40	0.95-2.43

Table 3-5. Magnitude of maximal spatial QRS vector in Chinese(mV). [M=male, F=female]

3. Magnitude of the projection of the maximal QRS vector onto the frontal, horizontal, and right sagittal planes [Fig 3-4] (Tables 3-6 & 3-7)

The magnitude of the projection of the maximal QRS vector onto the frontal and horizontal planes is greater in Caucasians than in Chinese in the groups aged less than forty, yet in the groups aged forty or more, it is greater in Chinese than in Caucasians, whereas, there is no such relationship in the right sagittal plane. The magnitude in men is significantly greater than that in women in both Caucasians ($p<0.001$) and Chinese ($p<0.001$) in all planes.

The age dependent magnitude decreases were found to be statistically significant for men in all planes both in Caucasians and Chinese ($p<0.001$ for Caucasians in all planes; $p=0.002$ for Chinese in the frontal plane and $p<0.001$ in other planes). In Chinese women, the age-related diminution of the QRS vector magnitude was found to be statistically significant only in the right sagittal plane ($p<0.001$); yet in Caucasian women, these changes were statistically significant in all

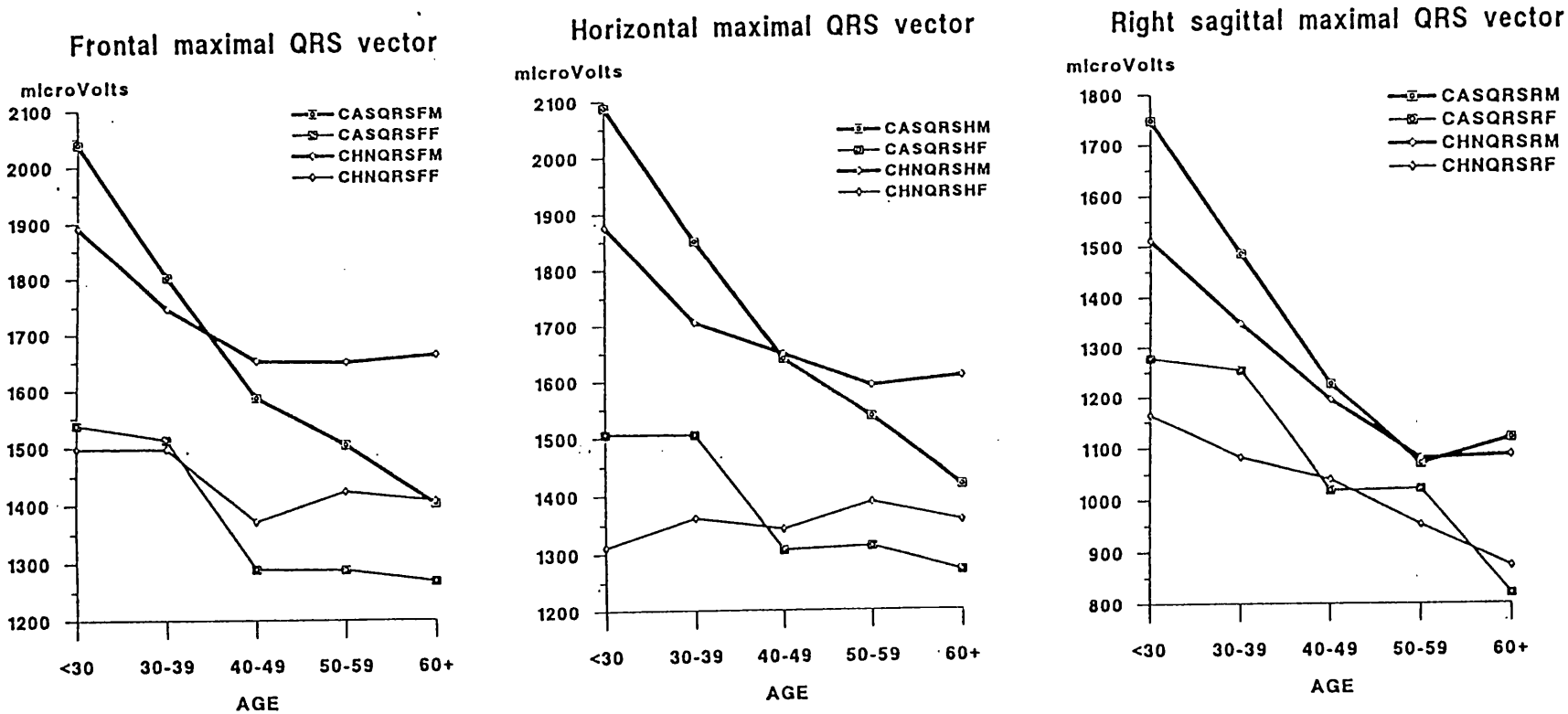


Fig. 3-4 Relationship between age and mean magnitude of the projection of the maximal QRS vectors onto the frontal (left panel), horizontal (middle panel) and right sagittal planes (right panel).

planes ($p<0.001$). For differences between Chinese and Caucasians, in the horizontal and right sagittal planes, the differences are statistically significant ($p<0.0001$), whereas in the frontal plane, there is no such significance.

4. Magnitude of the projection of the maximal T vectors onto the frontal, horizontal, and right sagittal planes [Fig 3-5] (Tables 3-8 & 3-9)

The magnitude of the projection of the maximal T vector onto all planes decreases significantly with advancing age for both sexes in Caucasians ($p<0.001$) as well as in Chinese ($p<0.001$). The magnitude in men is significantly greater than that in women in both races in all planes ($p<0.001$). There are significant differences between Caucasians and Chinese in all planes ($p=0.0037$ in the frontal plane, $p<0.0001$ in the horizontal and right sagittal planes).

5. Ranges of angle of the projection of the maximal QRS vector onto the frontal, horizontal, and right sagittal planes [Fig 3-6] (Tables 3-10 & 3-11)

Both for Caucasians and Chinese, the angle of the projection of the maximal QRS vector in the frontal plane was found to be distributed in a much narrower range than in the horizontal and right sagittal planes.

The maximal QRS vector angle decreases significantly with advancing age in the frontal plane ($p<0.001$) for both sexes and races, while it increases significantly in the horizontal and right sagittal planes ($p<0.001$) only in men for both races (i.e. the maximal QRS vector for

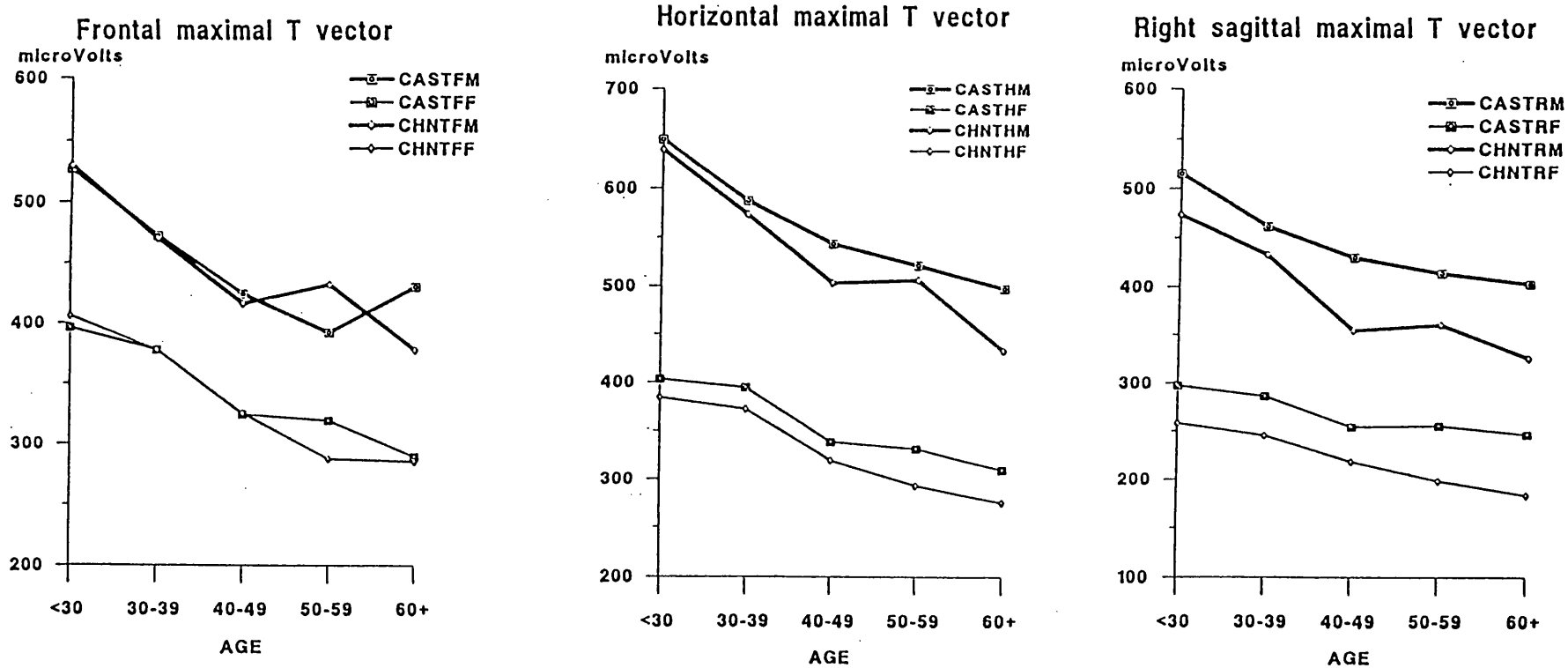


Fig.3-5 Relationship between age and mean magnitude of the projection of the maximal T vectors onto the frontal (left panel), horizontal (middle panel) and right sagittal planes (right panel).

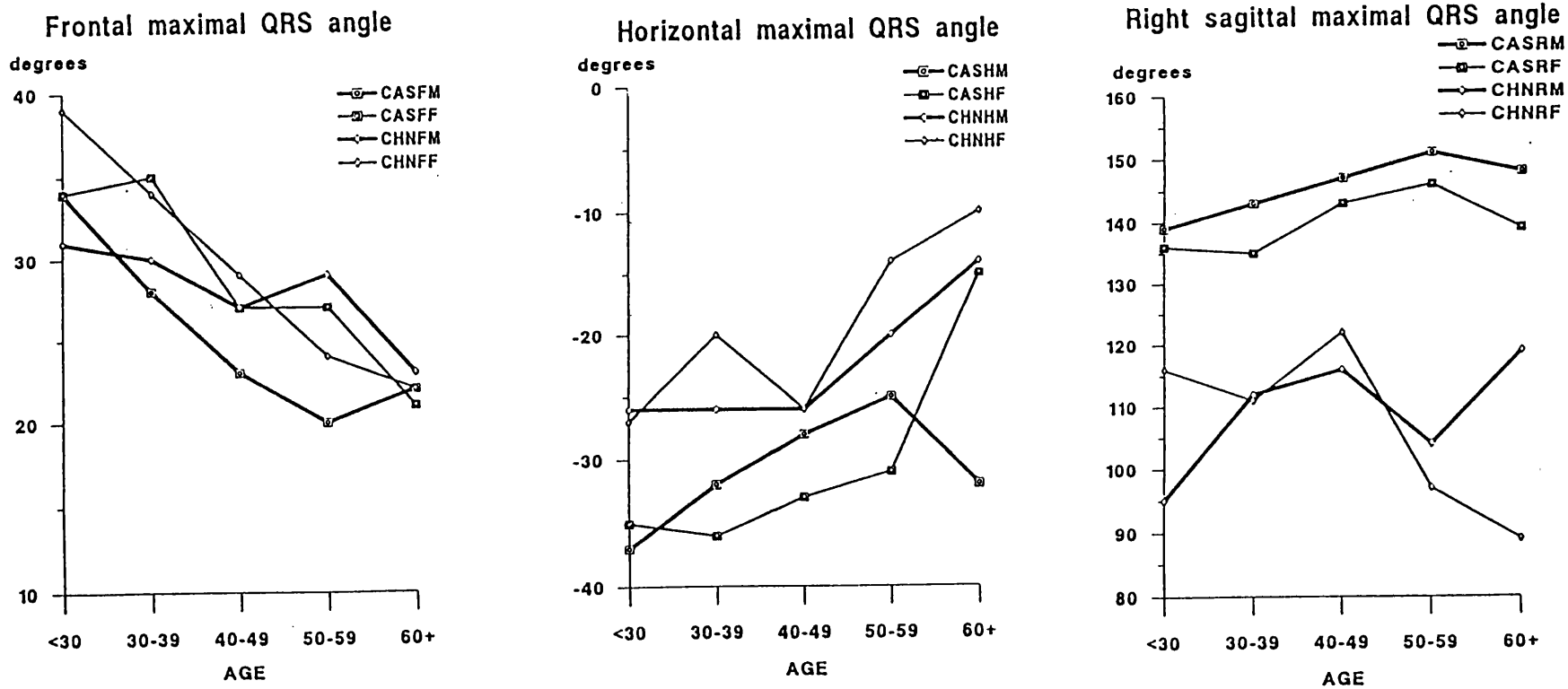


Fig.3-6 Relationship between age and mean angles of the projection of the maximal QRS vector onto the frontal (left panel), horizontal (middle panel) and right sagittal planes (right panel).

both races is directed more superiorly and anteriorly in the older age group). In both races, the maximal QRS vector angle in men is directed more superiorly and anteriorly than those in women.

6. Ranges of angle of the projection of the maximal T vector onto the frontal, horizontal, and right sagittal planes [Fig 3-7] (Tables 3-12 & 3-13)

The maximal T vector angles of women are directed more inferiorly and posteriorly than those of men in both Caucasians and Chinese. The trend of the maximal T angle is the same as that of the maximal QRS angle, except for those aged 60 or over. This apparent change may be related to the smaller number of individuals studied in the Caucasians aged 60 or over compared to other age groups.

7. Magnitude of initial 20 milliseconds vector in frontal, horizontal, and right sagittal planes (Tables 3-14 & 3-15)

For Caucasians, the magnitude of the initial 20 milliseconds vector in the frontal plane increases significantly with advancing age in both sexes ($p=0.002$ in men, $p<0.001$ in women). In the horizontal plane, the vector magnitude decreases significantly with advancing age in both sexes ($p< 0.001$ in men, $p=0.019$ in women). In the right sagittal plane, the vector magnitude decreases significantly with advancing age in both sexes ($p<0.001$ in men and women).

For Chinese, the magnitude of the initial 20 milliseconds vector in the frontal plane increases significantly with advancing age in both sexes ($p=0.005$ in men, $p=0.003$ in women). In the horizontal plane, no

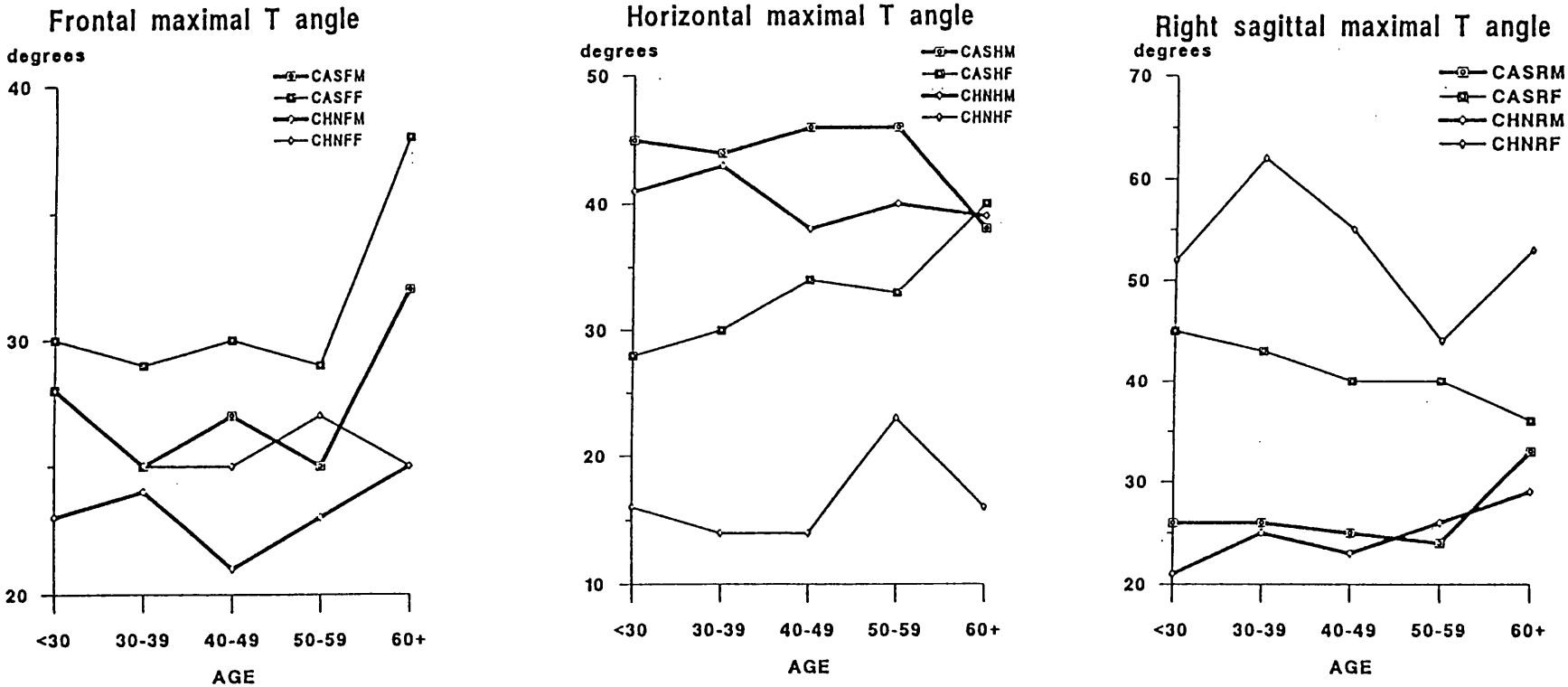


Fig.3-7 Relationship between age and mean angles of the projection of the maximal T vector onto the frontal (left panel), horizontal (middle panel) and right sagittal planes (right panel).

correlations with age were observed. In the right sagittal plane, the vector magnitude decreases significantly with advancing age in both sexes ($p=0.003$ in men, $p=0.002$ in women).

In all planes, the differences between Caucasians and Chinese are statistically significant in all age groups ($p<0.0001$ in the frontal and horizontal planes, $p=0.0195$ in the right sagittal plane).

8. Magnitude of initial 30 milliseconds vector in frontal, horizontal, and right sagittal planes (Tables 3-16 & 3-17)

For Caucasians, the magnitude of the initial 30 milliseconds vector increases significantly with advancing age in men only in the frontal plane ($p=0.009$), while in the right sagittal plane the vector magnitude was observed to decrease with advancing age in both sexes ($p<0.001$ in men and women).

For Chinese, the vector magnitude was observed to increase significantly with advancing age only in women in the horizontal plane ($p=0.008$), while in the right sagittal plane the vector magnitude was found to decrease significantly with advancing age in both sexes ($p<0.001$ in men and women). In the frontal plane, no correlation with age was observed.

There were statistically significant differences between the two races ($p<0.0001$ in the frontal and horizontal planes, $p<0.01$ in the right sagittal plane).

9. Ranges of initial 20 milliseconds vector angles in frontal, horizontal, and right sagittal planes (Tables 3-18 & 3-19)

There were no significant initial 20 milliseconds QRS vector angle changes with age in either sex or race.

10. Duration and amplitude of bites in frontal, horizontal, and right sagittal planes.

Bites were found in only 153 Caucasians. The mean of the amplitude was 123±70 μV in the frontal plane, 114±86 μV in the horizontal plane, and 104±75 μV in the right sagittal plane; the means of the durations were 20±8 ms in the frontal plane, 16±7 ms in the horizontal plane, and 18±9 ms in the right sagittal plane.

Bites were found in only 102 Chinese. The mean of the amplitude was 163±89 μV in the frontal plane, 140±105 μV in the horizontal plane, and 151±103 μV in the right sagittal plane; the means of the durations were 20±8 ms in the frontal plane, 18±9 ms in the horizontal plane, and 17±8 ms in the right sagittal plane.

There were statistically significant differences of amplitude between the two races ($p<0.001$ in the frontal and right sagittal planes, $p<0.05$ in the horizontal plane), whereas, there was no statistically significant difference in duration.

3.3.2 Scalar Measurements

1. P wave amplitude (Tables 3-20 & 3-21)

For Caucasians, the P wave amplitude in lead X increases significantly with advancing age in women [$p<0.001$]. In lead Z, P wave amplitude decreases significantly with advancing age in both sexes [$p<0.001$ in both sexes].

For Chinese, the P wave amplitude in lead X increases significantly with advancing age in both sexes [$p=0.001$ in men, $p<0.001$ in women]. In lead Y, P wave amplitude increased significantly with advancing age in women [$p=0.043$]. In lead Z, P wave amplitude decreases significantly with advancing age in men [$p<0.001$].

2. P wave duration (Tables 3-22 & 3-23)

For Caucasians and Chinese, P wave duration increases significantly with advancing age in both sexes [$p<0.001$ in all leads and both sexes].

3. R wave amplitude [Fig 3-8] (Tables 3-24 & 3-25)

For Caucasians in all three orthogonal leads, R wave amplitude decreases significantly with advancing age in both sexes [$p<0.001$ in all leads and both sexes except $p<0.005$ for women in lead X].

For Chinese in lead Y, R wave amplitude decreases significantly with advancing age in both sexes [$p=0.001$ in men, $p<0.001$ in women]. In

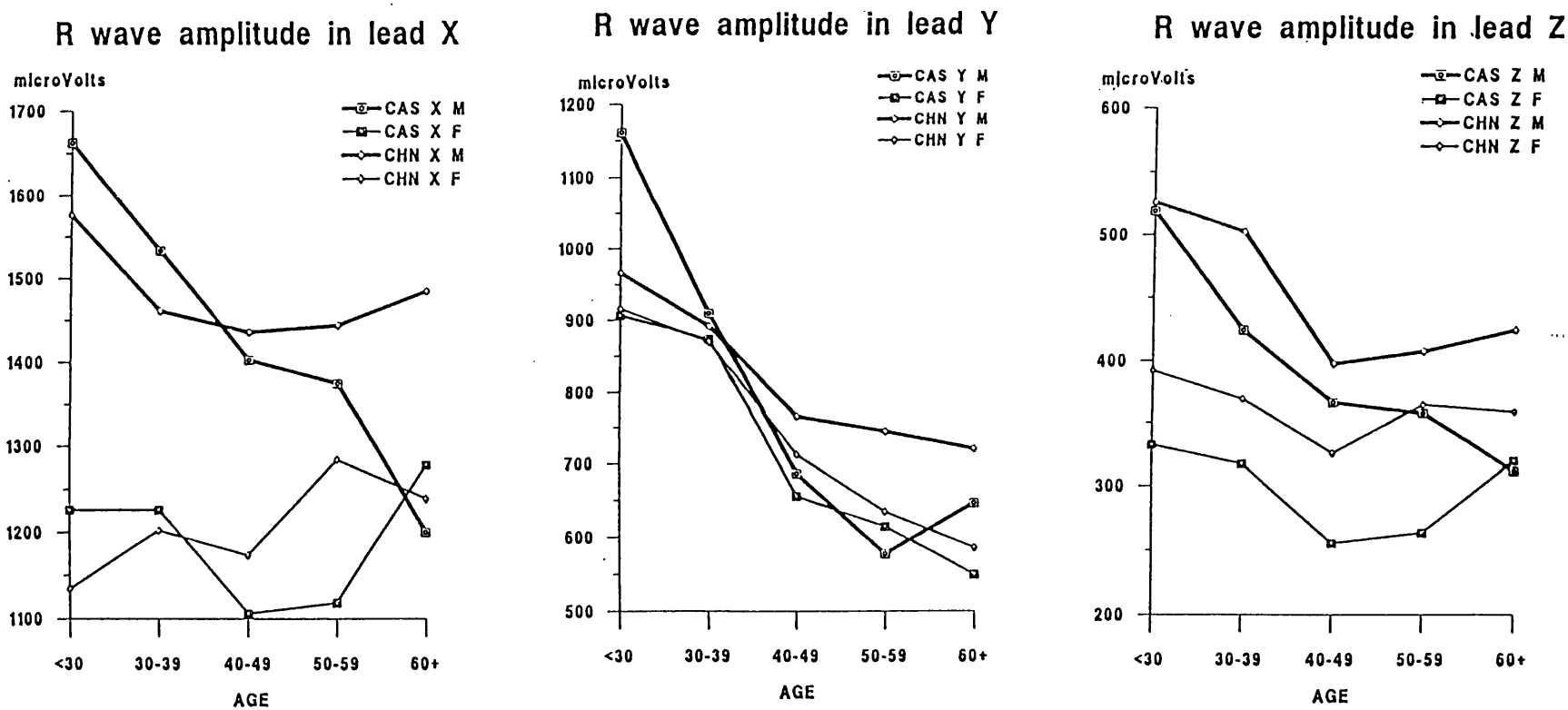


Fig.3-8 Relationship between age and mean R wave amplitude in three orthogonal leads:lead X (left panel), lead Y (middle panel), lead Z (right panel).

lead Z, R wave amplitude decreases significantly with advancing age in both sexes [$p<0.001$ in men, $p=0.019$ in women].

4. S wave amplitude [Fig 3-9] (Tables 3-26 & 3-27)

For Caucasians in lead Z, directed positively to the anterior, S wave amplitude decreases significantly with advancing age in both sexes [$p<0.001$]; while in lead X and Y, no such correlation with age was observed.

For Chinese in lead Z, S wave amplitude decreases significantly with advancing age only in men [$p<0.001$], while there were no such age dependent changes found in lead X and Y.

5. T wave amplitude [Fig 3-10] (Tables 3-28 & 3-29)

For Caucasians in all three orthogonal leads, T wave amplitude decreases significantly with advancing age in both sexes [$p<0.001$ in all leads and both sexes , except $p=0.023$ for women in lead Z].

For Chinese in all three orthogonal leads, T wave amplitude was found to decrease significantly with advancing age in both sexes [$p<0.001$ for men and women in lead X; $p=0.007$ for men and $p=0.001$ for women in lead Y; $p<0.001$ for men and $p=0.047$ for women in lead Z] .

6. Differences between two races in the amplitude of the R wave and S wave in leads X, Y, & Z.

There were significant differences in lead Z for both R wave and S wave amplitude in both sexes ($p<0.0001$), while in lead X, there were also

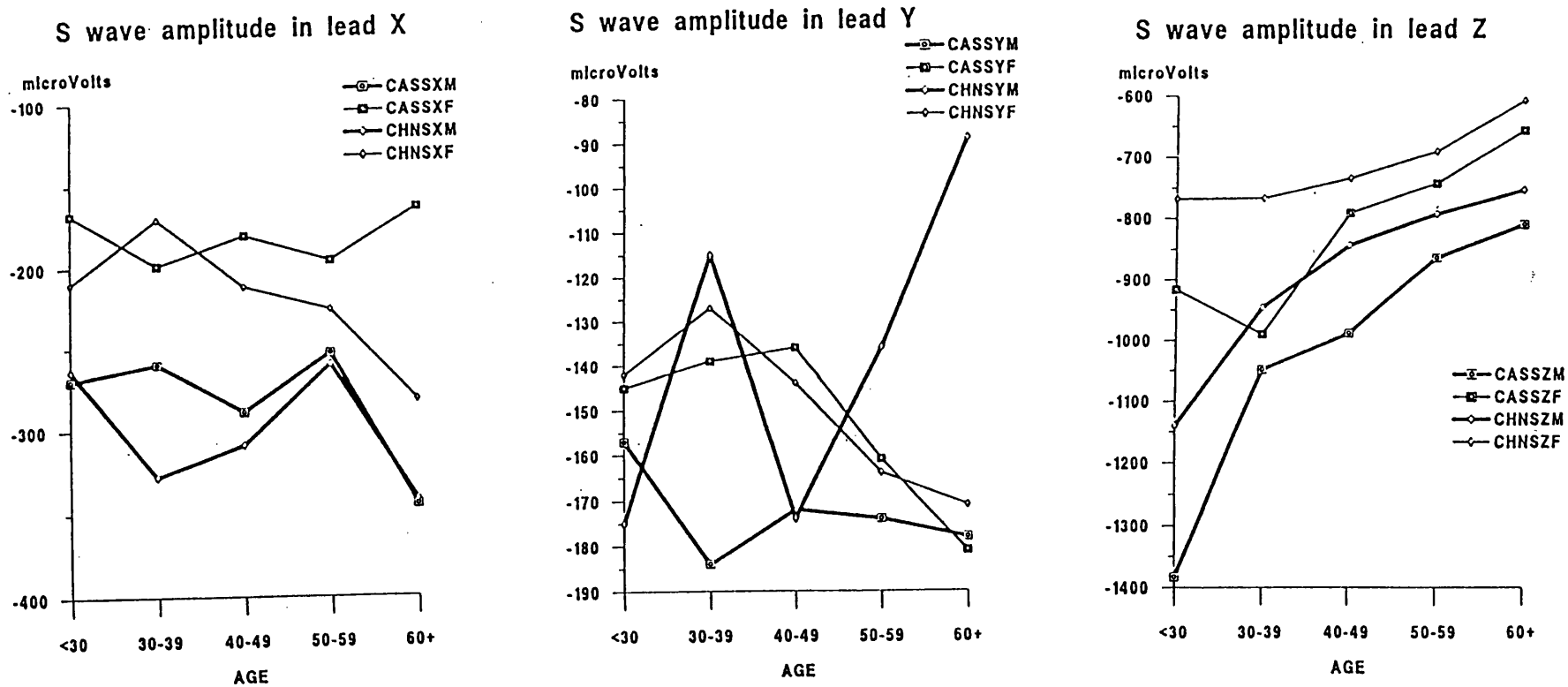


Fig.3-9 Relationship between age and mean S wave amplitude in three orthogonal leads: lead X (left panel), lead Y (middle panel), lead Z (right panel).

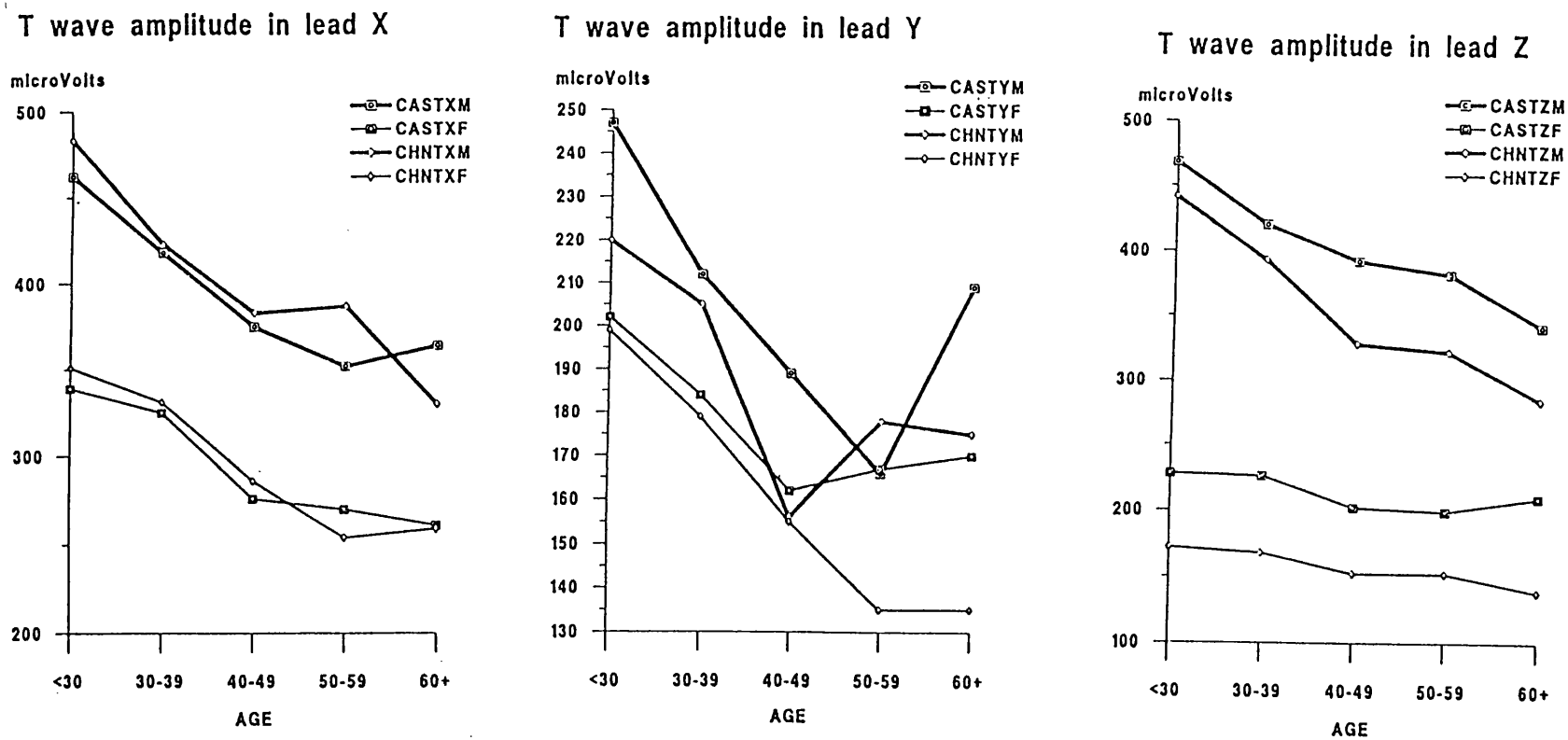


Fig.3-10 Relationship between age and mean T wave amplitude in three orthogonal leads: lead X (left panel), lead Y (middle panel), lead Z (right panel).

statistically significant differences in the S amplitude between Caucasians and Chinese in both sexes (p=0.0149).

3.4 DISCUSSION

3.4.1 General Discussion

The vectorcardiographic parameters measured included the directions of inscription of the QRS vector loops, the magnitude of initial 20 and 30 millisecond QRS vectors and the direction of the initial 20 millisecond QRS vector. These parameters were measured because of the importance of the initial portion of the QRS vector loop for the detection of myocardial infarction. Since the velocity of movement of the wave of depolarisation through the myocardium is much more variable than the direction of the initial forces, it is more difficult to quantify the effects of loss of myocardium on terminal forces (Chou, 1986).

Bites have been reported to be found more frequently in patients suffering from diabetes mellitus (Riff & Riff, 1974), adriamycin induced cardiomyopathy (Vitolo et al, 1982), and myocardial infarction (Morikawa et al, 1987) than in normal individuals. Vectorcardiographic bites have been claimed to represent localised small myocardial fibrotic lesions (Selvester, Palmersheim & Pearson, 1971). Nevertheless, vectorcardiographic bites can also be present in normal individuals without evidence of cardiovascular abnormalities (Pipberger, 1968). Therefore, an assessment of vectorcardiographic bites was also included in the present study.

Individuals with an ECG pattern of poor R wave progression over the praecordial leads were also included in the present study. It has been claimed that poor R wave progression is found in approximately 7% of patients with no disease or poor electrode placement, and is not related to patient morphology (Zema et al, 1980). Although the aetiology of poor R wave progression in normal individuals still remains uncertain and may be related to a complex interaction of variables, its presence in the normal adult population cannot be neglected and therefore 30 atypical chest pain volunteers with this pattern investigated by coronary arteriography were also included in the present study.

3.4.2 Influence of Age, Sex and Race on ECG and VCG measurements

Age, sex, and race have also been found to influence the normal electrocardiographic and Frank vectorcardiographic measurements, according to several previous reports (Draper et al, 1964; Macfarlane, Chen & Chiang, 1988; Mizuno, 1966; Nemati et al, 1978; Pipberger et al, 1967; Sotobata, Richman & Simonson, 1968). The single factor that most affects the variation in electrocardiographic measurements is age (Simonson, 1961), the influence of which is apparent from birth to death. An earlier report (Chen, Chiang & Macfarlane, 1989) also confirmed this finding using the 12 lead ECG in the presently studied group of Chinese individuals in whom body mass index contributed less than half of the variation in QRS amplitude compared to age. On the other hand, the influence of sex tends to be more important in young adulthood around puberty but has a much less important role to play at either extremes of age (Macfarlane & Lawrie, 1989). Age and sex have been found by practically all investigators to influence the normal

electrocardiographic parameters more strongly than all other factors (Simonson, 1961).

Significant differences of QRS wave and T wave amplitude as well as duration have been clearly demonstrated between American Whites and Blacks (Nemati et al, 1978). Racial variations in certain ECG parameters do exist particularly between praecordial QRS amplitudes in corresponding age and sex groups, e.g. between Chinese and Caucasians (Macfarlane, Chen, & Chiang, 1988). Since the derived VCG is synthesized from the conventional 12-lead ECG, logically these racial variations should also be present in the derived VCG as was confirmed in the present study, which has utilised the essentially identical ECG recording and analysing techniques as well as derived VCG synthesis method in both populations studied. Therefore, any differences between the two populations in the corresponding age and sex groups can only be derived from the racial variations and not from the processing procedures. Thus, this comparative study illustrates and confirms the necessity to include racial variations in ECG diagnostic criteria in order to achieve further improvement in the sensitivity and specificity of an automated ECG analysis program.

3.4.3 Direction of Inscription of the QRS Vector Loop

In our Chinese population and their Caucasian cohort, the directions of inscription of the QRS vector loop in the frontal, horizontal, and right sagittal planes (Tables 3-2 & 3-3 and Fig 3-2) were similar to those of various published studies which used the Frank lead system in Caucasians (Chou & Helm, 1967; Draper et al, 1964; Edenbrandt et al, 1987; Liebman et al, 1973; Lyon & Belletti, 1968; Macfarlane, Lorimer &

Lawrie, 1971; Nemati et al, 1978; Pipberger et al, 1967; Simonson et al, 1960; Sotobata, Richman & Simonson, 1968; Witham & Lahman, 1970).

3.4.4 Age Dependent Changes in derived Vectorcardiographic Magnitude Measurements

The most conspicuous age changes were found in magnitude measurements in various previous studies using the Frank lead system in Caucasians (Chou & Helm 1967; Edenbrandt et al, 1987; Liebman et al, 1973; Witham & Lahman, 1970). The results of the present study coincide with this finding. The age dependent magnitude change of the maximal spatial QRS vector was observed in both sexes (Table 3-5 & Fig 3-3).

For example, in men the mean value of the maximal spatial QRS vector magnitude decreases from 2.39mV (Caucasian) and 2.14mV (Chinese) in the under 30-year-old age group to 1.57mV (Caucasian) and 1.77mV (Chinese) respectively in the over 60-year-old age group, whereas in women, it decreases from 1.76mV (Caucasian) and 1.63mV (Chinese) to 1.37mV (Caucasians) and 1.48mV (Chinese) respectively in similar age groups. This trend of diminution of 34% in Caucasian versus 17% in Chinese for men and 22% in Caucasians versus 9% in Chinese for women further highlights the variation between the two races. These results are in concordance with the results of previous Caucasian Frank VCG studies (Draper et al, 1964; Macfarlane & Lawrie, 1989; Nemati et al, 1978; Pipberger et al, 1967; Sotobata, Richman & Simonson, 1968).

The magnitude of the maximal T vector was also found to decrease with advancing age in both sexes in the frontal, horizontal, and right sagittal planes. The age dependent magnitude change of the maximal T vector was more significant than that of the maximal QRS vector either in Chinese or in Caucasians.

In the planar projections of maximal QRS vectors, age dependent magnitude decreases were found to be significant in the frontal, horizontal, and right sagittal planes in Caucasians and Chinese men. This observation is similar to the studies using the Frank lead system by Draper et al (1964), Pipberger et al (1967) and Lyon et al (1968) in Caucasian men. In Chinese women this age dependent change was found to be significant only in the right sagittal plane. In the planar projections, the magnitude of the maximal QRS and T vectors was observed to decrease with advancing age in all planes for both sexes in Caucasians (Table 3-8) (Fig 3-5). The age dependent magnitude change of the maximal T vector was more significant than that of the maximal QRS vector in the horizontal and right sagittal planes, while in the frontal plane, this trend could be demonstrated only in the younger individuals aged less than forty (Fig 3-4 & 3-5).

3.4.5 Age Dependent Angular Changes of the Derived VCG

The ranges of angles of maximal QRS vectors were found to shift superiorly and anteriorly with advancing age in both sexes in Chinese, an observation which is similar to the results of previous studies using Frank VCGs in Caucasians.

3.4.6 Age Dependent Changes in Scalar Measurements

In the three orthogonal leads, the P wave duration was found to increase with age in both Chinese and Caucasians (Tables 3-20 & 3-21). This observation may be ascribed to slowed conduction through the specific intraatrial conduction systems or the atria. The slower conduction with advancing age may be due to sclerodegenerative processes affecting the conductive tissue or the atria themselves (Zoneraich & Zoneraich, 1971).

The most significant age dependent changes of the three orthogonal lead measurements were in the amplitude of the T wave. This was noted to decrease with advancing age in both Chinese and Caucasians. In lead X and Y, the amplitude of the R wave was smaller in the younger Chinese men and women than in Caucasians. Above age forty, no such relation existed. On the other hand, in lead Z, the amplitude of the R wave was greater in Chinese men and women than in Caucasians in each corresponding age group (Fig 3-8). This may imply that Chinese in general have a thinner chest wall, so that the amplitude measurements in the anteroposterior direction are significantly greater than in Caucasians.

3.4.7 Sex Differences in Derived Vectorcardiographic Magnitude Measurements

The present study also revealed that women had, on average, smaller QRS and T vector loops than men in all age groups (Table 3-8 & 3-9). There was also a statistically significant difference in magnitude in all planes in each age group. According to an autopsy study of the normal human heart by Kitzman et al (1988), body mass and body surface area (BSA) in adults of both sexes are better univariate predictors of normal

heart mass than is body height. When matched for age and BSA, however, normal heart weight is greater in men than in women and the correlations between heart weight and BSA were better in men than in women. In contrast to the well established correlation between BSA and heart mass, the relationship between age and heart mass in adults is somewhat controversial. The present study has demonstrated significantly larger QRS and T vector magnitudes in normal males than in normal females, which can be implied from the fact that the normal male has on average a greater cardiac mass and hence, cardiac electromotive forces than the normal female. The maximal QRS and T vectors showed a statistically significant difference in magnitude in the frontal, horizontal, and right sagittal planes in all age groups. There was also a statistically significant sex difference in the magnitude of maximal spatial QRS vector in all age groups, although in a previous study of Caucasians and Japanese by Sotobata et al (1968), it was found that the right sagittal maximal QRS and the frontal maximal T vectors failed to show significant sex differences in their magnitude. These variations may be ascribed to the racial difference and the small number of subjects in their study [102 women and 101 men]. The smaller QRS and T vector magnitude in women have been ascribed to differences in torso size, higher content of body fat, and smaller average heart size (Nemati et al, 1978; Sotobata, Richman & Simonson, 1968).

3.4.8 Sex Dependent Angular Changes of the Derived Vectorcardiogram

The maximal T vector angle is one of the parameters in which the most prominent sex differences were observed. In women, it is, in general, directed more posteriorly or more inferiorly in each plane than in men.

The directional difference in the horizontal plane is ascribed to the fact that the anteroposterior component has a larger sex difference than the leftward component because of constitutional factors (Nemati et al, 1978; Sotobata, Richman & Simonson, 1968). The present study reconfirmed this finding.

The observation of a more posterior orientation of the maximal T vector in women in the horizontal plane is comparable to findings by Simonson et al (1961) and Nemati et al (1978) in Caucasians. This observation may be inferred from the electrocardiographic observation of more marked sex differences in lead V2 and V3 than in the left praecordial leads with a negative mean T amplitude value for women and a positive mean value for men in lead V1 (Macfarlane, Chen & Chiang, 1988; Nemati, Doyle & McCaughan, 1978; Simonson et al, 1960). This is also consistent with the findings of Mizuno (1966) that in healthy Japanese women, there was a higher incidence of negative T waves in leads V1 through V3 than in men. This phenomenon was also found to exist in Chinese women. In 10% of Chinese women between 40-49 years, there was a negative T wave component in lead V2 and 6% had a similar finding in V3. These percentages are much higher than in apparently healthy Caucasian women who rarely have an inverted T wave in V2 and almost never have an inverted T wave in V3 (Macfarlane, Chen & Chiang, 1988).

3.4.9 Racial Differences in Derived Vectorcardiographic Magnitude Measurements

In a previous study of twelve-lead electrocardiograms in Chinese (Chen, Chiang & Macfarlane 1989, Macfarlane, Chen & Chiang, 1988), the

amplitude of the QRS complex was generally smaller in Chinese than in Caucasians. This implies that the magnitude of the spatial maximal QRS vector in Chinese should be smaller than that in Caucasians.

In the present study, the maximum QRS spatial vector had a mean magnitude in Chinese of 1.89±0.53 mV in men and 1.55±0.38 mV in women while in Caucasians it was 2.39±0.63 mV in men and 1.76±0.47 mV in women. This compares with 1.58±0.43 mV for white males and 1.87±0.45 mV for black males (Pipberger et al, 1967), as well as 1.35±0.36 mV for a mixed white and black female cohort using the Frank VCG (Nemati et al, 1978). These data present a confusing picture from small numbers of individuals in the various age groups. Clearly, maximum values will come from the younger age group but the number of younger blacks in the study of Pipberger et al (1967) was not stated. In any event, derived vectorcardiographic measurements may differ from actual Frank vectorcardiographic measurements in the same subject. It might be concluded, however, that derived VCG amplitudes are higher than corresponding Frank VCG measurements.

There were also statistically significant differences of the bite amplitude between the Caucasians and Chinese, whereas there was no difference in bite duration between the two races.

3.4.10 Summary of Similarities in the Derived Vectorcardiographic Measurements in the two Races

The similarities of the trends in various derived VCG parameters between Caucasians and Chinese can be summarised as follows:

Similarities:

(1) The dominant directions of inscription of the vector loops were the same in both races (i.e. frontal plane:clockwise, horizontal plane:counterclockwise, right sagittal plane: counterclockwise);
(2) The magnitude of the maximal spatial QRS and T vectors in men are greater than in women;
(3) The magnitudes of the spatial QRS and T vectors decrease with advancing age;
(4) The maximal QRS and T vector angles are directed more superiorly and anteriorly in the older age groups;
(5) The maximal QRS and T vector angle in men are directed more superiorly and anteriorly than those in women;
(6) The magnitude of the initial 20 milliseconds vector increases with advancing age in the frontal plane while it decreases in the right sagittal plane;
(7) There is no significant age dependent change of the initial 20 milliseconds vector angle.

3.4.11 Summary of Differences in the Derived Vectorcardiographic Measurements Between the two Races

The differences in the various derived VCG parameters between the two races can be summarised as follows:

Differences:

(1) There is no clockwise or figure of 8 inscription of the QRS vector loop for Chinese men in the horizontal plane;
(2) The magnitude of maximal spatial QRS and T vectors in Chinese is smaller than in Caucasians;
(3) The magnitude of the maximal spatial QRS vector is greater in Caucasians than in Chinese in the groups aged less than forty, whereas in the groups aged above forty it is greater in Chinese than in Caucasians. These phenomena also exist in the frontal and horizontal planar projections of the maximal QRS vectors. In the right sagittal plane, there is no such relationship;

(4) The age dependent diminution of the maximal QRS magnitude is only significant in the right sagittal plane for Chinese women but is significant in all planes for Caucasian women;
(5) The magnitude of the initial 20 milliseconds vector decreases significantly with advancing age for Caucasians but not for Chinese in the horizontal plane.

These discrepencies between the two races may be related to environmental, occupational, or dietary variations. In a previous 12-lead ECG study (Macfarlane, Chen & Chiang, 1988), the body mass index was found to be lower in the Chinese individuals compared to the Caucasians in the corresponding age and sex groups. It might be inferred that the difference in the magnitude of the QRS and T vector is due to the difference in the thoracic muscle mass and the fat content, but this cannot explain the discrepancy for the groups aged forty or more. This mandates further investigation and is beyond the scope of the present study.

Although the trends of the derived VCG parameters are in concordance with the previous observations using the Frank lead system in studies on Caucasians and Japanese (Draper et al, 1964; Mizuno, 1966; Nemati et al, 1978; Pipberger et al, 1967, Sotobata, Richman & Simonson, 1968), nevertheless, the quantitative derived VCG parameters and the derived VCG vector loop configuration are different from those of the Frank VCG measurements (Edenbrandt & Pahlm, 1988; Kors et al, 1990b; Rubel, Benhadid & Fayn, 1992). The effects of age and sex on the magnitude of the vector for both races are well demonstrated in the present study. Thus, further investigations to assess racially dependent derived VCG diagnostic criteria based on these normal limits are necessary. Various cardiac conditions, e.g. myocardial infarction,

ventricular hypertrophy, and bundle branch block, required to be included in any such study.

3.5 CONCLUSION

The data from the present study illustrate significant age, sex and racial variations in the derived VCG. These are of potential importance for diagnostic applications. The specificity of vectorcardiogram interpretation can be further enhanced by the utilisation of age, sex, and race-corrected normal ranges. The main use of these results is likely to lie in obtaining diagnostic vectorcardiographic criteria which are supplementary to those of the conventional scalar 12-lead ECG. Macfarlane and Edenbrandt (1992), in a short discussion article, showed how the derived vectorcardiogram could help to distinguish between different abnormalities when the diagnosis from the scalar 12-lead ECG alone was not clear. For example, in patients with non-specific poor R wave progression in the 12-lead ECG, some may have derived vectorcardiographic loops which appear normal, while others may have loops in which bites can be detected, suggesting that the low R waves are due to myocardial infarction. It is hoped that the present work will in future allow highly specific vectorcardiographic criteria to be derived from the 12-lead ECG for incorporation into an automated program for the analysis of the resting 12-lead ECG.

The significance of these observations is that they provide further evidence that race is an important factor that should be taken into consideration for an automated ECG interpretation program. In conclusion, it is hoped to incorporate age, race and sex dependent

derived VCG diagnostic criteria into the Glasgow Program in the future in order to improve its sensitivity and specificity.

CHAPTER 4

SOFTWARE BASED ARTIFICIAL NEURAL NETWORKS FOR THE DIAGNOSIS OF ATRIAL FIBRILLATION

4.1 INTRODUCTION

Atrial fibrillation (AF) is a common atrial arrhythmia, with a prevalence of approximately 2% in the general population (Dunn et al, 1989). It appears more commonly in the elderly, affecting about 5% of persons over 60 years of age, and its incidence increases with advancing age (Kannel et al, 1982). The importance of atrial fibrillation is emphasized by its association with between 6% and 24% of ischaemic strokes and with about 50% of cardioembolic strokes (Cerebral Embolism Task Force, 1986 & 1989). The definitive diagnosis of AF is usually made by electrocardiograms (ECGs) showing an absence of discrete atrial activity, an irregular undulating baseline, and a variable RR interval.

Before the introduction of computers for electrocardiogram interpretation of cardiac rhythm, each ECG had to undergo visual classification by experienced cardiologists. Since the ECG is one the most available and most cost-effective, non-invasive investigational tools for the diagnosis of cardiac rhythm in modern daily cardiological practice, demand has increased in recent years, resulting in additional work for cardiologists. On the other hand, the interpretation of cardiac rhythm from an ECG needs experience and expertise. For these reasons, it was hoped that the application of computers to ECG analysis could not

only save time but perhaps also improve the accuracy of the interpretation of cardiac rhythm. As far as is known, all programs for computer-assisted ECG rhythm interpretation use deterministic logic.

Although computer-assisted ECG interpretation was introduced in the early 1960s, rhythm analysis is still one of the problematic areas requiring further research. In Glasgow Royal Infirmary, three orthogonal leads X, Y, Z were initially used for the development of methods for the interpretation of cardiac rhythm (Taylor, 1973). Subsequently, the selection of three leads from the simultaneously recorded 12 conventional leads was utilised for rhythm analysis (Macfarlane, 1986b; Macfarlane et al, 1986a). Up until now, a diagnostic tree or a set of diagnostic criteria with sharp thresholds for classification have been used in essentially all computer-assisted rhythm interpretation programs (e.g. Macfarlane et al, 1990b). This deterministic logic might function satisfactorily if the measurements related to P waves or fibrillatory wave detection were accurate and reliable. However, isolated P waves can sometimes be missed and fibrillatory waves at higher heart rates can often not be detected by the computer or even by human eyes [Fig 4-1a].

Deterministic logic is currently applied to rhythm analysis in computer-assisted electrocardiographic interpretation methods developed in the department of Medical Cardiology of Glasgow Royal Infirmary (Macfarlane et al, 1990b). Recently, a locally developed artificial neural network simulation software package (Devine, 1990) has become available, and has been applied to the classification of electrocardiographic ST-T segments (Edenbrandt, Devine & Macfarlane, 1992) and detection of electrocardiographic left ventricular strain (Devine & Macfarlane, 1993). On the other hand, because artificial neural

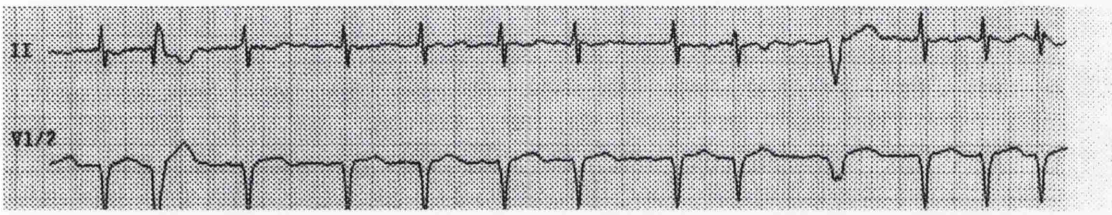


Fig.4-1a An ECG rhythm strip of atrial fibrillation with no identifiable P waves.

networks are not rule-based and can be applied in areas where rules are difficult to formalise or are ill-defined, they are theoretically particularly suitable for application to visual pattern recognition problems in computerised electrocardiography, such as AF. Therefore, it was decided to investigate whether the application of software based artificial neural networks might offer any advantage as compared to the existing deterministic logic for separating the AF from sinus rhythm and (supraventricular extrasystoles &/or ventricular extrasystoles) [Fig 4-1b].

4.2 METHODS OF RHYTHM ANALYSIS IN THE GLASGOW PROGRAM

At present, the rhythm interpretation logic is linked to the main wave measurement section of the computer-assisted ECG diagnostic Glasgow program (Macfarlane et al, 1990b). The QRS onsets and terminations used in wave typing are transferred to the rhythm program together with the measurement matrix of the 12-lead ECG and used as landmarks around which P wave recognition can be performed. Furthermore, the wave typing and pacemaker spike recognition logic which are in the wave measurement program also have their results transferred to the rhythm analysis section inside the Glasgow Program (Macfarlane, 1986b).

In the Glasgow Program, an 8 or 10 second strip of three simultaneously recorded leads chosen from the conventional 12-lead ECG is used for the analysis of cardiac rhythm (Macfarlane et al, 1986a; Macfarlane et al, 1990b). These 3 leads are selected on the basis of the study of P wave amplitudes of the average beats calculated by the main wave measurement program. Leads II and V1 are always chosen and a third

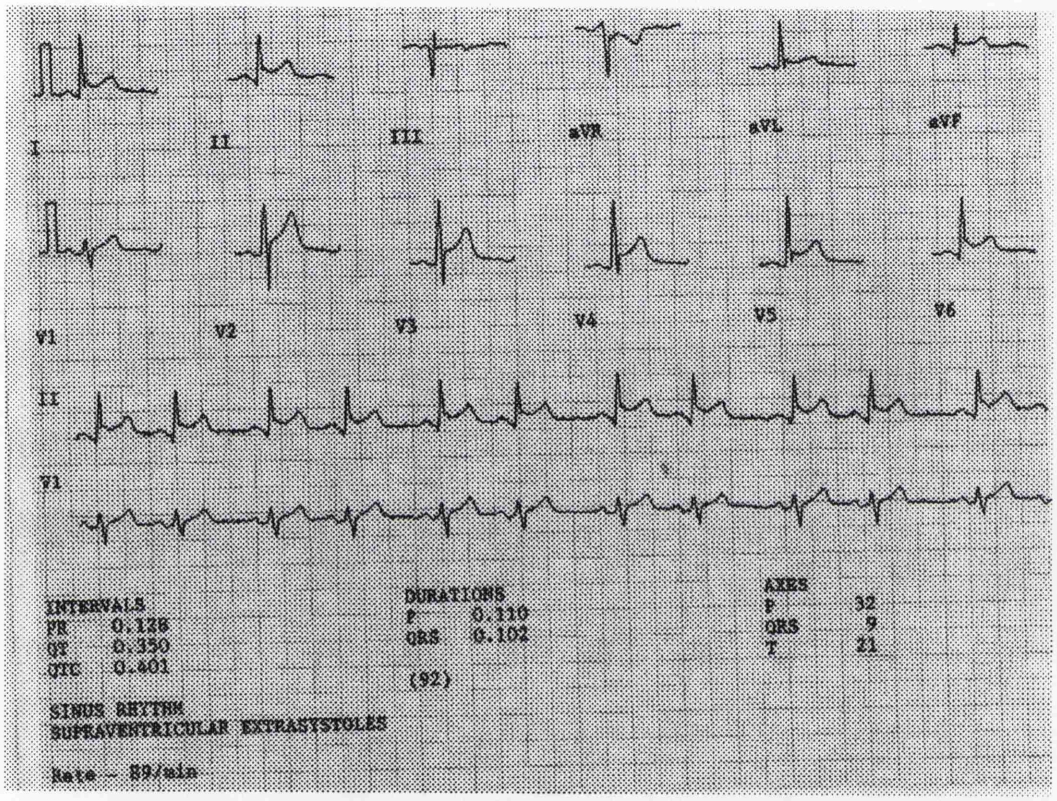


Fig.4-1b An ECG showing sinus rhythm with supraventricular extrasystoles.

lead is selected from I, III, aVF, and aVR, depending on the P wave amplitude. In general, this applies in the presence of an expected sinus rhythm. If atrial flutter has been detected in lead II, then leads III and V1 are the other two leads which would automatically be selected. On the other hand, if atrial flutter is not present as is the situation with over 99% of daily routine ECGs, the program will choose leads II, V1 and one other lead from I, III, aVF and aVR.

Irrespective of which three leads are selected, the configuration of the P wave in the corresponding average beat is studied. For each P wave configuration, a special category is assigned. This is subsequently used in P wave recognition which is undertaken by searching the first difference of the lead in an attempt to find a series of gradients which would conform to the selected P wave configuration. In other words, because the first difference of the P wave in lead II is normally different from that in V1, cognizance of this must be taken in any search for P waves. In the Glasgow Program, the sampling rate is 500 samples/second and the first difference used in the P wave search utilises a difference interval of 10 samples to enhance the amplitude of the signal. If no definite P wave is found in the average beat, such as would occur frequently in atrial fibrillation or other dysrhythmias, for instance complete atrioventricular block, the leads selected for analysis are II and V1, with two different P wave morphologies being adopted for the latter.

From the three selected leads, measurements on

(a) PR interval variability= {(minimum PR/maximum PR) x 100%},
(b) RR interval variability = % RR intervals which differ by more than 10% from the mean RR interval,
(c) Percentage regularity of dominant RR intervals
= 100% - {[(maximum RR - minimum RR) / mean RR] x 100%},

(d) Presence or absence of discrete P waves,
(e) Presence or absence of multiple P waves.
and so on (Taylor, 1973) are determined.

In the presence of atrial fibrillation, it is likely that fibrillatory "f" waves may, for the purpose of rhythm analysis, be detected as "P" waves and hence the concepts of multiple "P" waves and PR regularity etc can still be used in the Glasgow ECG interpretation program (Macfarlane et al, 1986a; Macfarlane et al, 1990b).

The Glasgow rhythm interpretation program is divided into six phases:

PHASE 1

The first phase focuses on the analysis of P waves given that the various QRS onsets and terminations are available as input data from the main interpretation program. The purpose of this phase is to filter out the vast majority of ECGs which show pure regular sinus rhythm. This is achieved on the basis of the obvious criteria of finding one P wave for each individual RR interval and a regular PR interval.

The program searches for P waves throughout most of the RR interval. In the course of assessing the presence or absence of sinus rhythm, a check is made for "flutter" waves and various sections of logic are used to detect "multiple P waves", "few P waves", "irregular PR interval" etc. If sinus rhythm is detected but the rate exceeds 100 or is lower than 60, further searching will be made to ensure that there are no other complicating factors. If even one RR interval has no P wave found, perhaps in error, further assessment is made in the next phase. If regular

sinus rhythm is found, the remaining sections of the rhythm analysis are not entered.

PHASE 2

In phase 2, an attempt is made to sort AF from any other irregular rhythm which might be present, e.g. SR + (SVEs &/or VEs). There is one other important additional test in which the RR intervals are searched for the presence of a definite P wave. This could be such as occurs in isolation in complete atrioventricular block, although it may only be found in every other RR interval. The average beat calculated in the main program is also used in searching for definite P waves.

However, the two approaches must be coupled together since various rhythms such as junctional rhythm and atrial fibrillation will produce an average beat with virtually no identifiable P wave as would also be the case in complete atrioventricular block, even although P waves will be present in the record.

PHASE 3

Phase 3 contains two mutually exclusive sections of logic, one which deals with regular rhythm and the other with irregular rhythm. The former handles rhythms such as atrial flutter with a regular ventricular response, or complete atrioventricular dissociation. The latter deals with abnormalities such as second degree sino-atrial or atrio-ventricular block or atrial bigeminy.

PHASE 4

Phase 4 is entered only if aberrant beats have been detected and differentiates between ventricular ectopics and supraventricular ectopics with aberrancy, non-sustained ventricular tachycardia etc.

PHASE 5

In phase 5, the discrimination between sinus arrhythmia and sinus rhythm with supraventricular extrasystoles is performed.

PHASE 6

The final stage of the rhythm analysis is to assess pacemaker spikes detected by the main program particularly when these do not occur regularly. The program is able to produce diagnostic statements which deal with regular or demand atrioventricular sequential or ventricular pacing as appropriate.

This final section of the analysis also decides whether a preliminary assessment of sinus rhythm can be retained or whether it should be changed to ectopic atrial rhythm given the presence of inverted P waves in the inferior leads II, III, and aVF. The list of possible diagnoses consists of two groups, one of which contains the dominant rhythm while the other includes various supplementary statements. Only one dominant rhythm, e.g. AF can be selected together with up to three supplementary statements, e.g. ventricular extrasystoles.

4.3 USE OF NEURAL NETWORKS IN GLASGOW RHYTHM ANALYSIS

The neural network simulation package was written in C++ language and developed locally on a DEC Microvax II minicomputer. The software was tested and debugged using a number of methods described elsewhere (Devine, 1990). A supervised feed-forward type neural network trained with a back-propagation algorithm was used in the present study. The training tolerance was set at 0.1, and the gain was set at 1.0.

It was decided to assess the use of an artificial neural network only in those cases which had passed phase 1 of the analysis, entered phase 2 and reached the point where existing logic tries to separate irregular sinus rhythm from atrial fibrillation. It was estimated that the number of ways through the logic at that point was over 20,000.

The number of neurons in the input layer and the hidden layers, as well as the number of hidden layers, were determined by trial and error during the initial development of the neural networks. The training process consisted of repeated sequential presentation of the measurements from training set ECGs to a variety of neural networks until the error between the output and the designated pattern could not be improved (i.e. the error could not be reduced as it had reached a minimum). In most of the circumstances for this study, the networks were fully trained after more than one hundred thousand iterations, i.e. all the ECGs in the training set were well separated into AF or Non-AF by a set of completely adjusted weights.

4.4 METHODOLOGY

3080 ECGs with irregular rhythm, consisting of atrial fibrillation, or sinus rhythm (SR) with supraventricular extrasystoles (SVEs) &/or ventricular extrasystoles (VEs), were selected from the Departmental data bank for use in the present study. These ECGs were recorded from the patients attending the outpatient clinics or admitted to Cardiology wards of the Royal Infirmary from 1981 to 1992.

Gold Standard Used in the Present Study

All the interpretations of ECGs were initially made by the deterministic logic in the Glasgow Program while ECGs were also reviewed by two experienced electrocardiographers in order to reach consensus for every rhythm diagnosis. The "gold standard" used in the current study was the electrocardiographers' diagnosis. i.e. all the ECG rhythm diagnoses were determined by the visual classification of two experienced electrocardiographers.

The ECGs were subsequently randomly divided into five sets (Table 4-1), each set containing both AF and Non-AF. The Non-AF group consisted of [SR + (SVEs &/or VEs)]. Regular rhythm such as normal sinus rhythm without any SVEs &/or VEs was not used in the present study.

The distribution of the ECGs in the five sets (A, B, C, D, & E) is shown in Table 4-1.

The way in which the various sets of ECGs are used in the present study is summarised in the flow diagrams (Figures 4-2a & 4-2b).

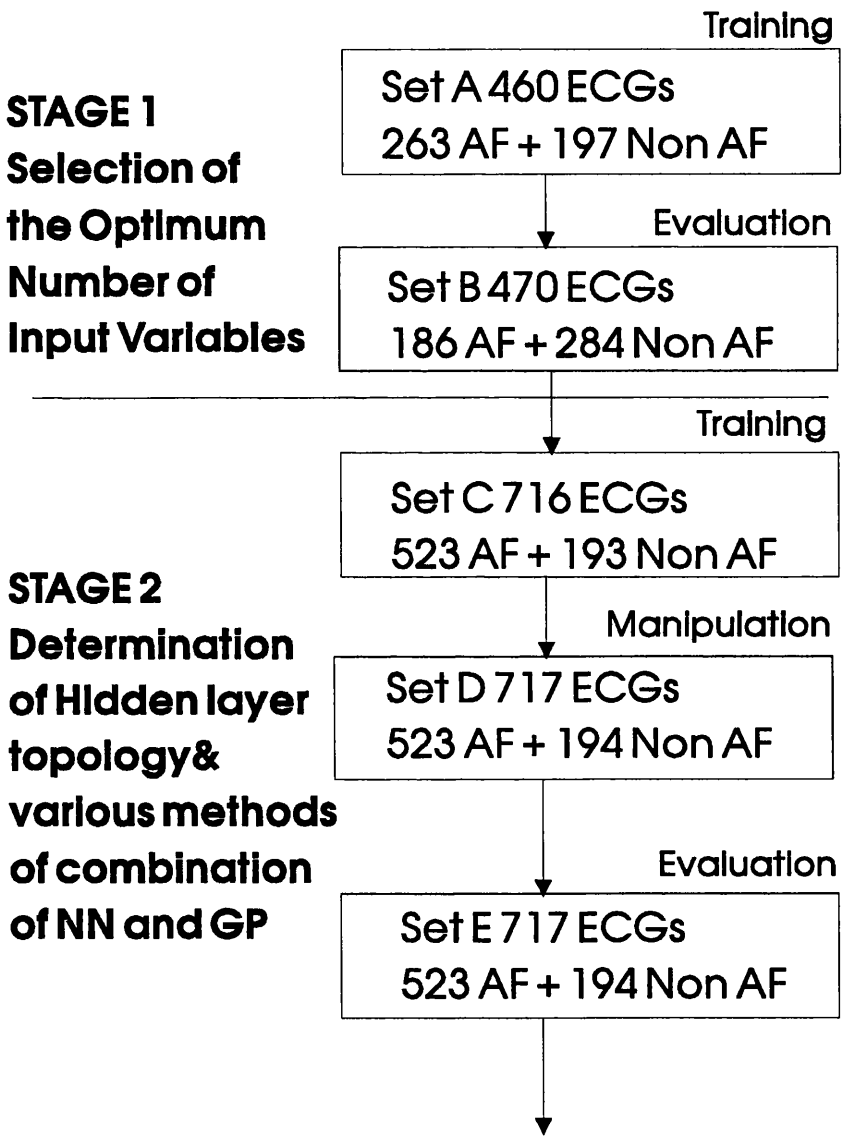


Fig.4-2a Flow Chart of Stages 1 & 2 in the Evaluation of Neural Networks for the Diagnosis of Atrial Fibrillation

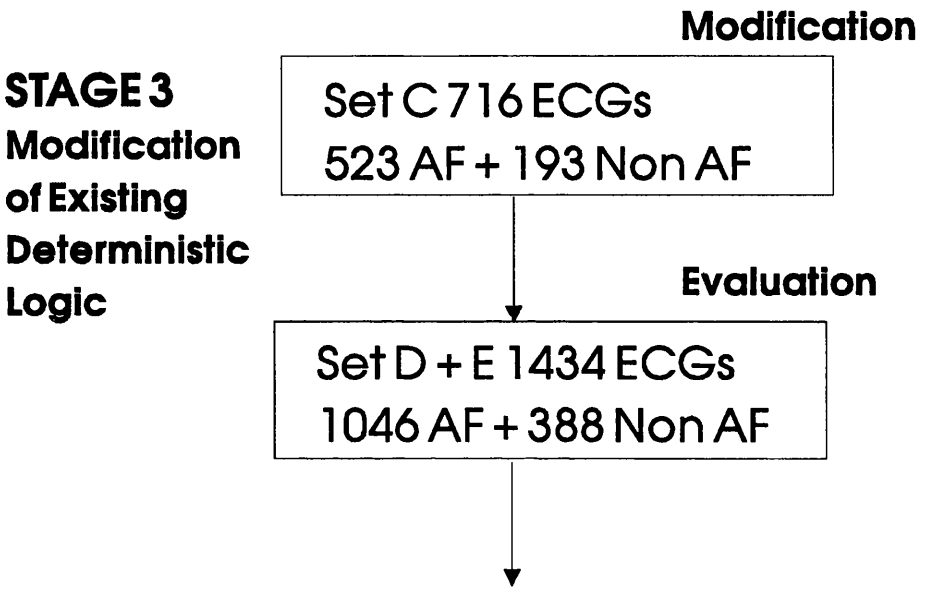


Fig. 4-2b Distribution of ECGs used in Stage3 of the Modification of Existing Deterministic Logic for the Diagnosis of Atrial Fibrillation

ECG Set	AF	Non-AF	Total
A	263	197	460
B	186	284	470
C	523	193	716
D	523	194	717
E	523	194	717
Total	2018	1062	3080

Table 4-1 Composition of the ECG sets used for the neural network diagnosis of atrial fibrillation.

4.4.1 STAGE 1

Selection of the Optimum Input Variables for the Artificial Neural Networks

4.4.1.1 MATERIALS

930 ECGs were divided into set A (460 ECGs) for the purpose of training and set B (470 ECGs) for the assessment of the performance of neural networks in order to determine the optimal number of input variables for separating AF and Non-AF.

4.4.1.2 METHODS

The neural networks were used in isolation for the diagnosis of atrial fibrillation. The input variables to the neural networks were selected from the existing rhythm analysis parameters used by the Glasgow Program, namely:

[1] RR interval variability;
[2] PR interval variability;
[3] Number of leads with definite P waves;
[4] Presence or absence of multiple P waves;
[5] Number of limb leads with definite P waves;
[6] Number of praecordial leads with definite P waves;
[7] Percentage regularity of dominant RR intervals;
[8] Maximum PR interval;
[9] Minimum PR interval;
[10] Presence of regular dominant rhythm;
[11] Number of samples in mean RR interval.

Only ECG parameters of importance for the separation of atrial fibrillation from SR+(SVEs &/or VEs) were chosen as the input variables to the network. Three groups of neural networks with 7, 9, and 10 input variables were used in order to determine the best topology of the input layer. The parameters selected were 1 to 7, 1 to 9, 1 to 10, respectively, from the above list. All the networks were designed with only a single output neuron. The number of hidden layers varied from one to two, and the number of neurons in the hidden layer(s) varied from 5 to 40.

The structure of the artificial neural network used for the diagnosis of atrial fibrillation in the present study is shown in Fig 4-3.

Training and Assessment Procedures

The ECGs of set A together with the corresponding rhythm coding (AF was assigned "1" and Non-AF was assigned "0") were fed into the networks for the purpose of training. Usually, it required around 100,000 iterations to have a single network reach a fully trained status. Then the ECGs of set B were run on these various well trained networks

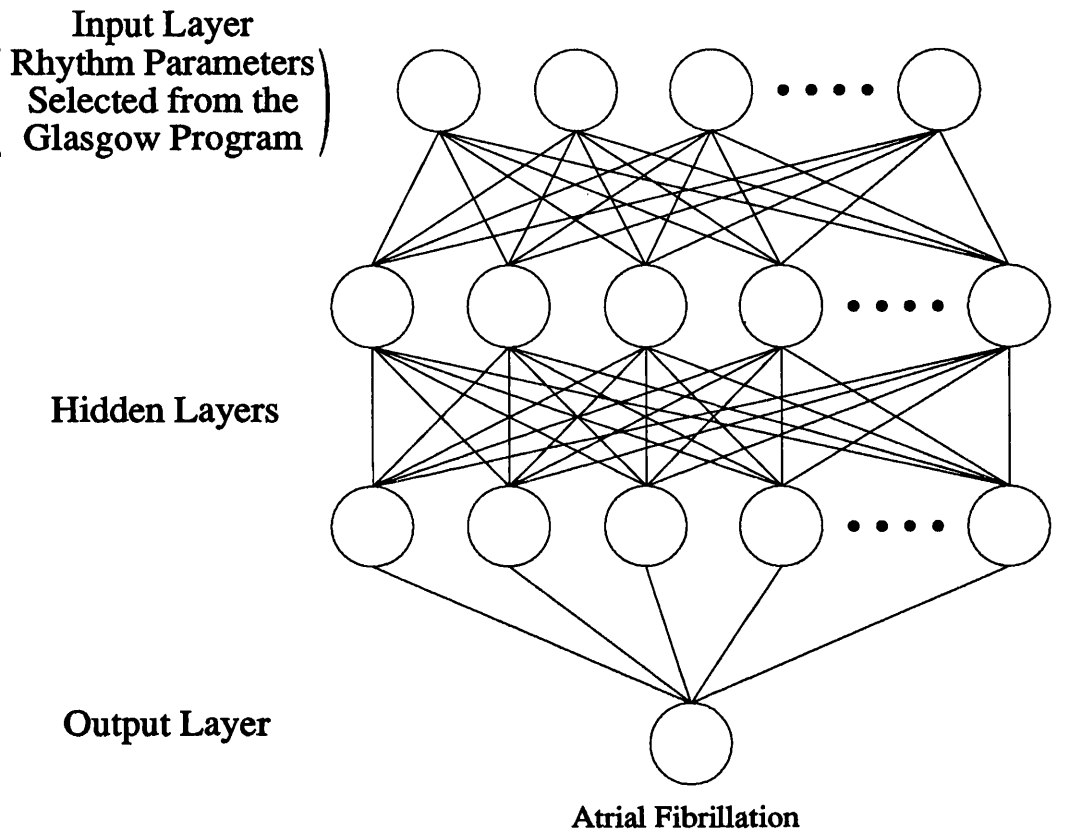


Fig 4-3. Structure of the Neural Network Used in the Diagnosis of Atrial Fibrillation.

to evaluate their performances. The output from each neural network was a numerical value ranging between 0.0 (Non-AF) and 1.0 (AF). The cutoff point 0.5 was arbitrarily selected to differentiate between AF and Non-AF, because the numerical output from the network for AF was found to cluster in the range 0.5 to 1 and for non-AF in the range 0 to 0.5 (see Fig. 4-4). Thereafter, the selection of the optimum number of input variables for the best performing artificial neural network was determined by reviewing the Association Index and specificity based on the assessment of set B ECGs using this discriminating point with value 0.5.

Sensitivity, Specificity and Association Index

In the present study, sensitivity (Se), specificity (Sp), and Association Index (AI) (Rautaharju, Blackburn & Warren, 1976) were defined and calculated as follows:

$$\text{Sensitivity} = \frac{\text{No of True AF detected by the specific method}}{\text{Total no of True AF in the sample}} \times 100\%$$

$$\text{Specificity} = \frac{\text{No of True Non-AF detected by the specific method}}{\text{Total no of True Non-AF in the sample}} \times 100\%$$

$$\text{Association Index} = \text{Sensitivity} + \text{Specificity} - 100\%$$

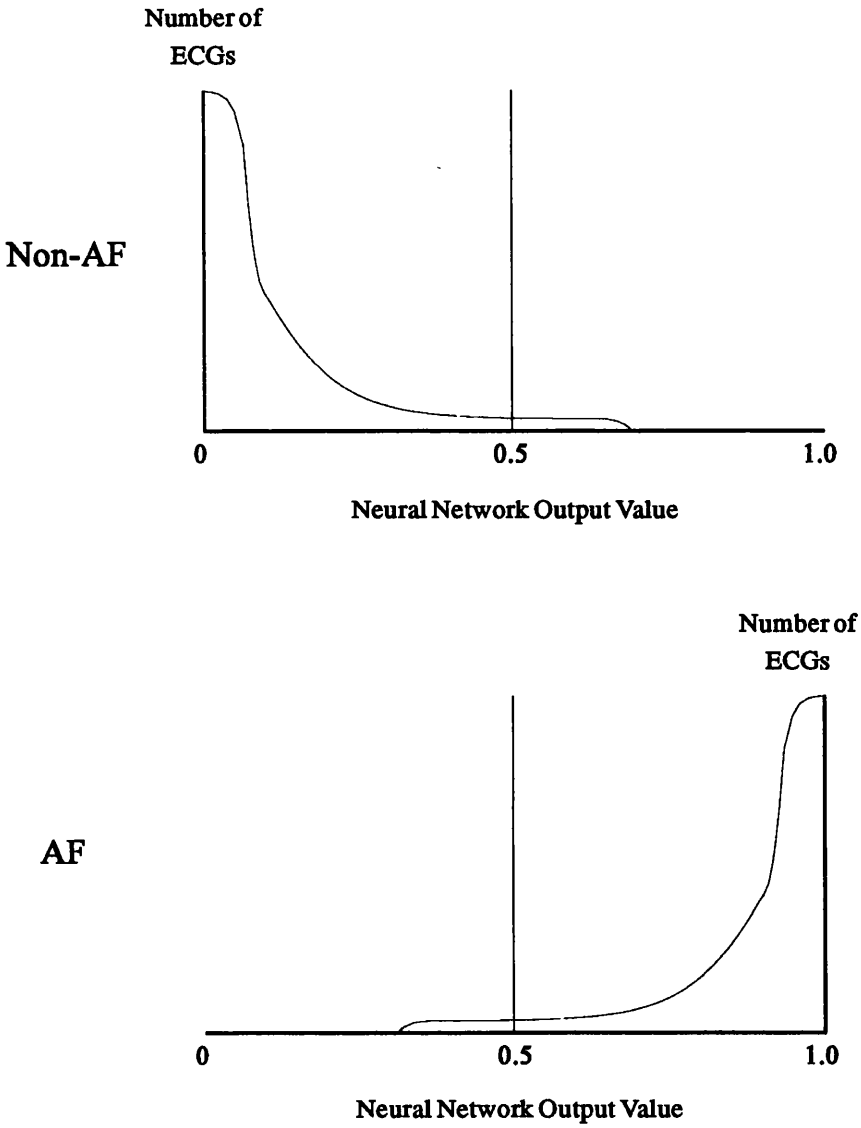


Fig. 4-4 Schematic illustration of the distribution of the typical outputs from a Neural Network used to separate AF from non-AF cases.

4.4.1.3 RESULTS

The results from the ECGs of set B assessed by the neural networks with 7, 10 and 9 input variables but all with a single output are shown in Tables 4-2, 4-3, 4-4, respectively.

From the assessment of ECGs of set B, the artificial neural networks with 9 input variables were found to have the best performance as compared to the networks with 7 and 10 input variables (Tables 4-2, 4-3 & 4-4). Therefore, it was decided to use the neural networks with 9 input parameters for the diagnosis of atrial fibrillation in further studies. The 9 ECG measurements input to the optimal performing neural network for the diagnosis of atrial fibrillation were as follows:

(1) RR interval variability;
(2) PR interval variability;
(3) Number of praecordial leads with definite P waves;
(4) Number of limb leads with definite P waves;
(5) Presence or absence of multiple P waves;
(6) Percentage regularity of dominant RR intervals;
(7) Number of leads with definite P waves;
(8) Maximum PR interval;
(9) Minimum PR interval.

The optimum performing neural network was selected from Table 4-4 on the basis of the Association Index and the specificity. Although the neural network with a single hidden layer containing 30 neurons had the best Association Index as compared to the double hidden layer network (92.0% versus 91.7%), its specificity (94.7%) was lower than the other network with 30-30 configuration in the hidden layer (94.7% versus 96.5%). As a consequence, the artificial neural network with the optimal

performance was selected to be that with 2 hidden layers each with 30 neurons.

Number of neurons in hidden layers	Sensitivity %	Specificity %	Association Index %
20	86.0	98.3	84.3
30	89.2	96.5	85.7
40	88.7	97.6	86.3
20-20	82.3	98.3	80.6
30-30	88.2	97.9	86.1
30-30 v2*	90.3	98.3	88.6
40-40	86.6	98.6	85.2

Table 4-2 The effect of varying design of hidden layers on the neural network with 7 input variables assessed by ECGs of set B.*v2 denotes a network with different initial weights compared to another of identical design.

Number of neurons in hidden layers	Sensitivity %	Specificity %	Association Index %
20	86.6	98.3	84.9
30	87.1	95.8	82.9
40	85.5	95.8	81.3
20-20	87.6	98.6	86.2
30-30	84.9	97.6	82.5
30-30 v2	83.9	99.3	83.2
40-40	87.1	98.3	85.4

Table 4-3 The effect of varying design of hidden layers on the neural network with 10 input variables assessed by ECGs of set B.*v2 denotes a network with different initial weights compared to another of identical design.

Number of neurons in hidden layers	Sensitivity %	Specificity %	Association Index %
5	96.8	94.4	91.2
10	97.3	94.0	91.3
15	93.5	94.4	87.9
20	96.2	94.4	90.6
30	97.3	94.7	92.0
40	97.3	91.9	89.2
5-5	96.2	94.7	90.9
10-10	94.6	96.1	90.7
15-15	95.2	94.7	89.9
20-20	94.6	94.0	88.6
20-20v2	93.0	95.8	88.8
30-30	95.2	96.5	91.7
30-30v2	95.2	96.1	91.3
40-40	93.0	95.4	88.4

Table 4-4 The effect of varying design of hidden layers on the neural network with 9 input variables assessed by ECGs of set B. *v2 denotes a network with different initial weights compared to another of identical design.

The effect of varying the design of the hidden layer(s) on the network with 9 input variables is shown in Table 4-4. For the original Glasgow program, sensitivity was 94.6% (176/186) and specificity was 98.6% (280/284), while for the optimum performing artificial neural network, sensitivity was 95.2% (177/186) and specificity was 96.5% (274/284).

4.4.2 STAGE 2a

Determination of the Optimal Topology of the Hidden Layer in the Neural Network

4.4.2.1 MATERIALS

2150 ECGs were divided into a training set of 716 ECGs (Set C) and two test sets, each being composed of 717 ECGs (Sets D and E) which were almost equally randomly separated.

4.4.2.2 METHODS

In this stage, the number of input variables for the neural networks was held constant at 9, because this class of network performed better than the other two groups of neural networks using a different number of input parameters. For 30 separate neural networks used in the training process, more than 100,000 iterations were required for each network to reach the fully trained status, i.e. the rhythm of each ECG in the set C could be well recognised by the neural network without mistake. Then the fully trained networks were assessed by using ECGs from a separate set D. The architecture of the hidden layer was selected at the initial development of the neural networks. The fully trained neural networks with a varying number of neurons in the hidden layers and different numbers of hidden layers were subsequently tested by set D through the assessment of the Association Index etc in order to select the best topology of the hidden layers.

The numerical outputs from the neural networks for cases of AF were found to cluster in the range 0.5 to 1 and for the Non-AF in the range 0 to 0.5, and so 0.5 was selected as the cutoff point. Therefore, it was thought unnecessary to plot thirty Receiver Operating Characteristic curves, one for each of the neural networks. The best performing neural network with the optimum topology of hidden layers was then selected by reviewing the Association Index (Rautaharju, Blackburn & Warren, 1976) based on the assessment of set D ECGs using 0.5 as the discriminating point.

4.4.2.3 RESULTS

The effect of the varying design of the hidden layer(s) on the networks with 9 input variables is shown in Table 4-5. The artificial neural network with the best performance was found to be the network with 2 hidden layers each with 50 neurons inside and only a single output. The sensitivity of 89.9% and specificity of 92.8% yielded an Association Index of 82.7%. A single Receiver Operating Characteristic curve with varying cut off points from 0.1 to 0.9 was plotted for this neural network to assess the selection of the cutoff point 0.5. The false negative and false positive percentages of the results were determined by comparing the numerical output of the network with the true diagnosis of the ECG inside set D. The Receiver Operating Characteristic curve for this neural network for the diagnosis of atrial fibrillation is shown in Fig 4-5. The optimal cut off point indeed appeared to be 0.5. This 9v-50-50-1 neural network was then used in subsequent studies.

Number of neurons in hidden layers	Sensitivity %	Specificity %	Association Index %
5	92.7	83.5	76.2
5-5	91.6	89.2	80.8
10	93.3	84.0	77.3
10-10	92.4	84.5	76.9
15-15	91.2	86.1	77.3
20	93.3	77.8	71.1
20-20	89.1	85.1	74.2
30	91.2	88.7	79.9
30-30	91.0	85.6	76.6
40	92.7	87.6	80.3
40-40	92.9	87.6	80.5
50	92.2	86.6	78.8
50-50	89.9	92.8	82.7
50-50v2	92.0	84.0	76.0
50-50v3	90.6	88.7	79.3
60	94.0	89.2	83.2
60-60	89.9	92.3	82.2
60-60v2	92.5	88.1	80.6
70	91.8	85.1	76.9
70-70	90.6	87.6	78.2
80	94.3	85.1	79.4
30-40	91.6	82.5	74.1
30-50	90.4	88.1	78.5
30-60	90.4	87.1	77.5
40-50	90.8	89.2	80.0
40-60	90.3	87.6	77.9
40-70	91.0	84.5	75.5
50-60	89.9	83.5	73.4
50-70	88.0	88.1	76.1
50-40	89.5	86.1	75.6

Table 4-5 Effect of varying design of hidden layer of neural network for the diagnosis of atrial fibrillation on the set D ECGs. *v2 v3 denotes networks with different initial weights compared to others of identical design.

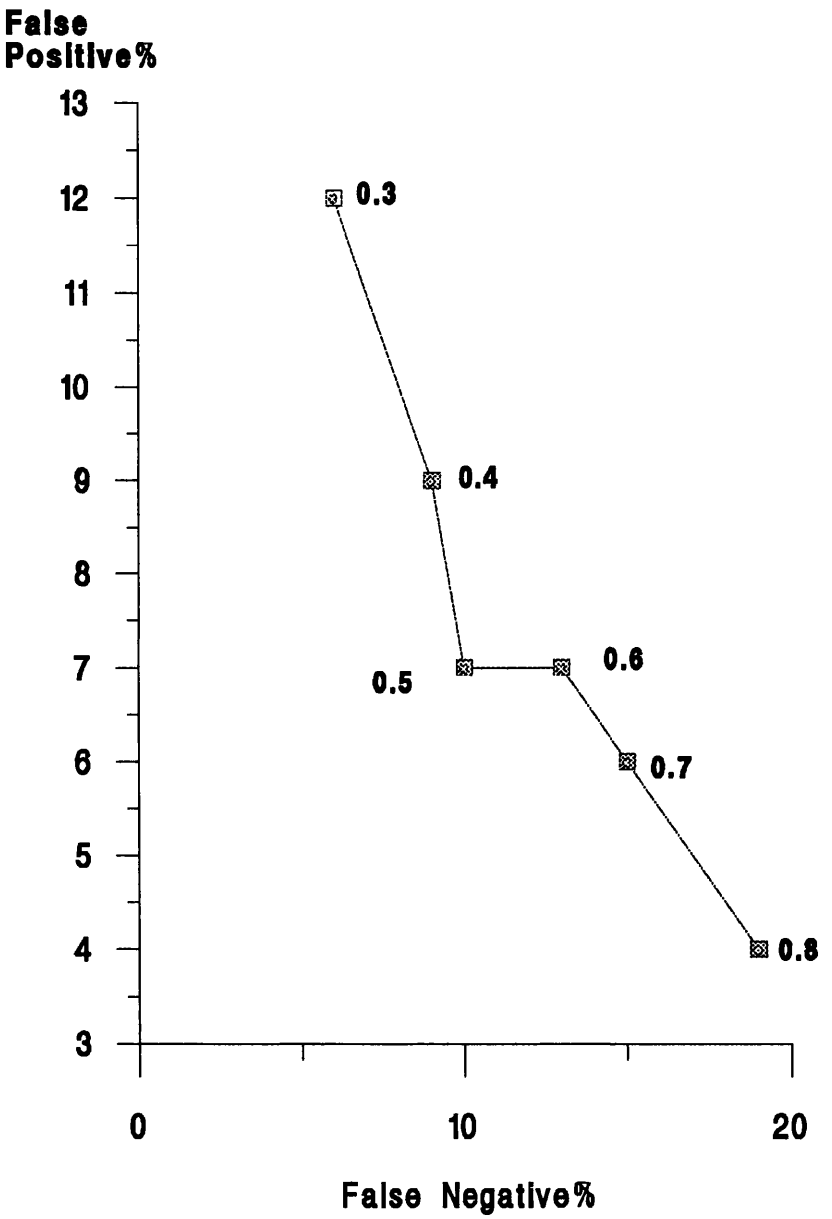


Fig.4-5 Receiver Operating Characteristic curve of the best neural network in stage 2a.

It was felt that the network with 60 hidden neurons in a single hidden layer was too non specific (89.2%), although the Association index was 83.2%. The Association Index was in general lower in the test set of stage 2 compared to similar networks in stage 1 where a different test set of ECGs was studied.

4.4.3 Stage 2b

Various Methods of Combining the Interpretation from the Deterministic Logic and Output from the Artificial Neural Networks

4.4.3.1 MATERIALS

The materials used were the same as described in 4.4.2.1.

4.4.3.2 METHODS

Several attempts were made to improve the sensitivity and specificity of the results for the diagnosis of atrial fibrillation by various combinations of the Glasgow Program interpretation and artificial neural network outputs. These methods were as follows:

[Option 1] If the Glasgow Program diagnosed "AF", this was accepted and the diagnosis of the best neural network (9v-50-50-1) was used for the remainder, i.e. those ECGs initially diagnosed as "Non-AF" by the Glasgow Program were subsequently classified by the single best performing neural network.

[Option 2] If the Glasgow Program diagnosed "SR + (SVEs &/or VEs)", this was accepted and the neural network diagnosis was used for the remainder, i.e. those ECGs initially classified by the Glasgow Program as not having "SR + (SVEs &/or VEs)" were subsequently classified by the single best performing neural network.

[Option 3] The addition of the neural network output to the Glasgow Program interpretation. The interpretation of Glasgow Program was assigned a value "1" if "AF", and "0" if "Non-AF". The diagnosis of the Glasgow Program and the output of the neural network were then summed to give a value between "0" and "2". Then, the best discriminating point was selected by plotting an ROC curve.

The best cutoff point was also found to be 0.5 in various combinations of the interpretation from the Glasgow Program and the output from the best neural network (9v-50-50-1), even in option 3 [Fig 4-6].

717 ECGs (523 AF + 194 non-AF) of set E were then subsequently used to assess the performance of the various combinations of interpretation from the deterministic logic and output from the neural network, i.e. a further final test set which had not been involved in any part of network design, selection of cut off point, etc., was used.

4.4.3.3 RESULTS

The results of the various methods of combining the outputs from the Glasgow Program and the best neural network for the diagnosis of atrial fibrillation on test set D are shown in Table 4-6. The sensitivity and

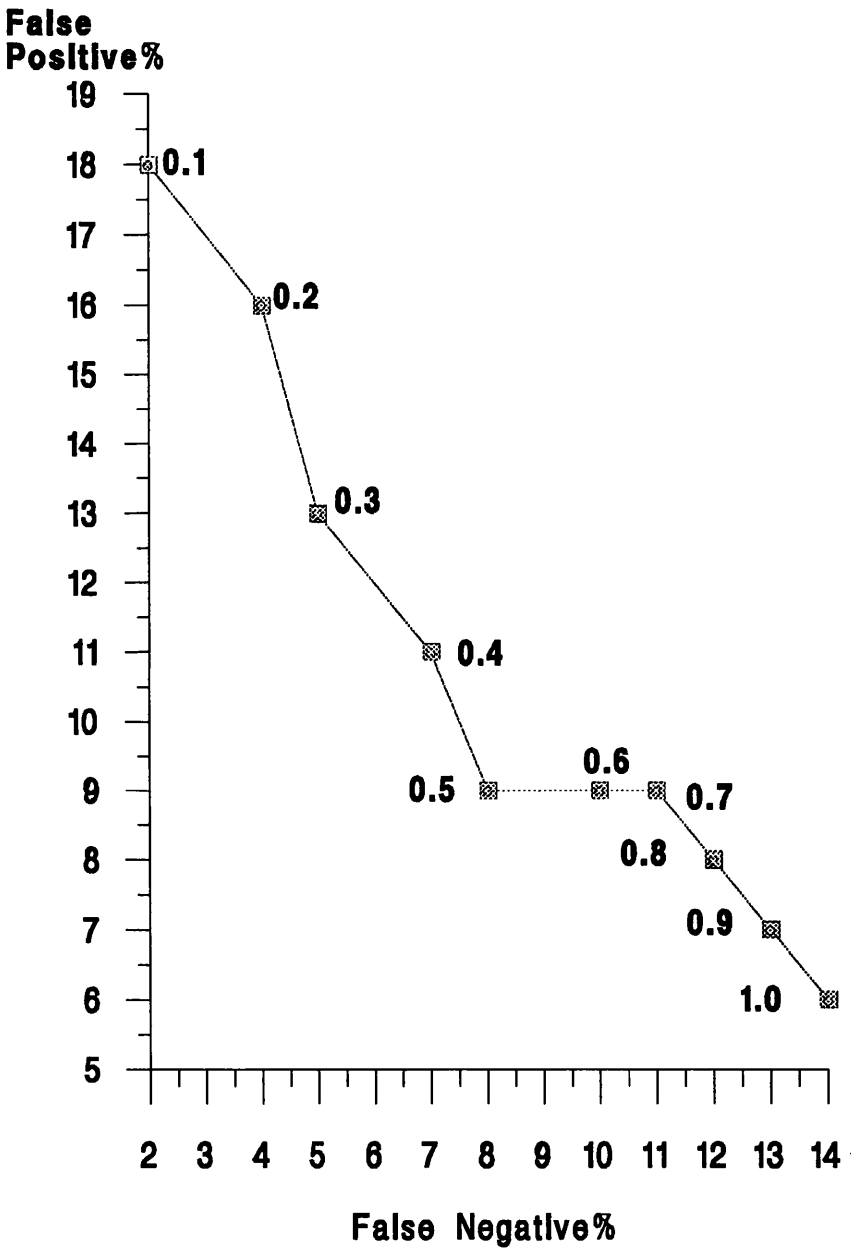


Fig.4-6 Receiver Operating Characteristic curve derived from combining results from Glasgow Program and the best neural network.

specificity of the original Glasgow Program was 86.2% and 93.8%, respectively (i.e. Association Index 80.0%), whereas, the sensitivity and specificity of the best neural network was 89.9% and 92.8%, separately (i.e. Association Index 82.7%). The options 1 and 3 did improve the sensitivity (Table 4-6) but jeopardised the specificity (93.8% for the Glasgow Program & 92.8% for the best neural network versus 90.7% for the two best options).

The effect of combining the deterministic logic and the best performing network was then assessed on the basis of the study of set E ECGs (Table 4-7). The performance of the best neural network (9 input variables with two 50-neuron hidden layers and a single output) alone (Association Index 84.3%) exceeded that of the various combinations (Association Index 83.2%, 81.8%, and 83.2% for option 1, 2, and 3, respectively) and the original Glasgow Program itself (Association Index 80.8%). It was thus demonstrated that the combination of the deterministic logic and the best neural network did not lead to any appreciable improvement compared to the use of a neural network alone.

Test Set D	Sensitivity	Specificity	AI
GP	86.2	93.8	80.0
NN	89.9	92.8	82.7
Option 1	92.3	90.7	83.0
Option 2	83.7	95.9	79.6
Option 3	92.4	90.7	83.1

Table 4-6 Results of the comparison among various combinations of the original Glasgow Program (GP) and the best neural network (NN) on ECGs of test set D.

Test Set E	Sensitivity	Specificity	AI
GP	88.5	92.3	80.8
NN	92.0	92.3	84.3
Option 1	94.0	89.2	83.2
Option 2	86.4	95.4	81.8
Option 3	94.0	89.2	83.2

Table 4-7 Results of the various combinations of the Glasgow Program (GP) and the best performing network (NN) on the ECGs of set E.

4.4.4 STAGE 3

Modification of Existing Deterministic Logic in the Glasgow Program

It was noted that certain patterns of the combinations of erroneous interpretations occurred in the deterministic approach, e.g. **SINUS RHYTHM + FIRST DEGREE AV BLOCK + SVEs** was sometimes reported instead of **ATRIAL FIBRILLATION**. It was postulated that if these patterns could be detected and the report modified prior to the final output, then the sensitivity might be improved without jeopardising the specificity.

4.4.4.1 MATERIALS

Set C (716 ECGs - 523 AF + 193 Non-AF) was used as a training set to select the optimal method of modifying the existing diagnostic deterministic logic in the Glasgow Program, and the 1434 ECGs of both sets D and E were used as an independent test set.

4.4.4.2 METHODS

After careful review of the ECGs of set C, there were two strategies A and B proposed for the modification of the deterministic logic. Basically the difference lay in the varying degree of prolonged PR interval.

These proposed modifications of the deterministic logic were as follows:

(1) Sinus Rhythm* + SVEs + PR interval > K ms

(2) Ectopic Rhythm + SVEs + PR interval > K ms

(3) Irregular Supraventricular Rhythm or Irregular Rhythm

Note that the Sinus Rhythm* includes sinus arrhythmia, bradycardia, tachycardia, etc. ms= milliseconds.

In strategy A, K=220; while in strategy B, K=240.

Both strategies were subsequently assessed using the 716 ECGs of set C and the best method of modification was selected by comparing the Association Indices among the Glasgow Program, the best performing neural network (9 input variables with two 50-neuron hidden layers and a single output), and the modified Glasgow Program using both strategy A and B. Then the remaining 1434 ECGs of both set D and E (1064 AF and 388 not AF) were used as a test set to evaluate the performance of the modified Glasgow Program in comparison with the original Glasgow Program and the best neural network.

McNemar's test of the significance of changes was used to evaluate the statistical significance of the improvement in sensitivity and specificity

before and after the modification of the Glasgow Program. A P value less than 0.05 is regarded as statistically significant.

4.4.4.3 RESULTS

The deterministic logic in the Glasgow Program modified by the strategy A (i.e. PR > 220ms) was found to perform better for the detection of atrial fibrillation in ECGs of set C than that modified by Strategy B (i.e. PR > 240ms) (Association Index: 85.2% versus 84.5%).

Both modified Glasgow Programs were observed to perform better than either the original Glasgow Program (Association Index: 80.0%) or the best neural network (Association Index: 82.7%) based on the study of ECGs of set C (Table 4-8).

Set C	Sensitivity	Specificity	AI
Original GP	86.2	93.8	80.0
NN	89.9	92.8	82.7
Strategy A	91.9	93.3	85.2
Strategy B	91.2	93.3	84.5

Table 4-8 Results of the comparison among modification of the existing deterministic logic by strategy A or B, neural network (NN) and Original Glasgow Program (GP) using set C as a training set.

Therefore, Strategy A was adopted for the modification of the original Glasgow Program. The final evaluation was subsequently performed on the remaining sets D and E (1434 ECGs). Results are presented in Table 4-9.

The modified Glasgow Program (Association Index 83.8%) was found to be the best among all the methods after assessment of the ECGs of set D & E (Table 4-9).

Set D & E	Sensitivity	Specificity	AI
Original GP	87.3	93.1	80.4
NN	91.0	92.6	83.6
Modified GP	92.3	91.5	83.8

Table 4-9 Results of the final assessment of the modified Glasgow Program (GP) using test sets D and E.

McNemar's test for the significance of changes was applied to evaluate the changes in sensitivity and specificity of the Glasgow Program before and after the modification as well as to compare the modified Glasgow Program with the best performing neural network (9v-50-50-1). The Tables for use in testing the significance of changes are shown below (Table 4-10, 11, 12, & 13).

It was observed that there was a significant improvement in the sensitivity for the detection of atrial fibrillation by the modified Glasgow Program as compared to the original Glasgow Program (McNemar's test: $X^2 = 27.04$, $P < 0.01$) and to the best performing neural network (McNemar's test: $X^2 = 6.82$, $P < 0.05$). On the other hand, the original logic was significantly more specific than the modified program (McNemar's test: $X^2 = 6$, $P < 0.05$), but there was no significant difference in the specificity of the modified logic compared to the best network (McNemar's test: $X^2 = 1$, $P = NS$).

It can be seen that the performance of the neural network for the detection of atrial fibrillation is not superior to the use of modified deterministic logic, but is better than the original deterministic logic.

	Original(TP)	Original(FN)	Total
Modified(TP)	890	76	966
Modified(FN)	24	56	80
	914	132	1046

X^2=27.04, P <0.01.

Table 4-10 Table for the McNemar's test of significance of changes of the sensitivity for the detection of AF between the modified and the original logic. [TP: True Positive, FN: False Negative]

	Original(TN)	Original(FP)	Total
Modified(TN)	355	0	355
Modified(FP)	6	27	33
	361	27	388

X^2=6, P <0.05.

Table 4-11 Table for the McNemar's test of significance of changes of the specificity for the detection of AF between the modified and the original logic. [TN: True Negative, FP: False Positive]

	NN(TP)	NN(FN)	Total
Modified(TP)	942	24	966
Modified(FN)	9	71	80
	951	96	1046

X^2=6.82, P <0.05.

Table 4-12 Table for the McNemar's test of significance of changes of the sensitivity for the detection of AF between the modified logic and the best performing network. [TP: True Positive, FN: False Negative]

	NN(TN)	NN(FP)	Total
Modified(TN)	349	6	355
Modified(FP)	10	23	33
	359	29	388

X^2=1, P =NS.

Table 4-13 Table for the McNemar's test of significance of changes of the specificity for the detection of AF between the modified logic and the best performing network. [TN: True Negative, FP: False Positive]

4.5 DISCUSSION

Currently, deterministic logic is applied in the Glasgow rhythm analysis program and is suitable for the identification of certain rhythm abnormalities with well defined exact limits for specific measurements, e.g. first degree atrioventricular conduction block (PR>0.22s) tachycardia (rate>100) or bradycardia (rate<60) for heart rate. However,

atrial fibrillation represents a continuum in respect of heart rate, RR variability, PR regularity etc. and deterministic criteria may well not be optimum. Therefore the purpose of the present study was to evaluate the feasibility of the application of neural networks alone or in combination with the original logic at a specific point in the diagnostic logic in order to improve the detection of atrial fibrillation.

4.5.1 Design of the Present Study

First of all, it should be noted that in the present study the neural network was applied at a particular point inside the deterministic logic of Phase 2 of the Glasgow Program where the decision on atrial fibrillation versus SR+(SVEs and/or VEs) is made. Therefore, for instance, regular rhythms, or atrial flutter with fixed AV block if detected correctly, would not be dealt with by these neural networks which were trained on ECGs that reached the selected decision point in the logic of the Glasgow Program. Secondly, it should also be noted that the specificity for detecting AF in the present study was assessed against SR + (SVEs and/or VEs) and not on regular sinus rhythm. Therefore, the specificity would almost certainly have been higher if regular sinus rhythm had been included in the test set. Thirdly, in an earlier report (Macfarlane, 1986b) from this institute on the accuracy of reporting atrial fibrillation using the Glasgow Program, a sensitivity of 98.5% was indicated, based on a study of 94 cases of atrial fibrillation. The sensitivity of Glasgow Program for the present ECGs was obviously lower (86.2% and 88.5% for test sets D and E, respectively), even after modification (92.3% for both test sets D and E), because the present study was based on a large number of more "complex" ECGs from patients with a wide variety of acute and chronic

cardiac diseases and probably gives a fairer reflection of the accuracy of reporting atrial fibrillation in a tertiary referring hospital population.

4.5.2. Reasons for Using Three Sets of ECGs in Stage 2

In stage 2, the reasons why the three sets of ECGs (C, D, E) were required are as follows:

(1) It was necessary to select the best design and optimal cutoff point for the best neural network output and to manipulate the various combinations of the deterministic logic results with the neural network output through utilising sets C and D.
(2) Set E had to be used as a fifth, completely independent set for the final assessment of the performance of the best neural network chosen from the above mentioned process, as well as of the modified Glasgow deterministic program.

4.5.3 Neural Network versus Deterministic Logic versus Combination Method

The performance of the best neural network for the diagnosis of atrial fibrillation was shown to be better than the original deterministic program in respect of sensitivity (92.0% versus 88.5%) while preserving the specificity (92.3%) (Table 4-7). However, the technique of combining the output from the neural network with the interpretation from the deterministic program did not further improve the performance for the detection of atrial fibrillation (Table 4-7). Nevertheless, after suitable adjustment of the existing deterministic criteria, the deterministic program was found to have a statistically significantly (Tables 4-10 & 4-

12) improved sensitivity (92.3%) in detecting atrial fibrillation as compared to either the original program (sensitivity= 87.3%) or the best neural network (sensitivity= 91.0%) (Table 4-9), but with a decreased specificity (91.5% versus 92.6% & 93.1% respectively) (Tables 4-9, 4-11, & 4-13).

Thus, the conclusion of this study was that assessment of the neural network in the diagnosis of atrial fibrillation had acted as a catalyst for the improvement of the original logic, which ultimately had a performance essentially equivalent to that of the network.

The results also show that the performance of all the approaches varies little depending on the test set used. For example, the same network tested on set D, set E as well as set D & E combined had sensitivities of 89.9%, 92.0% and 91.0% with corresponding specificity of 92.8%, 92.3% and 92.8%. However, the performance of neural networks with the same topology but different initial weights is not reproducible so that it is the relative performance of different approaches within the same test set that is of importance rather than absolute values of sensitivity etc.

4.5.4 Experience Concerning the Neural Network Approach

Several points can be drawn from the present study.

First of all, many years of experience required to produce a section of deterministic logic for reporting atrial fibrillation can be replaced in a rather short time by an artificial neural network which can achieve an essentially equivalent performance to the original logic.

The second point would seem to be that, in view of the foregoing, if further improvements are to be obtained, then new input parameters will have to be taken into consideration. For instance, the difference in area under the curves after superimposing two RR intervals of similar length in a Holter recording (Cubanski et al, 1993) might in the future be a useful new input parameter for the detection of atrial fibrillation. The possibility is that the availability of new input measurements would enhance the performance of both deterministic logic and an artificial neural network but the time required for an appropriate neural network to be developed might well be less.

According to Erb (1993), the number of hidden neurons in a network can be critical to the performance of the neural network. If there are too few neurons inside the hidden layer of the neural network, the neural network will lack the power it needs to classify patterns in the analysing data. If there are too many neurons in a hidden layer, specific patterns will be memorised. Memorisation handicaps the ability of the neural network to generalise. In this case, an artificial neural network has so many degrees of freedom that it responds with memorised facts rather than with an estimated value based upon the general features of the facts (Maren, Harston & Pap, 1991). Memorisation is analogous to fitting a set of N data points with an N-degree polynomial. A perfect fit results, but the fit's ability to interpolate is diminished. In a neural network, complete memorisation is said to occur when the number of hidden neurons equals the number of facts used to train the neural network.

Kolmogorov's theorem predicts that twice the number of input neurons plus one is a sufficient number of hidden neurons to compute any arbitrary continuous function (Hecht-Nielsen, 1987 & 1990), but this

topology is not necessarily the one with the best performance. This was clearly demonstrated in the present study. For example, the topology of the hidden layer of the best performing neural network in stage 1 is 9v-30-30-1 and not the 9v-20-1 (Table 4-4), while in stage 2, the best performing network is 9v-50-50-1 and not the 9v-20-1 (Table 4-5).

It has been learned from the present study that although the increase in the number of hidden neurons or hidden layers could shorten the time needed for completion of the whole training process because of fewer iterations, it took longer to have every single training run completed, and this did not necessarily improve the performance of the neural network.

Some of the characteristics of neural networks were also learned from the training process. Different training sets can lead to different results (internal representations or weights) even from neural networks with the same topology. The selection of training set ECGs has a significant, profound effect on the performance of the networks. It was also demonstrated that the greater the variation of the patterns of atrial fibrillation in the training set, the more the time needed to complete the training process.

This study also confirmed the previous report by Dassen et al (1993) that the selection of the training set and the input variables used to train the neural network has considerable influence on the final result, i.e. even with the same topology of the hidden and output layers, different input parameters or different initial weights for the same input variables can lead to a completely different result. It was also observed that it took longer to train a more complex neural network, but this did not necessarily yield a better result. The number of input variables, different

initial weights, and various topologies of the hidden layer could produce different effects on the performance of artificial neural networks.

4.6 CONCLUSIONS

(1) The application of a trained artificial neural network improved the sensitivity but slightly decreased the specificity for the detection of atrial fibrillation as compared to the use of the original deterministic logic.
(2) The performance of neural network assisted logic (i.e. various combinations of the neural network output and the results from the deterministic logic) was not superior to the use of a neural network alone in the diagnosis of atrial fibrillation as assessed in the specific test set (Tables 4-6 & 4-7).
(3) The modification of the original logic can achieve a noticeable improvement in detecting atrial fibrillation compared to both a neural network and the original deterministic logic.
(4) The choice of diagnostic parameters appears to be more important than using either a neural network or deterministic logic in the diagnosis of atrial fibrillation.
(5) The selection of test material is also an important factor that can affect the result. Various test sets can produce different results when tested by the same neural network.

This study only involved a small part of the rhythm analysis, and there are still areas that could be investigated. It might be possible in the near future to incorporate multiple neural networks at various points within the deterministic logic framework in order to enhance the overall accuracy of the rhythm interpretation in computer assisted ECG analysis. Whether or not this is a realistic goal remains an area for further investigation and beyond the scope of the present study.

CHAPTER 5

SOFTWARE BASED ARTIFICIAL NEURAL NETWORKS FOR THE COMPUTER ASSISTED ECG DIAGNOSIS OF MYOCARDIAL INFARCTION

5.1 INTRODUCTION

Myocardial infarction is one of the major cardiovascular causes of death in Western industrialised countries. Although myocardial infarction can only be definitely diagnosed by pathological evidence of myocardial necrosis and scarring, clinically it is usually detected by the medical history, subjective symptoms, physical signs, changes in the electrocardiogram, and a rise in cardiac enzymes. Of these, the electrocardiogram is the most readily available, relatively inexpensive, and reproducible non-invasive investigational tool for the objective diagnosis of myocardial infarction.

Considerable efforts have been made to improve the accuracy of diagnosing myocardial infarction clinically. For instance, stepwise discriminant analysis (Pozen, Stechmiller & Voigt, 1977), logistic regression (Pozen et al, 1980), recursive partition analysis (Goldman et al, 1982), pattern recognition (Patrick et al, 1976 & 1977), and artificial neural networks (Baxt, 1991 a & b) have been applied in different phases of data analysis and decision making for the detection of myocardial infarction. These reports mainly focused on the clinical data analysis and decision making.

The performance of the neural network is claimed to be better than that of physicians both in sensitivity (97.2% versus 77.7%) and in specificity (96.2% versus 84.7%) (Baxt, 1991b). However, the study by Baxt (1991b) included the medical history, symptoms, physical signs, and ECG ST and T wave changes as input variables of the neural networks. It was thought that the human observations, such as subjectively observed signs, might have interobserver variability which could lead to a decrease in accuracy of the neural network diagnostic system. The magnitude of these inaccuracies will increase in parallel with interobserver variabilities which in turn will increase when the neural networks are used outside their place of development. In contrast, software based neural networks, which adopt the ECG measurements as inputs, should have the same accuracy even outside the laboratory where they were developed.

On the other hand, the ECG provides unique and objective non-invasive evidence of myocardial ischaemia and infarction as compared to other subjective symptoms such as chest pain, diaphoresis, etc. Therefore, only ECG measurements were used for the neural network diagnosis of myocardial infarction in the present study.

The definitive diagnosis of myocardial infarction, either in the acute phase or the post myocardial infarction period, is of therapeutic and prognostic importance (AIMS Trial Study Group, 1988; Gruppo Italiano per lo Studio della Streptochinasi nell'infarcto Miocardico (GISSI), 1987; ISIS-2, 1988). In the emergency department or general wards, where chest pain cases are usually first seen by relatively inexperienced junior medical staff, computer-assisted ECG interpretation can provide supplementary objective evidence to aid in the preliminary differential diagnosis of myocardial infarction before cardiac enzyme assay results

are available. Frank vectorcardiographic criteria for both anterior (Starr et al, 1974) and inferior (Starr et al, 1976) myocardial infarction have been proposed to improve the ECG diagnosis. Although lack of superiority of the vectorcardiogram over the electrocardiogram in detecting inferior wall myocardial infarction regardless of time since infarction has also been claimed by one group of investigators (Lui et al, 1987), other researchers have claimed improved diagnosis of inferior myocardial infarction using either the Frank vectorcardiogram (Chou, 1986) or the derived vectorcardiogram (Edenbrandt et al, 1990).

Computer-assisted ECG interpretation of myocardial infarction is still an area where much research is being undertaken to improve the sensitivity and specificity of the diagnosis. Most computer ECG analysis programs use either deterministic logic or statistical techniques for the diagnosis of myocardial infarction.

The main advantage of the application of the artificial neural networks over the traditional deterministic logic techniques is their self-learning capacity, i.e. they do not require any background knowledge from their designers, and they also perform relatively better in a complex relationship representation, i.e. it takes many years to develop a deterministic logic program as compared to several months to train a set of neural networks to achieve almost the same level of performance.

Therefore, the aim of this part of the study was to evaluate whether the use of software based neural networks and the addition of derived vectorcardiographic measurements to the list of 12-lead ECG variables input to the neural network are beneficial or not in the computer ECG diagnosis of myocardial infarction.

5.2 MATERIALS AND METHODS

A total of 515 ECGs were recorded from individuals who had suffered from either old or acute myocardial infarction (255 anterior myocardial infarction- 196 males and 59 females; 260 inferior myocardial infarction- 210 males and 50 females). All patients had been admitted to the Cardiology and Cardiac Surgery Wards in Glasgow Royal Infirmary between October 1991 and September 1993. ECGs were recorded from the acute cases in the coronary care unit within 5 days after the onset of chest pain, and from those with an old infarct at least 3 months post infarction.

The mean age of myocardial infarction patients included in the present study was as follows:

AGE (years)	anterior MI	inferior MI
Mean±SD	57±15	58±10
Male	56±16	58±11
Female	60±8	61±7

Table 5-1 Mean age of myocardial infarction patients included in the present study. (SD= standard deviation)

All the patients selected for this study had undergone cardiac catheterisation and left ventriculography to delineate their coronary artery anatomy and left ventricular wall motion abnormalities. Some patients recruited from the surgical ward also had their infarct confirmed by the operative finding of myocardial scarring and fibrosis.

The clinical criteria for the inclusion of the myocardial infarction patients were as follows:

1.Typical clinical history and physical signs.

a. Typical Clinical History.[Maximum:2 points]

i. Chest pain at the time of initial examination. [2 points]

ii. Associated symptoms : [any one of the following four or any of their combinations: 1 point]

Nausea and vomiting,

Profuse diaphoresis,

Sensation of terror or impending doom,

Syncope.

b. Physical signs: [any one of the following or any of their combinations: 1 point]

i. Palpable dyskinesia,

ii. Hypotension,

iii. Soft first heart sound,

iv. Paradoxical splitting of second heart sound,

v. Accentuated third or fourth heart sound,

vi. Murmur of mitral regurgitation,

vii. Signs and symptoms of heart failure.

2. Serial electrocardiographic changes. [Maximum: 3 points]

a. ST segment deviation [2 points]

b. T wave abnormalities [1 point]

c. Alterations of QRS complex [2 points]

d. Conduction disturbances [1 point]

3. Serial changes of cardiac enzymes. [Maximum: 3 points]

a. LDH1/LDH2 > 1 or Total LDH > 2 x Normal [2 points]

b. CKMB isoenzyme > 2.2 x Normal [2 points]

c. AST > 2 x Normal [1 point]

4. Echocardiographic studies. [1 point]

a. Wall motion abnormalities

5. Radionuclide studies. [Maximum 4 points]

a. Thallium-201 perfusion scan : filling defect

[2 points]

b. Technetium-99 pyrophosphate infarct-avid scan : hot spot [2 points]

6. Cardiac catheterisation and angiography.

[Maximum: 3 points]

a.Wall motion abnormalities [any of hypokinesia , akinesia, dyskinesia or aneurysm: 2 points]

b.High grade [>75%] stenosis of coronary arteries [1 point]

7. Operative findings at Surgery. [Maximum: 6 points]

a. Scar [3 points]

b. Aneurysm [3 points]

If the total score is more than 6, then the diagnosis of myocardial infarction is made.

Anterior, anteroseptal, apical, and lateral myocardial infarction for the purpose of the current study were classified as **ANTERIOR MYOCARDIAL INFARCTION**. There were therefore two categories used in the present study: **ANTERIOR** and **INFERIOR MYOCARDIAL INFARCTION**.

In addition to the above ECGs from patients with myocardial infarction, 144 ECGs from clinically validated left ventricular hypertrophy patients (91 males and 53 females) with valvular or hypertensive heart disease without evidence of myocardial infarction were obtained. All patients had two dimensional and M-mode echocardiographic evidence of left ventricular hypertrophy. The criteria for left ventricular hypertrophy on the M-mode echocardiogram were left ventricular mass greater than 132 g/m^2 for males and 109 g/m^2 for females, respectively (Huwez, 1991; Pringle, 1990). The mean age of the left ventricular hypertrophy patients was 57±12 years, while for 91 males the mean age was 57±12 years, and for 53 females the mean age was 59±12 years.

A total of 580 ECGs from normals were randomly selected from the 1555 ECGs in the data bank of the department of Medical Cardiology, Glasgow Royal Infirmary. These ECGs were recorded from apparently healthy individuals recruited for studies on the normal limits of the ECG during 1981-1992. They were employees of the Strathclyde Regional Council. All of them had a normal medical history, physical examination, chest radiograph and echocardiogram. None had a cardiovascular abnormality or other disease known to affect the cardiovascular system e.g. endocrine disease, diabetes mellitus (Macfarlane, Chen & Chiang, 1988). In addition, 30 individuals with an electrocardiographic pattern of poor R wave progression over the praecordial leads were also included in the normal controls. They were being investigated for chest pain and all had a normal coronary arteriogram and left ventriculogram.

All ECG signals were recorded by digital-to-analogue conversion with a locally developed electrocardiograph (Watts & Shoat, 1987) with 500 samples/second (500 Hz) directly from either the patients in the wards or apparently healthy volunteers in the normal clinics. The digitised data were subsequently transmitted to a central computer and later retrieved for analysis. Various measurements of durations and amplitudes of the electrocardiographic QRS and ST-T waves as well as derived vectorcardiographic parameters were thereafter obtained using the signal processing software inside the Glasgow Program (Macfarlane et al, 1990b). These selected electrocardiographic and derived vectorcardiographic measurements were then used as input variables to the neural networks. All electrocardiograms were interpreted by the deterministic logic of the Glasgow Program. The derived vectorcardiographic leads were also synthesized inside the program using the inverse Dower method (Edenbrandt & Pahlm, 1988b). Subsequently,

derived vectorcardiographic parameters were also extracted for input to the neural networks.

The artificial neural network simulation package has been described in detail in the previous chapters. Selected electrocardiographic and derived vectorcardiographic measurements (vide infra & Table 5-2) extracted by the Glasgow Program were fed into the input layer of various neural networks with different topologies for the purpose of training.

Input variables	AMI	IMI
ECG QRS parameters(Q)	21	9
ECG ST-T parameters(ST)	21	9
dVCG parameters(V)	4	3
ECG & dVCG Measurements	AMI	IMI
QRS+ST-T+dVCG	46	21
QRS+ST-T	42	18
QRS+dVCG	25	12
QRS	21	9

Table 5-2 Total numbers of selected electrocardiographic and derived vectorcardiographic parameters used as input to the neural networks for the diagnosis of myocardial infarction. (QRS = measurements from QRS complex, ST-T= measurements from ST-T segment, dVCG= measurements from derived VCG)

The total number of ECG and derived VCG parameters used as input to the neural networks for the diagnosis of anterior and inferior myocardial infarction is shown in Table 5-2.

The ECG and derived VCG measurements used as the input parameters to the neural networks for the diagnosis of myocardial infarction were as follows:

A. Anterior Myocardial Infarction

(1) Q wave amplitude in leads I, aVL, V_2-V_6
(2) Q wave duration in leads I, aVL, V_2-V_6
(3) R wave amplitude in leads I, aVL, V_2-V_6
(4) negative T wave amplitude in leads I, aVL, V_2-V_6
(5) ST segment amplitude in leads I, aVL, V_2-V_6
(6) positive T wave amplitude in leads I, aVL, V_2-V_6
(7) maximal angle of initial 20 millisecond vector in the horizontal plane
(8) the area bounded by the initial 30 millisecond QRS vector in the horizontal plane
(9) amplitude of VCG bite in the horizontal plane
(10) duration of VCG bite in the horizontal plane

B. Inferior Myocardial Infarction

(1) Q wave amplitude in leads II, III, aVF
(2) Q wave duration in leads II, III, aVF
(3) R wave amplitude in leads II, III, aVF
(4) negative T wave amplitude in leads II, III, aVF
(5) ST segment amplitude in leads II, III, aVF
(6) positive T wave amplitude in leads II, III, aVF
(7) X axis intercept in the frontal plane [Fig 5-1]
(8) maximal QRS vector angle in the frontal plane
(9) the area bounded by the initial 30 millisecond QRS vector in the frontal plane

There were 4 experiments included in this study.

In experiment 1, the neural networks were trained and tested in the two group situation (myocardial infarction versus normal) with only a single output neuron used for diagnosis.

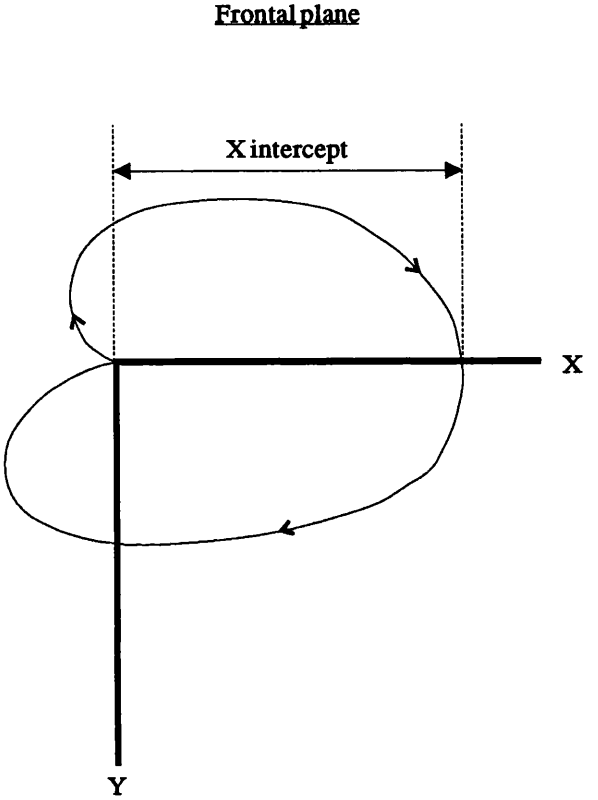


Fig. 5-1 X intercept of the derived VCG in the frontal plane used as input to the neural network for the diagnosis of inferior myocardial infarction.

In experiment 2, the neural networks were trained in the two group situation (myocardial infarction versus normal) and tested in the three group situation (myocardial infarction versus normal versus left ventricular hypertrophy) with only a single output neuron used for diagnosis.

In experiment 3, the neural networks were trained and tested in the three group situation (myocardial infarction versus normal versus left ventricular hypertrophy) with only a single output neuron used for diagnosis.

In experiment 4, the neural networks were trained and tested in the three group situation (myocardial infarction versus normal versus left ventricular hypertrophy), but with three output neurons for "Normal", " Myocardial Infarction", and " Left Ventricular Hypertrophy", respectively. The diagnosis associated with the neuron having the highest output value was selected as the classification.

The topology of hidden layers varied from one to three layers with the number of neurons inside each layer being adjusted between 5 and 100. The neural networks in experiments 1 to 3 had only a single output neuron with the classification of "Myocardial Infarction" or "Non-myocardial infarction".

In the training process, supervised feed-forward networks with a back propagation algorithm (Rumelhart, Hinton & Williams, 1986) were used. In all experiments except experiment 4, each ECG of the training set was assigned a classification of myocardial infarction (1) or non-myocardial infarction (0) and fed into the neural networks for training. In

experiment 4, each ECG in the training set was given a classification of myocardial infarction (1, 0, 0), normal (0, 1, 0), or left ventricular hypertrophy (0, 0, 1). Initial weights were randomly selected by the computer and were adjusted automatically in order to reduce the errors in representing the designated ECG (myocardial infarction or non-myocardial infarction except in experiment 4 where normal, myocardial infarction, or left ventricular hypertrophy were used) during the repeated exposure to the training set ECGs.

After more than 100,000 iterations in around 48 hours of learning for each network, all the cases in the training set were well learned and the neural network had the best internal representations (optimum weights) for each assigned classification (i.e. each ECG in the training set was correctly recognised by the neural network as myocardial infarction or non-myocardial infarction, except in experiment 4, where the ECG was recognised as normal, myocardial infarction, or left ventricular hypertrophy).

The tolerance of the training was set at 0.1. The gain (which is the convergence factor used in the back-propagation learning algorithm) of the network was set at 1.0. After the neural networks were fully trained, ECGs from a separate test set were subsequently used to assess the performance of the neural network.

The various combinations of ECG and derived VCG parameters used in the input layer included selected QRS, ST-T, and derived VCG measurements. The output from the neural network was a value ranging from 0 to 1. In the evaluation process, a threshold value of 0.5 was used to discriminate between myocardial infarction and non-myocardial

infarction (i.e. if the network output was below 0.5 then there was no infarct deemed to be present, otherwise myocardial infarction was reported by the network). Because the numerical outputs from the neural network for myocardial infarction were clustered from 0.5 to 1, and for non-myocardial infarction from 0 to 0.5, it was unnecessary to plot a Receiver Operating Characteristic curve to select the optimal discriminating point (cf. Fig. 4-4). There were three numerical outputs from experiment 4, one for myocardial infarction, one for left ventricular hypertrophy, and another for normal. Each value also ranged from 0 to 1. The highest value among the three classifications was chosen as the diagnosis of the neural network.

The topology of two examples of neural networks studied for the diagnosis of myocardial infarction used in these experiments is shown in Figures 5-2 and 5-3.

Definitions of the sensitivity, specificity, and Association Index (AI) for the diagnosis of myocardial infarction used in the current study are as follows:

Se	=	Sensitivity of myocardial infarction diagnosis in myocardial infarction patients.
N Sp	=	Specificity of myocardial infarction diagnosis in normals.
L Sp	=	Specificity of myocardial infarction diagnosis in left ventricular hypertrophy patients.
O Sp	=	Overall specificity of myocardial infarction diagnosis in normals and left ventricular hypertrophy patients.
AI	=	Se + N Sp -100%
O AI	=	Se + O Sp-100%

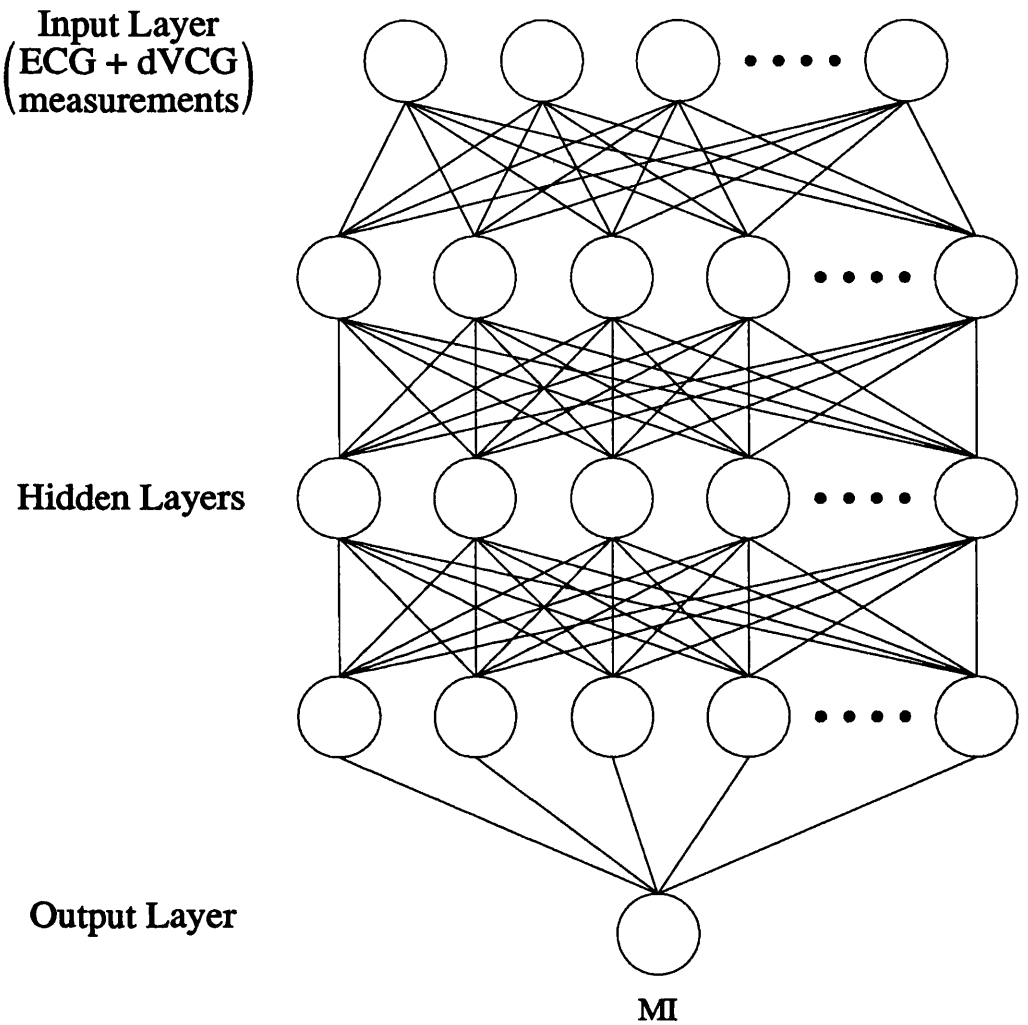


FIG 5-2. Topology of the Neural Networks used in Experiments 1, 2 and 3.

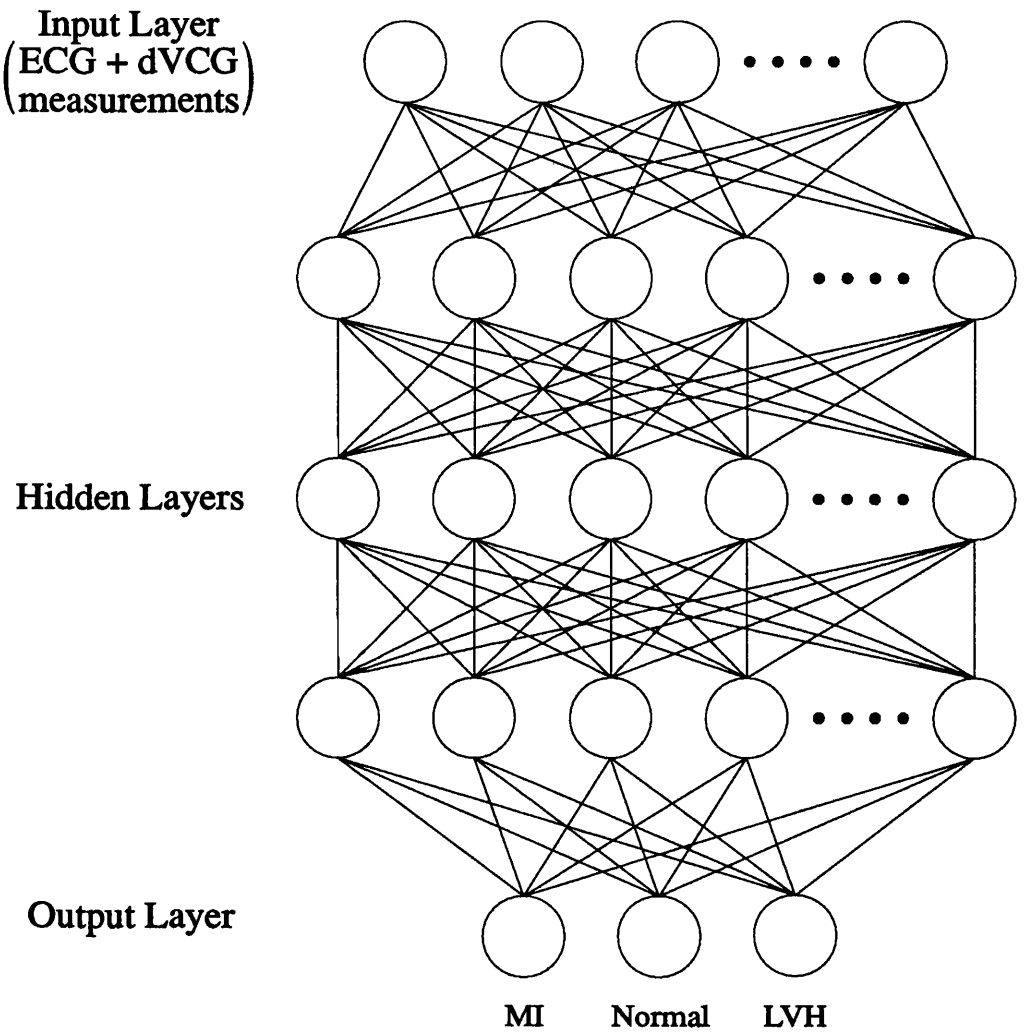


FIG 5-3. Topology of the Neural Networks used in Experiment 4.

5.3 EXPERIMENTS 1 and 2

TRAINING IN THE TWO-GROUP SITUATION
(Myocardial Infarction versus Normal)
ARTIFICIAL NEURAL NETWORKS WITH ONLY ONE OUTPUT NEURON

The aim of these experiments was to assess the usefulness of a neural network trained in the two-group situation (myocardial infarction versus normal) for the computer ECG diagnosis of myocardial infarction but tested separately **both** in the two-group **and** the three-group situation (myocardial infarction versus normal versus left ventricular hypertrophy).

5.3.1 EXPERIMENT 1

TESTING IN THE TWO-GROUP SITUATION
(Myocardial Infarction versus Normal)

The aim of this experiment was to test, in the two-group situation (Myocardial Infarction versus Normal), the performance of the neural networks trained in the two-group situation for the diagnosis of myocardial infarction.

5.3.1.1 Materials

Anterior Myocardial Infarction

The training set consisted of 80 ECGs from patients with clinically validated anterior myocardial infarction and 80 ECGs from normals

(including 30 ECGs showing poor R wave progression over the praecordial leads from apparently healthy individuals being investigated for atypical chest pain but shown to have normal coronary arteriograms).

The test set consisted of 101 ECGs from patients with clinically validated anterior myocardial infarction and 100 ECGs from normal subjects. (Table 5-3)

	Anterior MI	Normal
Training Set	80	80
Test Set	101	100

Table 5-3 The composition of the training set and test set ECGs used for the neural network diagnosis of anterior myocardial infarction in experiment 1.

Inferior Myocardial Infarction

The training set consisted of 100 ECGs from patients with clinically validated inferior myocardial infarction and 100 ECGs from normals.

The test set consisted of 108 ECGs from patients with clinically validated inferior myocardial infarction and 100 ECGs from normal subjects. (Table 5-4)

	Inferior MI	Normal
Training Set	100	100
Test Set	108	100

Table 5-4 The composition of the training set and test set ECGs used for the neural network diagnosis of inferior myocardial infarction in experiment 1

5.3.2 EXPERIMENT 2

TESTING IN THE THREE-GROUP SITUATION (Myocardial Infarction versus Normal versus Left Ventricular Hypertrophy)

The aim of this experiment was to test, in the three-group situation (myocardial infarction versus normal versus left ventricular hypertrophy), the performance of the neural networks trained in the two-group situation for the diagnosis of myocardial infarction.

5.3.2.1 Materials

Anterior Myocardial Infarction

The training set consisted of 80 ECGs from patients with clinically documented anterior myocardial infarction and 80 ECGs from normals (including 30 ECGs showing poor R wave progression over praecordial leads from apparently healthy individuals investigated for the atypical chest pain but found to have normal coronary arteriograms).

The test set consisted of 101 ECGs from patients with clinically documented anterior myocardial infarction, 84 ECGs from the patients with clinically validated left ventricular hypertrophy, and 100 ECGs recorded from normal subjects. (Table 5-5)

	Anterior MI	Normal	LVH
Training Set	80	80	0
Test Set	101	100	84

Table 5-5 The composition of the training set and test set ECGs used for the neural network diagnosis of anterior myocardial infarction in experiment 2.

Inferior Myocardial Infarction

The training set consisted of 100 ECGs recorded from patients with clinically documented inferior myocardial infarction and 100 ECGs recorded from normals.

	Inferior MI	Normal	LVH
Training Set	100	100	0
Test Set	108	100	84

Table 5-6 The composition of the training set and test set ECGs used in the neural network diagnosis of inferior myocardial infarction in experiment 2.

The test set consisted of 108 ECGs recorded from patients with clinically documented inferior myocardial infarction, 84 ECGs recorded from the patients with clinically validated left ventricular hypertrophy, and 100 ECGs recorded from normal subjects. (Table 5-6)

5.3.3 Methods

A total of 306 neural networks (167 for anterior myocardial infarction and 139 for inferior myocardial infarction) with a different number of input variables and various hidden layer topologies were fully trained and subsequently evaluated using ECGs from a separate test set. The method of training is described in section 5.2.

In experiment 1, only two subsets of neural networks with or without derived VCG input measurements (QRS+ST-T+derived VCG

measurements versus QRS+ST-T measurements) were trained and then evaluated.

In experiment 2, there were four subsets of neural networks trained with various combinations of different ECG and derived VCG input variables. Left ventricular hypertrophy cases were also included in the test set. The 4 subsets of neural networks for the diagnosis of anterior and inferior myocardial infarction are shown in the lower part of Table 5-2.

5.3.4 The Performance of Glasgow Program on The Training Set

The sensitivity and specificity of the Glasgow Program based on the training set used in experiments 1 and 2 for the diagnosis of anterior and inferior myocardial infarction are shown in Table 5-7.

Training Set	Anterior MI	Inferior MI
Sensitivity	75%	92%
Specificity	100%	100%

Table 5-7 Performance of the Glasgow Program for the diagnosis of anterior and inferior myocardial infarction on the training set.

5.3.5 Electrocardiographers' Interpretation of the Test Set

The same two test sets of ECGs with the composition shown in Tables 5-5 and 5-6 used for the assessment of the performance of the neural networks were also interpreted by two electrocardiographers without the knowledge of age, sex and clinical information. Only 12 average beats and two rhythm strips were given to the electrocardiographers and they

were asked to classify the ECG as myocardial infarction, normal or left ventricular hypertrophy. Thereafter, the usual statistical formulae for the calculation of specificity and sensitivity were used to assess their performance. Their results were averaged and compared to the results from the original Glasgow Program and the best neural networks.

5.4 RESULTS OF EXPERIMENTS 1 AND 2

5.4.1 ANTERIOR MYOCARDIAL INFARCTION

A total of 167 artificial neural networks for the diagnosis of anterior myocardial infarction were evaluated on the test set ECGs.

The performance of all 55 neural networks using 46 input variables (QRS+ST-T+dVCG) for the diagnosis of anterior myocardial infarction in experiments 1 & 2 is shown in Tables 5-8 a, b, & c (see Appendix 2).

The performance of all 52 neural networks using 42 input variables (QRS+ST-T) for the diagnosis of anterior myocardial infarction in experiments 1 & 2 is shown in Tables 5-9 a, b, & c (see Appendix 2).

The performance of all 30 neural networks using 25 input variables (QRS+dVCG) for the diagnosis of anterior myocardial infarction in experiments 1 & 2 is shown in Table 5-10 (see Appendix 2).

The performance of all 30 neural networks using 21 input variables (QRS) for the diagnosis of anterior myocardial infarction in experiments 1 & 2 is shown in Table 5-11 (see Appendix 2).

The results of the best performing networks, Glasgow Program and electrocardiographers for the diagnosis of anterior myocardial infarction are shown in Table 5-16.

For the diagnosis of anterior myocardial infarction, the best performing neural networks were as follows:

QRS+ST-T+dVCG (46v): 46v-5-5-1,
QRS+ST-T (42v): 42v-80-90-1,
QRS+dVCG (25v): 25v-90-90-1,
QRS (21v): 21v-30-30-1.

AMI	GP	Neural Network				H E
		Q+ST+V	Q+ST	Q+V	Q	
Se	89%	97%	95%	91%	88%	85%
N Sp	100%	99%	100%	97%	99%	100%
AI	89%	96%	95%	88%	87%	85%
L Sp	87%	50%	49%	77%	80%	95%
O Sp	94%	77%	77%	88%	90%	98%
O AI	83%	74%	72%	79%	78%	83%

Table 5-16 Result of the comparison among the Glasgow Program, the best performing neural networks, and human experts in the diagnosis of anterior myocardial infarction.(Q= QRS measurements, ST= ST-T measurements, V= derived VCG measurements, HE=human experts)

From Table 5-16, it can be seen that the electrocardiographers performed better than either the deterministic Glasgow program or the artificial neural networks in terms of the specificity of diagnosing anterior

myocardial infarction in the left ventricular hypertrophy cases (i.e. the electrocardiographers had less false positive cases probably because they knew the ECGs were from single disease cases). With respect to sensitivity of anterior myocardial infarction, the electrocardiographers' interpretations were inferior to both automated approaches but were 100% specific in the normal controls as was the deterministic approach.

The neural networks not using ECG ST-T measurements as input variables performed better with respect to the specificity of anterior myocardial infarction diagnosis in the left ventricular hypertrophy cases compared to the networks that did use them.

When ECG ST-T parameters were used as the input variables to the networks, the addition of derived VCG parameters seemed not to lead to any significant improvement in the detection of anterior myocardial infarction.

5.4.2 INFERIOR MYOCARDIAL INFARCTION

A total of 139 artificial neural networks for the diagnosis of inferior myocardial infarction were evaluated by the test set ECGs.

The performance of all 45 neural networks using 21 input variables (QRS+ST-T+dVCG) for the diagnosis of anterior myocardial infarction in experiments 1 & 2 is shown in Tables 5-12 a, b, & c (see Appendix 2).

The performance of all 36 neural networks using 18 input variables (QRS+ST-T) for the diagnosis of inferior myocardial infarction in experiments 1 & 2 is shown in Tables 5-13 a, b, & c (see Appendix 2).

The performance of all 30 neural networks using 12 input variables (QRS+dVCG) for the diagnosis of inferior myocardial infarction in experiments 1 & 2 is shown in Table 5-14 (see Appendix 2).

The performance of all 30 neural networks using 9 input variables (QRS) for the diagnosis of inferior myocardial infarction in experiments 1 & 2 is shown in Table 5-15 (see Appendix 2).

For the diagnosis of inferior myocardial infarction, the best performing neural networks were as follows:

QRS + ST-T + dVCG (21v):	21v-20-20-1,
QRS + ST-T (18v):	18v-80-1,
QRS + dVCG (12v):	12v-70-70-70-1,
QRS (9v):	9v-10-10-10-1.

The results of the best performing networks, Glasgow Program and electrocardiographers for the diagnosis of inferior myocardial infarction are shown in Table 5-17.

From Table 5-17, it can be seen that the electrocardiographers' interpretations were inferior to those of the artificial neural networks but superior to the deterministic Glasgow Program in terms of sensitivity in diagnosing inferior myocardial infarction. The specificity of diagnosing inferior myocardial infarction by the electrocardiographers was equivalent to the Glasgow Program and the neural networks using QRS+ST-T+dVCG or QRS+ST-T measurements as input variables.

The use of ST-T measurements as additional input parameters to the neural networks impaired the specificity of the inferior infarction

diagnosis in the left ventricular hypertrophy cases. The Glasgow Program was superior both to the electrocardiographers and the neural networks in terms of inferior infarct specificity in the left ventricular hypertrophy cases.

IMI	GP	Neural Network				H E
		Q+ST+V	Q+ST	Q+V	Q	
Se	65%	87%	87%	84%	80%	72%
N Sp	100%	100%	100%	94%	96%	100%
AI	65%	87%	87%	78%	76%	72%
L Sp	88%	79%	73%	82%	85%	81%
O Sp	95%	90%	88%	89%	91%	91%
O AI	60%	77%	75%	73%	71%	63%

Table 5-17 Result of the comparison among the Glasgow Program (GP), the best performing neural networks, and human experts in the diagnosis of inferior myocardial infarction.(Q= QRS measurements, ST= ST-T measurements, V= derived VCG measurements, HE= human experts)

5.4.3 General Results

There were no significant differences in the performance of the neural networks trained in the two-group situation for the diagnosis of anterior or inferior myocardial infarction if using the same set of input variables and the same number of hidden layers when tested both in the two-and three-group situations.

The results suggested that neural networks are superior to the deterministic criteria for the diagnosis of anterior myocardial infarction

and inferior myocardial infarction in a two-group situation by improving the sensitivity while preserving the specificity in normals.

In the three-group situation, the incorporation of ECG ST-T parameters improved the diagnosis of myocardial infarction but jeopardised the specificity of myocardial infarction diagnosis in left ventricular hypertrophy cases; the addition of derived VCG measurements to the input variables of the neural network using ECG QRS parameters produced better results than the network using ECG QRS parameters only but was not so good as compared to the neural network using the ECG QRS and ST-T measurements together.

5.5 EXPERIMENT 3

TRAINING AND TESTING IN THE THREE-GROUP SITUATION (Myocardial Infarction vs Normal vs Left Ventricular Hypertrophy)

ARTIFICIAL NEURAL NETWORKS WITH ONLY ONE OUTPUT NEURON

The aim of this experiment was to evaluate whether the performance of the neural networks could be improved after adding ECGs from left ventricular hypertrophy cases to the training set used in experiments 1 and 2.

5.5.1 MATERIALS

Anterior Myocardial Infarction

The training set consisted of 80 ECGs recorded from patients with clinically validated anterior myocardial infarction and 80 ECGs recorded from normals (including 30 ECGs showing poor R wave progression over praecordial leads from apparently healthy individuals admitted for the investigation of atypical chest pain but found to have normal coronary arteriograms) and 42 ECGs recorded from patients with clinically documented left ventricular hypertrophy.

The test set consisted of 101 ECGs recorded from patients with clinically documented anterior myocardial infarction, 42 ECGs recorded from the patients with clinically proven left ventricular hypertrophy, and 100 ECGs recorded from normal subjects (Table 5-18).

	Anterior MI	Normal	LVH
Training Set	80	80	42
Test Set	101	100	42

Table 5-18 The composition of the training set and test set ECGs used for the neural network diagnosis of anterior myocardial infarction in experiment 3.

Inferior Myocardial Infarction

The training set consisted of 100 ECGs recorded from patients with clinically proven inferior myocardial infarction, 42 ECGs recorded from patients with clinically documented left ventricular hypertrophy, and 100 ECGs recorded from normals.

	Inferior MI	Normal	LVH
Training Set	100	100	42
Test Set	80	100	42

Table 5-19 The composition of the training set and test set ECGs used in the neural network diagnosis of inferior myocardial infarction in experiment 3.

The test set consisted of 80 ECGs recorded from patients with clinically proven inferior myocardial infarction, 42 ECGs recorded from the patients with clinically validated left ventricular hypertrophy, and 100 ECGs recorded from normal subjects (Table 5-19).

5.5.2 RESULTS OF EXPERIMENT 3

Anterior Myocardial Infarction

The performance of the artificial neural networks using varying designs of input and hidden layers but with only a single output for the diagnosis of anterior myocardial infarction was assessed. Full details of results of all neural networks are given in Tables 5-20 to Table 5-23 (Appendix 2).

The best performing neural networks were as follows:

QRS+ST-T+dVCG (46v):	46v-40-40-1,
QRS+ST-T (42v):	42v-20-1,
QRS+dVCG (25v):	25v-30-30-1,
QRS (21v):	21v-40-40-1.

The best results for the neural network diagnosis of anterior myocardial infarction in experiment 3 are summarised in Table 5-24.

The use of derived VCG measurements in addition to the use of QRS measurements as input variables to the neural networks did not lead to any significant improvement in the sensitivity of diagnosing anterior myocardial infarction if ST-T measurements were already used as inputs to the network. There was a drop in sensitivity after addition of the derived VCG measurements to the input.

The original Glasgow Program performed better than each of the neural networks. The specificity of the neural networks was comparable to that of the deterministic program in the present study, but there was a slight decrease of sensitivity in detecting anterior myocardial infarction.

AMI	GP	Q+ST+V	Q+ST	Q+V	Q
Se	89%	87%	88%	85%	82%
L Sp	86%	86%	84%	86%	86%
N Sp	100%	99%	99%	100%	100%
O Sp	96%	96%	94%	96%	96%
O AI	85%	83%	82%	81%	78%

Table 5-24 Results of the comparison among the Glasgow Program and the best performing neural networks with varying input variables for the diagnosis of anterior myocardial infarction in experiment 3. (Q=QRS measurements, ST=ST-T measurements, V=derived VCG measurements)

Inferior Myocardial Infarction

The performance of the artificial neural networks using varying designs of input variables and hidden layers but with only a single output for the diagnosis of inferior myocardial infarction was assessed. Full details of

results of all neural networks are given in Table 5-25 to Table 5-28 (see Appendix 2).

The artificial neural networks using the same set of input parameters but with different topologies of the hidden layer do not have a significant difference in performance in diagnosing inferior myocardial infarction, i.e. 21v-20-20-1 and 21v-50-1 have no pronounced difference in performance. The results for the best neural network diagnosis of inferior myocardial infarction in experiment 3 are summarised in Table 5-29.

The best performing neural networks were as follows:

QRS+ST-T+dVCG (21v):	21v-20-1,
QRS+ST-T (18v):	18v-10-10-1,
QRS+dVCG (12v):	12v-50-50-1,
QRS (9v):	9v-50-50-1.

IMI	GP	Q+ST+V	Q+ST	Q+V	Q
Se	59%	76%	79%	70%	74%
L Sp	83%	80%	80%	80%	79%
N Sp	100%	99%	99%	96%	99%
O Sp	95%	94%	94%	92%	93%
O AI	54%	70%	73%	62%	67%

Table 5-29 Results of the comparison among the Glasgow Program and the best performing neural networks with varying input variables for the diagnosis of inferior myocardial infarction in experiment 3. (Q=QRS measurements, ST= ST-T measurements, V= derived VCG measurements)

The addition of the derived VCG measurements to the input variables in this series of networks led to a decrease in sensitivity but preservation of

the specificity for the diagnosis of inferior myocardial infarction. The networks using QRS+ST-T measurements as input variables had the best performance in the diagnosis of inferior myocardial infarction. In general, the use of a neural network with this design is better than the Glasgow Program in the diagnosis of inferior myocardial infarction.

5.6 EXPERIMENT 4

TRAINING AND TESTING IN THE THREE-GROUP SITUATION (Myocardial Infarction vs Normal vs Left Ventricular Hypertrophy)
ARTIFICIAL NEURAL NETWORKS WITH THREE OUTPUT NEURONS

The aim of this experiment was to evaluate whether the alteration of the output layer topology of the neural networks could lead to any improvement or not in the performance of diagnosing myocardial infarction.

5.6.1 MATERIALS

Anterior Myocardial Infarction

The training and test sets are the same as those in experiment 3 (Table 5-18).

Inferior Myocardial Infarction

The training and test sets are the same as those in experiment 3 (Table 5-19).

5.6.2 METHODS

Standard statistical formulae were used to calculate the sensitivity and specificity of the diagnosis of anterior and inferior myocardial infarction, and left ventricular hypertrophy as well as specificity in normals.

The definitions and methods of calculations of these terms are shown as follows:

AMI Se=	Sensitivity of anterior myocardial infarction diagnosis in anterior myocardial infarction patients
IMI Se=	Sensitivity of inferior myocardial infarction diagnosis in inferior myocardial infarction patients
LVH Se=	Sensitivity of left ventricular hypertrophy diagnosis in left ventricular hypertrophy patients
AMI Sp=	Specificity of anterior myocardial infarction diagnosis in non-anterior myocardial infarction patients (normal+LVH)
IMI Sp=	Specificity of inferior myocardial infarction diagnosis in non-inferior myocardial infarction patients (normal+LVH)
LVH Sp=	Specificity of left ventricular hypertrophy diagnosis in non-left ventricular hypertrophy patients (normal+MI)
N Sp=	Specificity of myocardial infarction and left ventricular hypertrophy diagnosis in normals.

5.6.3 RESULTS OF EXPERIMENT 4

Anterior Myocardial Infarction

Full details of the results of all neural networks using a varying set of input variables and different designs of hidden layers with three outputs

for the diagnosis of anterior myocardial infarction, left ventricular hypertrophy and normal are shown in Table 5-30 to Table 5-33 (see Appendix 2).

The best performing neural networks were as follows:

QRS+ST-T+dVCG (46v):	46v-20-3,
QRS+ST-T (42v):	42v-50-50-3,
QRS+dVCG (25v):	25v-10-3,
QRS (21v):	21v-20-20-3.

The results of the best neural networks for the diagnosis of anterior myocardial infarction are summarised in Table 5-34. It was demonstrated that the changes in the number of neurons in the hidden layers had no profound effect on the performance of the neural networks for the diagnosis of anterior myocardial infarction.

	GP	Q+ST+V	Q+ST	Q+V	Q
AMI Se	89%	84%	86%	84%	83%
AMI Sp	91%	96%	93%	96%	94%
LVH Se	69%	76%	67%	71%	45%
LVH Sp	86%	88%	92%	85%	87%
N Sp	100%	89%	93%	74%	81%

Table 5-34 Comparison of the results of the best performing neural networks with the Glasgow Program (GP) for the diagnosis of anterior myocardial infarction in experiment 4. (Q= QRS measurements, ST= ST-T measurements, V= derived VCG measurements)

The use of derived VCG measurements as supplementary input variables to the neural networks led to a decrease of specificity in normal and left

ventricular hypertrophy cases (cf. Q versus Q+V and Q+ST versus Q+ST+V), although there was an increase of specificity in the anterior myocardial infarction cases. The Glasgow program was better with respect to the specificity of normals and the sensitivity of anterior myocardial infarction. The artificial neural networks using ST-T measurements as input parameters performed better with respect to sensitivity and specificity in all cases compared to networks not using them (cf. Q+ST versus Q and Q+ST+V versus Q+V).

Inferior Myocardial Infarction

Full details of the results of artificial neural networks using varying sets of input variables and different designs of hidden layers with three outputs for the diagnosis of inferior myocardial infarction, left ventricular hypertrophy and normal are shown in Table 5-35 to Table 5-38 (see Appendix 2).

The best performing neural networks are as follows:

QRS+ST-T+dVCG (21v):	21v-50-3,
QRS+ST-T (18v):	18v-50-50-3,
QRS+dVCG (12v):	12v-10-10-3,
QRS (9v):	9v-30-3.

It was also found that the changes in the number of neurons in the hidden layers had no profound effect on the performance of the neural networks in the diagnosis of inferior myocardial infarction. The results of the best neural networks for the diagnosis of inferior myocardial infarction are summarised in Table 5-39.

There was no improvement in the specificity of inferior myocardial infarction diagnosis in normals using this topology of neural network. Artificial neural networks of this design were superior to the Glasgow Program in terms of sensitivity in diagnosing inferior myocardial infarction, but the Glasgow Program was better than neural networks of this design with respect to specificity in all diagnoses (i.e. there were less false positives in inferior myocardial infarction, left ventricular hypertrophy and normal cases).

	GP	Q+ST+V	Q+ST	Q+V	Q
IMI Se	59%	75%	79%	69%	70%
IMI Sp	95%	94%	95%	94%	94%
LVH Se	69%	67%	71%	64%	36%
LVH Sp	96%	93%	89%	91%	88%
N Sp	100%	94%	90%	98%	84%

Table 5-39 Comparison of the results of the best performing neural networks with the Glasgow Program (GP) for the diagnosis of inferior myocardial infarction in experiment 4. (Q= QRS measurements, ST= ST-T measurements, V= derived VCG measurements)

5.7 DISCUSSION

5.7.1 General Discussion

Certain phenomena can be observed from this study during the training of the neural networks which are also in concordance with the previous study by Dassen et al (Dassen et al, 1993). The performance of the

training improved and the time required for each run increased when the number of the neurons in the hidden layer increased (i.e. the number of iterations needed to reach the fully trained state decreased). However, the performance in the identification of myocardial infarction in the test set ECGs did not improve as the number of neurons in the hidden layers grew, but this did cause slower training and running, and required more computer memory.

5.7.2 The Two-group Situation versus The Three-group Situation in Training and Testing

Two-group situation

From these experiments, it has been demonstrated that for the population studied, in anterior myocardial infarction (Table 5-16), the Glasgow Program (Overall Association Index= 83%) performed similarly to the electrocardiographers (Overall Association Index= 83%). The neural network using QRS and derived VCG measurements (Overall Association Index= 79%) was the best, and the network using QRS measurements only (Overall Association Index= 78%) next, but their specificity for the diagnosis of anterior myocardial infarction in the left ventricular hypertrophy cases was not acceptable. In inferior myocardial infarction (Table 5-17), the Glasgow Program (Overall Association Index= 60%) performed worse than the electrocardiographers (Overall Association Index= 63%), while the neural network using QRS, ST-T and derived VCG measurements (Overall Association Index= 77%) was the best, and the network using QRS only (Association Index= 71%) the worst.

Because in both deterministic logic based ECG interpretation and human expert ECG analysis, left ventricular hypertrophy ECGs may be interpreted as myocardial infarction, left ventricular hypertrophy ECGs were also included in the non-myocardial infarction group in the present study to make the assessment more realistic. With respect to the specificity of diagnosing myocardial infarction in the left ventricular hypertrophy cases, the neural network was found to be inferior to the deterministic Glasgow Program. For example, in the diagnosis of anterior myocardial infarction, the neural network was 50% specific using all the ECG and derived VCG parameters as input variables versus 87% specificity of the Glasgow Program and 95% specificity of the human expert (Table 5-16). With respect to the diagnosis of inferior myocardial infarction in left ventricular hypertrophy cases, the neural network was 79% specific versus 88% specificity of the Glasgow Program and 81% specificity of the human expert (Table 5-17).

Three group situation with the topology of a single output

The neural networks trained in the three-group situation were better in specificity but worse in sensitivity and association indices than the networks trained in the two-group situation for the diagnosis of anterior myocardial infarction (Tables 5-16 & 5-24). Similarly, for the diagnosis of inferior myocardial infarction, the sensitivity and association indices of the neural networks trained in three group situation were worse and the specificity better than those of the networks trained in the two-group situation (Tables 5-17 & 5-29).

The three-group situation with the topology of three output neurons

The neural networks using three output neurons were not superior in the diagnosis of myocardial infarction because of their poor specificity in the normal cases in comparison with the networks using only a single output. Therefore, this design was shown to be unsuitable for the diagnosis of myocardial infarction as demonstrated in the present study (Tables 5-34 & 5-39). Bortolan et al (1991) reported that a set of 39 measurements (37 ECG variables plus age and sex) from each of 2446 ECGs was input to train a neural network. The best network had the topology of only a single hidden layer and five output neurons. On the basis of a test set of 820 ECGs, it was shown to produce a specificity (i.e. sensitivity in normals) of 80.9% which was unacceptable. The present study also confirmed that the use of three output topology for the neural network could jeopardise the specificity in normals.

5.7.3 Comparison with Other Studies

In the two reports from the University of Lund in Sweden, Heden et al (1993) and Reddy et al (1993) have claimed that the improvement gained from adding ST-T measurements and derived VCG parameters to QRS measurements is less for anterior myocardial infarction networks than inferior myocardial infarction networks (Tables 5-40 & 5-41).

Nevertheless, from the results of the current study, the trend of improvements when adding either ST-T or derived VCG measurements to the input variables of the neural networks for the diagnosis of either anterior or inferior myocardial infarction was found to be concordant, (cf Tables 5-16 & 5-17) i.e. the addition of derived VCG parameters to the

QRS measurements improved the association index from 76% to 78% in inferior myocardial infarction and from 87% to 88% in anterior myocardial infarction, but the combination of QRS and ST-T parameters led to a more pronounced improvement from 76% to 87% in inferior myocardial infarction and from 87% to 95% in anterior myocardial infarction. Furthermore, the combination of both QRS & ST-T measurements and the derived VCG parameters resulted in an association index of 87% for inferior myocardial infarction and 96% for anterior myocardial infarction (Tables 5-16 & 5-17). In the Lund study, the detailed performance of the neural networks for the diagnosis of anterior myocardial infarction was not quoted, but the results for their inferior myocardial infarction networks are shown in Table 5-40.

IMI	QRS	QRS+ST	QRS+dV	QRS+dV+ST
Se	68%	78%	75%	84%
Sp	95%	95%	95%	95%
AI	63%	73%	70%	79%

Table 5-40 The performance of the Lund neural network for the diagnosis of inferior myocardial infarction in the two group situation *(Adapted from Heden et al 1993)*.

The study of Reddy et al (1993) (Table 5-41) used the old (1987) Glasgow criteria. However, the new revised Glasgow deterministic criteria (1992) were used in the present study for the comparison with neural networks and electrocardiographers' interpretations. Therefore, it is not possible to draw any further conclusions from this comparison.

The neural networks seemed to perform much better in our study than the Lund study when comparisons of two group performance were made

(Table 5-16: Association Index= 96%, Sensitivity= 97%, Specificity= 99% in the diagnosis of anterior myocardial infarction), yet the different sampling rate of ECG recording (250 samples/second in Lund versus 500 samples/second in Glasgow), different ECG signal processing techniques, different ECG and derived VCG parameters used for the input layers, and different populations of patients studied could easily produce completely different results even when the same design and configuration of neural networks is used. However, the trend when using the neural network approach in the diagnosis of inferior myocardial infarction was similar in both centres.

AMI	Glasgow Criteria	Lund NN (QRS)	Expert
Se	68%	79%	82%
Sp	97%	97%	93%
AI	65%	76%	75%

Table 5-41 Comparisons of the performances of the old Glasgow criteria (1987), Lund neural network, and electrocardiographer in the Lund study for the diagnosis of anterior myocardial infarction *(Adapted from Reddy et al 1993).*

In a recent report, Pahlm et al (1994) claimed that appending derived VCG parameters (X intercept, early rotation, direction of maximum vector, duration of early superior forces and direction of the initial 30 ms vector) to the traditional scalar ECG measurements (e.g. Q duration, Q amplitude and Q/R ratio) could have a significant improvement in diagnostic yield in separating cases of inferior myocardial infarction from the combined group of normals and cases with anterior myocardial infarction. Our data do not consistently confirm this finding. This difference may be due to the different training and testing materials, e.g.

the inclusion of left ventricular hypertrophy cases, as well as the different signal processing technique used. On the other hand, it is also possible that in the current study the ECG variables used had already performed maximally for the diagnosis of myocardial infarction, so that no significant improvement could be shown after adding the derived VCG measurements.

5.7.4 The Addition of Derived VCG Measurements versus ST-T Measurements

The use of QRS measurements alone as the input variables to the neural network had the best specificity for the diagnosis of myocardial infarction in left ventricular hypertrophy cases. However, the performance was worse than the original Glasgow Program. The addition of derived VCG parameters (specificity= 82% in inferior myocardial infarction and 77% in anterior myocardial infarction), ST-T measurements, and both together to the input layer jeopardised the specificity of the neural networks for the left ventricular hypertrophy cases. The addition of ST-T measurements led to greater improvement compared to the addition of derived VCG parameters. In the diagnosis of inferior myocardial infarction, even the worst performing neural network did better than the deterministic Glasgow Program in the two group situation, while in anterior myocardial infarction, the worst two subsets of neural networks were comparable to the Glasgow Program in the two group situation. From the present observation, the addition of the derived VCG measurements made no difference to the performance in the two group situation.

In the present study, it was demonstrated that the addition of derived VCG measurements as input variables to a neural network had a variable effect on performance in the diagnosis of either anterior or inferior myocardial infarction whether or not left ventricular hypertrophy cases were used in the test set.

This may imply that:

(1) The information needed to reach the diagnosis of myocardial infarction can be fully supplied using the carefully selected 12-lead ECG parameters only, so that the addition of the derived VCG measurements does not offer any further benefit.

(2) The derived VCG provides the same diagnostic information as the ECG, i.e. because the derived VCG is synthesized from the 12-lead ECG, it is clearly the same data but with a different form of display.

(3) The software used to extract data from the derived VCG could have been improved or extended, i.e. it was possibly inadequate.

On the other hand, the neural networks using QRS+ST-T measurements as input variables were better in performance than the neural networks using QRS measurements only. Therefore, the neural networks using QRS+ST-T measurements were used in the present study.

5.7.5 Comparison Among Electrocardiographers, Deterministic Logic and Artificial Neural Network in the Diagnosis of Myocardial Infarction

The results for the detection of myocardial infarction performed blindly by the experienced electrocardiographers using only 12-lead ECG print outs without derived VCGs are also shown in Table 5-16 & 5-17. In the diagnosis of clinically proven anterior myocardial infarction cases by the human expert, the sensitivity of myocardial infarction detection was worst (85%) compared to the Glasgow Program (89%) and the neural networks using QRS+ST-T+dVCG, QRS+ST-T, QRS+dVCG, QRS only (97%, 95%, 91%, 88%). However, their specificity of myocardial infarction detection in the normal cases was similar to that of the Glasgow Program and the neural networks using QRS+ST-T (100%), while their specificity of myocardial infarction diagnosis in the left ventricular hypertrophy cases was the best (95%). In the clinically documented inferior myocardial infarction cases, the sensitivity of the ECGs interpreted by the human expert (72%) was intermediate to that of the Glasgow Program (65%) and the neural networks (87%, 87%, 84%, 80%), but the specificity of the expert in the normal cases was the same as Glasgow program and the neural networks using ST-T measurements (100%). The specificity of inferior myocardial infarction detection in the left ventricular hypertrophy cases by the human experts (81%) was worse than that of the Glasgow Program (88%) and neural networks not using ST-T measurements (QRS+dVCG 82%, QRS only 85%) but was superior to the networks using ST-T input variables.

This suggests that the neural networks performed better than either the human experts or the deterministic program in the two-group situation,

but in the three-group situation, the specificity of the neural network for myocardial infarction in left ventricular hypertrophy cases decreased significantly.

Evaluation of the specificity of the diagnosis of myocardial infarction with neural networks using only normal cases is very misleading. e.g. the specificity was 96-100% in normals but 74-88% in left ventricular hypertrophy cases. The major drawback for the neural network in isolation approach for the diagnosis of myocardial infarction as learned from the current study was that the specificity of reporting inferior myocardial infarction dropped from 88% for the Glasgow Program to 73% for the network in left ventricular hypertrophy ECGs using QRS and ST-T parameters as input variables and to 85% for the network using only QRS parameters, although the sensitivity for the diagnosis of inferior myocardial infarction improved from 65% to 87% and 80%, respectively. The use of ECG ST-T measurements as input variables may lead to a 12% (85% to 73%) decrease of the specificity in inferior myocardial infarction detection for the left ventricular hypertrophy cases.

Thereafter, it was decided to include the ECGs from the left ventricular hypertrophy cases in the training set to see if this approach could improve the performance of the neural networks. In anterior myocardial infarction (Table 5-24), the sensitivity of the neural network (88%) in the tested population was similar to that of the Glasgow Program (89%), whereas in inferior myocardial infarction (Table 5-29), the sensitivity of the neural network (79%) was significantly better than the Glasgow Program (59%), and the performance was better (73% versus 54%). These results were calculated on the basis of only a single output with two classifications of myocardial infarction and non-myocardial

infarction, respectively. Thus, the addition of left ventricular hypertrophy cases to the training sets has different effects on the diagnosis inferior and anterior myocardial infarction.

5.7.6 Limitations of the Present Study

In the current study, the ECG data base was assembled on clinically validated normal individuals and patients suffering from myocardial infarction and left ventricular hypertrophy. The clinical diagnoses served as a diagnostic "Gold Standard".

The inherent limitation of this gold standard is that the normal individuals may not all have been free of cardiac disease, since the normal resting 12-lead ECG could not be used to differentiate among what was normal, clinically silent ischaemia, or silent myocardial infarction. It was demonstrated in the Framingham study by Kannel & Abbott (1984) that 25% of all myocardial infarctions are clinically silent. Furthermore, some so called "normal" PRWP individuals who have normal large epicardial coronary arteries demonstrated by coronary arteriography may have unrecognised microvascular (Syndrome X) or myocardial disease. There are also problems with the correct identification of ventricular hypertrophy and myocardial infarction in the living data base used in the present study. The presence of angiographic evidence of ventricular wall motion abnormalities, e.g. hopokinesia, akinesia, or dyskinesia, does not necessarily imply myocardial infarction. Hibernating and/or stunned myocardium can also present with hypoknesia or akinesia (Rahimtoola, 1989).

The definition of ventricular hypertrophy infers an increase in the size of myocardial cells, but does not necessarily imply an increase in ventricular mass (Laks, Morady & Swan, 1969). Ventricular mass in fact is a variable mix of hypertrophied myocardial cells, oedema, interstitial infiltrates, and fibrosis.

The same caveat applies to the diagnosis of myocardial infarction. The definite pathologic gold standard of myocardial infarction is the histologic evidence of myocardial cell necrosis, myocardial fibrosis, or both, not necessarily wall motion abnormalities on the left ventriculogram, nuclear scan, or echocardiogram. Wall motion abnormalities can also be present in the chronic ischaemic cardiac situations like myocardial stunning and hibernating myocardium. However, they do not always indicate the presence of a myocardial scar. Nevertheless, the data base was collected as objectively as possible.

Finally, the alteration of the number of output neurons from one to three did not lead to any improvement in the performance of the neural networks as compared to the original deterministic logic in both anterior and inferior myocardial infarction (Table 5-34 & 5-39). It was therefore thought that the implantation of the neural networks at a specific point in the deterministic logic of the Glasgow Program might preserve the advantage of higher myocardial infarction diagnostic specificity of the deterministic logic in the left ventricular hypertrophy cases and also the gain in sensitivity offered by the neural networks. This approach will be discussed in the following chapter.

5.8 CONCLUSIONS

Several conclusions can be drawn from the present study.

(1) In the diagnosis of anterior myocardial infarction, the neural networks trained in the three-goup situation but with only a single output neuron were superior both to the networks trained in the two-group situation and the networks trained in the three-group situation but with three output neurons. On the other hand, in the diagnosis of inferior myocardial infarction, the neural networks trained in the two-group situation were superior to the networks trained in the three-group situation.

(2) The use of three output topology was not helpful in the present study especially with respect to the specificity of normals, confirming Bortolan's report (1991) that the use of 5 outputs led to a lower normal specificity of only 80.9% which is completely unacceptable.

(3) In the most realistic test situation of having normal+myocardial infarction+left ventricular hypertrophy, the best results were obtained from 42v-20-1 in the diagnosis of anterior myocardial infarction and 18v-80-1 in the diagnosis of inferior myocardial infarction, i.e. networks with a single hidden layer.

(4) Neural networks can significantly improve the diagnosis of both anterior and inferior myocardial infarction in the two-group situation, but less so in anterior myocardial infarction.

(5) The neural network approach is better than human experts only in the two-group situation for the computer assisted ECG diagnosis of myocardial infarction.

(6) The use of ST-T measurements in addition to QRS data as input parameters to the neural networks can improve sensitivity as well as specificity in the detection of myocardial infarction. On the other hand, the addition of derived VCG parameters has no significant benefit in the detection of myocardial infarction if ST-T measurements have already been used. Finally, the addition of derived VCG measurements can be beneficial in the diagnosis of myocardial infarction if only 12-lead ECG QRS parameters are already in use.

(7) Finally, the gold standard (clinical diagnosis) used in the present study is the inherent limitation of the performance. There are clinically well documented myocardial infarction and left ventricular hypertrophy cases with normal ECGs included in this study. This is also an inherent limitation of the electrocardiographic classification versus the clinical diagnosis.

CHAPTER 6

SOFTWARE BASED NEURAL NETWORKS IMPLANTED IN DETERMINISTIC LOGIC FOR THE COMPUTER ASSISTED ECG DIAGNOSIS OF MYOCARDIAL INFARCTION

6.1 INTRODUCTION

It was found from the previous experiments that the neural network approach had a higher sensitivity whereas the deterministic approach had the better specificity in diagnosing myocardial infarction. In particular, the Glasgow Program was superior to the neural network approach in the specificity of diagnosing myocardial infarction in left ventricular hypertrophy cases. It was also learned from the experience gained using neural networks for the diagnosis of atrial fibrillation, that although there was no benefit in combining the neural network output and the deterministic logic interpretation for the detection of atrial fibrillation, it was still ultimately possible to improve the performance of the computer assisted ECG interpretation program. Therefore it was thought that the incorporation of the best neural networks into the existing deterministic program with possible modification of the existing logic might have some advantage over using either alone for the computer assisted ECG diagnosis of myocardial infarction. For this reason, two further experiments were undertaken.

6.2 EXPERIMENT 5

GLASGOW PROGRAM WITH NEURAL NETWORK IMPLANTATION

Aim of the study

The purpose of the present experiment was to evaluate whether the incorporation of a neural network into the deterministic program could improve the performance of the program in the diagnosis of myocardial infarction.

6.2.1 ANTERIOR MYOCARDIAL INFARCTION

6.2.1.1 MATERIALS

Full clinical details of the training and test sets ECGs used in experiment 5 are described in Chapter 5 section 5.2 and Table 5-18. The composition of the test set is shown in Table 6-1.

	AMI	LVH	Normal
Test Set	101	42	100

Table 6-1 Composition of the test set ECGs used in experiment 5 for the diagnosis of anterior myocardial infarction.

6.2.1.2 METHODS

The best performing neural networks, one for the neural network using QRS+ST-T measurements as inputs and one for the neural network using QRS measurements only, were selected from experiment 3 in Chapter 5. The best performing neural network for each group was as follows:

QRS+ST-T measurements: 42v-20-1,
QRS measurements: 21v-40-40-1.

The results of Table 5-24 indicate that the network using QRS+ST-T measurements should be adopted for further study.

6.2.1.3 METHOD OF IMPLANTATION

The technique adopted for implantation was as follows. In the original deterministic Glasgow Program, the logic (see Appendix 3, Table 6-2) for the diagnosis of definite Q waves (VQ1-VQ4 in Table 6-2) in the anteroseptal (V2,V3,V4) leads was replaced by a single neural network, while other deterministic logic e.g. for the detection of low R waves, was left unaltered in the program.

The revised logic functioned in the following way for anterior myocardial infarction:

(1) Definite praecordial Q criteria were replaced by an artificial neural network.
(2) If the output from the neural network is true (>0.5)
then if there is a Q wave in V3 or V4
set reason statement "Q waves in anterior leads" true
else
set reason statement "low R waves in anterior leads" true.

(3) If the output from the neural network is false, other criteria for anterior myocardial infarction are checked, e.g. reversed R waves. Set appropriate reason statement if necessary.

(4) Continue with remainder of the logic.

6.2.1.4 RESULTS

The results of incorporating the best neural network (42v-20-1) into the Glasgow Program for the diagnosis of anterior myocardial infarction (AMI) on the test set are shown in Table 6-3.

AMI	Se	N Sp	L Sp	O Sp	AI
GP	89%	100%	86%	96%	85%
NN	88%	99%	84%	94%	82%
NN+GP	93%	99%	74%	92%	85%

Table 6-3 Results after Implementation of the best performing neural network (42v-20-1) into the Glasgow Program for the diagnosis of anterior myocardial infarction on the test set (AMI= anterior myocardial infarction, L=LVH, N=normal, O=overall, GP= Glasgow Program, NN= neural network, NN + GP= neural network implanted in the GP)

The replacement of the definitive Q deterministic criteria by a neural network enhanced the sensitivity of diagnosing anterior myocardial infarction from the 89% of the original Glasgow Program to 93% (Table 6-3) with a decrease of overall specificity in normals and left ventricular hypertrophy cases from the 96% of the original Glasgow Program to 92% but the same overall performance was preserved (Association Index = 85%). Compared to the use of neural network in isolation, there was a 5% improvement in the sensitivity of detecting anterior myocardial

infarction but a 2% decrease in overall specificity in normals and left ventricular hypertrophy cases. In particular, there was a 12% and 10% decrease of specificity in left ventricular hypertrophy cases compared to the original Glasgow Program and neural network in isolation, respectively. In contrast, there was no change of specificity in normals. Therefore, the source of the decline in overall specificity came from the decrease in left ventricular hypertrophy cases (see Table 6-3).

6.2.2 INFERIOR MYOCARDIAL INFARCTION

6.2.2.1 MATERIALS

Full clinical details of the training and test sets ECGs used in experiment 5 for diagnosing inferior myocardial infarction are described in Chapter 5 section 5.2 and Table 5-4. The composition of the test set is shown in Table 6-4.

	IMI	LVH	Normal
Test Set	80	42	100

Table 6-4 Composition of the test set ECGs used in experiment 5 for the diagnosis of inferior myocardial infarction.

6.2.2.2 METHODS

It was clearly demonstrated in the previous experiments in chapter 5 that the neural networks trained in the three-group situation were worse in performance than the neural networks trained in the two-group situation (see Tables 5-17 & 5-29).

The best performing neural networks, one for the neural network using QRS+ST-T measurements as inputs and one for the neural network using QRS measurements only, were therefore selected from experiment 2 in Chapter 5. The best performing neural network for each group was as follows:

QRS+ST-T measurements: 18v-80-1,
QRS measurements: 9v-10-10-10-1.

These two best performing networks were then incorporated into the deterministic Glasgow Program one at a time for further study. The performance of this new program was then evaluated on the test set.

6.2.2.3 METHOD OF IMPLANTATION

In the original deterministic program, the diagnosis of myocardial infarction is carried out after left bundle branch block and Wolff-Parkinson-White pattern are excluded. The definite Q wave criteria are classified by their degree of significance from Q1 to Q4 (see Appendix 3, Table 6-5). The evaluation of the criteria Q1 to Q4 are carried out in the inferior leads II, III, and aVF in the original Glasgow Program.

In this experiment, the logic of definite Q wave criteria (Q1-Q4) was replaced by the best performing neural network of group 1 (18v-80-1) or group 2 (9v-10-10-10-1) one at a time. If the neural network did not detect inferior myocardial infarction, then other "Non-Q criteria" (see Appendix 3, Table 6-6) were tested in the inferior leads II, III and aVF.

6.2.2.4 RESULTS

The results of incorporating the best neural network of each group one at a time into the Glasgow Program for the diagnosis of inferior myocardial infarction on the test set are shown in Table 6-7.

IMI	Se	NSp	LSp	OSp	AI
GP	59%	100%	93%	98%	57%
NN(Q+ST)	84%	100%	74%	92%	76%
NN(Q)	74%	96%	88%	94%	68%
GP+NN(Q+ST)	86%	100%	74%	92%	78%
GP+NN(Q)	80%	96%	86%	93%	73%

Table 6-7 Results on the test set for the diagnosis of inferior myocardial infarction following the implantation of the best performing neural network of each category one at a time (Q= QRS measurements, ST= ST-T measurements) into the Glasgow Program. (IMI= inferior myocardial infarction, GP= Glasgow Program, NN= neural network, NN + GP= neural network implanted GP, Se= sensitivity, Sp= specificity, N= normal, LSp= IMI Sp in left ventricular hypertrophy cases, OSp= Overall Sp, AI= Association Index)

In the diagnosis of inferior myocardial infarction (see Table 6-7), the implementation of the neural network into the Glasgow Program led to a dramatic 27% increase (for the neural network using QRS+ST-T measurements as input variables) and 21% increase (for the neural network using QRS measurements only) in the sensitivity of inferior myocardial infarction detection and a significant increase in the overall performance of the program (21% and 16% increase of the Association Index for the neural networks using QRS+ST-T and QRS measurements, respectively).

It can also be seen that the neural network using QRS+ST-T measurements (Association Index= 78%) as input variables was better than the neural network using QRS measurements only (Association Index= 73%). On the other hand, the neural network using only QRS measurements had a higher specificity of diagnosing inferior myocardial infarction in the left ventricular hypertrophy cases but also had a lower specificity of 96% in normals as compared to 100% specificity in either the Glasgow Program or the network using QRS+ST-T measurements as input variables.

The implementation of the neural network using QRS+ST-T measurements as input variables into the Glasgow Program had the advantage of enhanced sensitivity in reporting inferior myocardial infarction while preserving the specificity in both normal and left ventricular hypertrophy cases compared to the use of a neural network in isolation (Table 6-7).

6.3 EXPERIMENT 6

NEURAL NETWORK IMPLANTED IN MODIFIED GLASGOW PROGRAM

The result of Table 6-3 showed that the specificity of reporting anterior myocardial infarction in left ventricular hypertrophy cases needed to be improved. After careful review of the ECGs, it was thought that modification of the existing deterministic logic might improve the performance of the Glasgow Program with the neural network implanted. Therefore, the aim of this next experiment was to evaluate

the performance of the neural network implanted into the deterministic program together with some modification of the logic for the diagnosis of myocardial infarction.

6.3.1 ANTERIOR MYOCARDIAL INFARCTION

6.3.1.1 MATERIALS

The same test set as that of experiment 5 was used for the current study.

6.3.1.2 METHODS

Training set ECGs were carefully reviewed by two experienced electrocardiographers and consensus was reached on how to modify certain parts of the deterministic logic which still contained the same neural network (42v-20-1) for diagnosing anterior myocardial infarction. Whereas previously the neural network replaced Q wave criteria, in this experiment the network was combined with the original Q criteria in "OR" logical fashion. It was thought this might enhance sensitivity. On the other hand, to improve specificity, some additional logic checking on R wave amplitude in the anterior leads was added.

The method of using the neural network, together with the modification of the logic, was as follows:

1. If either
   the deterministic logic definite Q criteria = true
   or
   neural network output > 0.5 (= true)
   then
   set reason statement "Q waves in anterior leads"
   = true.

2. If atypical Q criteria = true
and definite Q criteria = false
then
set reason statement "atypical Q"= true.

4. If low R wave criteria = true
and definite Q criteria = false
and atypical Q criteria = false
then
set reason statement "low R wave" = true

5. If neural network = true
and low R wave criteria = false
and no Q in V2 or V3
then
set reason statement "low R wave"= true
set reason statement "Q waves in anterior leads"
= false

6. If low R wave criteria = true
and reversed R wave criteria = false
and{(RV2 > 0.2 mV and QV2 = 0) or
(RV3 >0.2 mV and QV3 = 0)}
then
set reason statement "low R wave" = false.

7. If neural network = true
and reversed R wave criteria = false
and {RV2 > 0.2 mV and QV2=0 and
RV3 >0.2 mV and QV3=0}
and {|Q| > 0.15 mV or |Q/R| > 1/3
in any 2 of I, aVL, V5, V6}
then
set "Q waves in lateral leads" = true

6.3.1.3 RESULTS

Table 6-8 shows that after the modification, there was a decrease in sensitivity for the detection of anterior myocardial infarction from 93% to 90% while the specificity of 99% was preserved in normal cases and a

small gain of 2% specificity in left ventricular hypertrophy cases (76%) was obtained in comparison with experiment 5. Although the performance (Association Index= 82%) decreased compared to (i) the use of deterministic logic alone (Association Index= 85%) and (ii) the use of neural network implanted Glasgow Program before the modification of the logic (Association Index= 85%), the overall specificity (92%) is similar to that before the modification (92%) but was inferior to the original deterministic logic (96%). This version of the neural network implanted in the deterministic program together with the first modification of the logic was subsequently used in the evaluation of the second test set.

AMI	Se	N Sp	LSp	O Sp	AI
GP	89%	100%	86%	96%	85%
NN+GP	93%	99%	74%	92%	85%
NN+GP+M	90%	99%	76%	92%	82%

Table 6-8 Results of neural network implanted into the modified deterministic program for the diagnosis of anterior myocardial infarction evaluated on the test set. (GP= Glasgow Program, NN= neural network, M=modifications)

6.3.2 Inferior myocardial infarction

6.3.2.1 MATERIALS

The same test set as that of experiment 5 was used in the current study.

6.3.2.2 METHODS

The result in Table 6-7 has shown that the specificity of inferior myocardial infarction in the left ventricular hypertrophy cases is 74%, a figure which was not acceptable. Therefore, after careful review of ECGs

from the training set, it was decided to add the following modified deterministic criteria for a small inferior Q wave in order to improve the specificity of inferior myocardial infarction (IMI) diagnosis in the left ventricular hypertrophy cases. The implementation was otherwise unchanged, i.e. a neural network replaced Q wave criteria while "non-Q criteria" remained.

Criteria

If the neural network reports **inferior myocardial infarction**
and
|QaVFamp| < 0.075mV and RaVFamp > 0.15mV
then
set **inferior myocardial infarction as false.**

The detailed method of implantation was as follows:

1. If the neural network output > 0.5 (= true),
   then
   set IMI = true.
2. If IMI= true
   and | QaVFamp | < 0.075mV and RaVFamp > 0.15mV
   then
   set IMI = false.
3. If output from the neural network = true
   and QaVFamp= 0,
   then
   set reason statement "definite Q" = false
   and reason statement "Q equivalent" =true.
4. If the neural network output = false,
   then
   "non-Q criteria" are used.

6.3.2.3 RESULTS

The results of incorporating the neural network into the Glasgow Program before and after the addition of the small inferior Q criterion for the diagnosis of inferior myocardial infarction are shown in Table 6-9.

IMI	Se	NSp	LSp	OSp	AI
GP	59%	100%	93%	98%	57%
GP+NN(Q+ST)	86%	100%	74%	92%	78%
GP+NN(Q)	80%	96%	86%	93%	73%
GP+M+NN(Q+ST)	79%	100%	88%	96%	75%
GP+M+NN(Q)	76%	97%	88%	94%	70%

Table 6-9 Results after the implantation of the neural networks (NN) (Q+ST=QRS+ST-T, Q=QRS) into the deterministic Glasgow Program (GP) together with the addition of the modified small inferior Q criterion (M) for the diagnosis of inferior myocardial infarction (IMI).

After the addition of the modified small inferior Q criterion to the neural network implanted Glasgow Program, it can be seen from Table 6-9 that there was a 14% increase (from 74% to 88%) in specificity for reporting inferior myocardial infarction in left ventricular hypertrophy cases in the neural network using QRS+ST-T measurements and a 2% increase (from 86% to 88%) in specificity in the neural network using QRS parameters only. There were corresponding increases in overall specificity. The improvement in specificity was, as expected, offset by a drop in sensitivity of 7% and 4%, respectively, still giving an increase of 20% and 17% compared to the original program.

As a consequence, this led to corresponding small decreases of 2% and 4% respectively of the overall specificity as compared to the original deterministic logic. However, the performance was significantly better than that of the original deterministic program (i.e. the Association Index improved from 57% to 75%).

6.3.2.4 FURTHER MODIFICATION

From Table 6-9, it was demonstrated that the Glasgow Program implanted with the neural network using QRS+ST-T measurements (18v-80-1) was superior in performance to the program implanted with the neural network using QRS parameters only (9v-10-10-10-1). Therefore, only the former was used in further assessment.

Despite the improvement seen in Table 6-9, it was felt that it might still be possible to achieve further enhancements. Therefore, ECGs from the training set were again reviewed and an additional Q/R ratio criterion in lead aVF was then added in the hope of reducing the false positive diagnosis of inferior myocardial infarction in the left ventricular hypertrophy cases and normal controls.

The simplified procedure of implementation of the neural network together with the modification of the logic for the diagnosis of inferior myocardial infarction is shown in Fig.6-1.

The method (M1) of further modification to the section 6.3.2.2 step 2 was as follows:

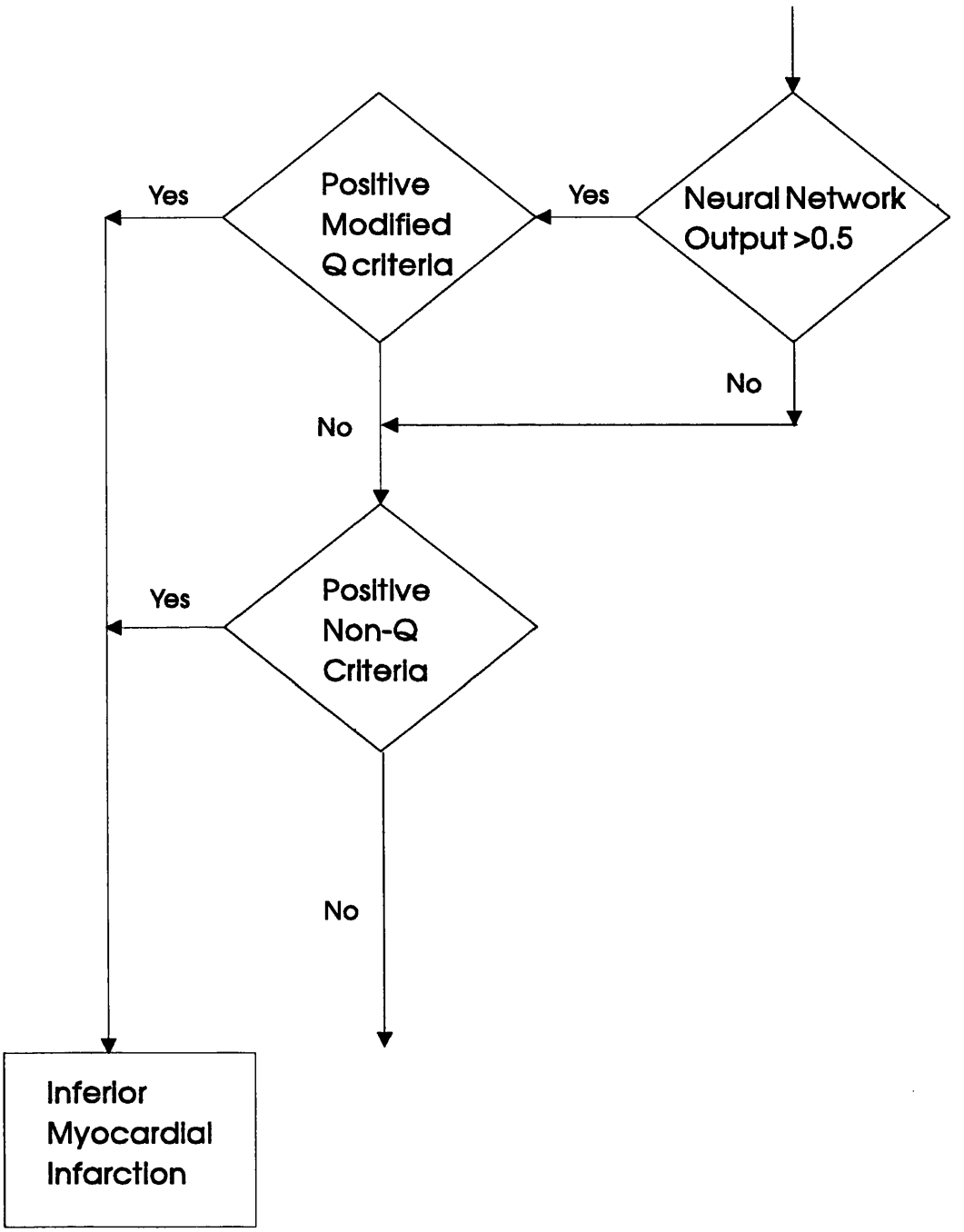


Fig. 6-1 Flow Chart for the Implantation of a Neural Network into the Deterministic Logic for the diagnosis of Inferior Myocardial Infarction.

If IMI=true
and
{ | Q | aVF < 0.075mV & R aVF > 0.075mV}
or
{ | Q/R | aVF < 17.5% & RaVF > 0
& | Q | aVF > 0.05mV & QaVF duration < 30ms}
then
set IMI = false.

Results

The result of the further modification involving lead aVF for the diagnosis of inferior myocardial infarction is shown in Table 6-10. It can be seen that there was a minor decrease (1%) in inferior myocardial infarction detection while the specificity of the inferior myocardial infarction diagnosis was preserved in both normal controls and left ventricular hypertrophy cases. In other words, there was no apparent gain from this further modification.

IMI	IMI Se	N Sp	L Sp	O Sp	AI
M1	78%	100%	88%	96%	74%

Table 6-10 Results of the neural network (QRS+ST-T) implanted in the modified deterministic program (M1) for the diagnosis of inferior myocardial infarction on the test set.

6.4 FINAL ASSESSMENTS

6.4.1 THE SECOND TEST SET

Because some modifications had been based on assessing training sets, it was decided to have a further evaluation on a completely new locally acquired test set.

6.4.1.1 MATERIALS & METHODS

The second test set consisted of 52 ECGs recorded from patients with clinically documented inferior myocardial infarction, 74 ECGs recorded from cases of clinically proven anterior myocardial infarction, 230 ECGs recorded from normals, and 60 ECGs recorded from patients with clinically validated left ventricular hypertrophy (Table 6-11). All the ECGs from the patients and normals were recruited and clinically validated using methods described previously in Chapter 5, section 5.2.

Both the best anterior and inferior myocardial infarction neural networks (42v-20-1 and 18v-80-1, respectively) together with small inferior Q and low praecordial R modifications were implanted simultaneously into the existing deterministic program designated as M1, and evaluated on the second test set ECGs as shown in Table 6-11.

	AMI	IMI	LVH	Normal
Final Test Set	74	52	60	230

Table 6-11 Composition of ECGs used in the second test set for the modified neural network implanted deterministic program in the diagnosis of myocardial infarction. (AMI= anterior myocardial infarction, IMI= inferior myocardial infarction, LVH= left ventricular hypertrophy.)

Separately, another method (M2) of neural network incorporation into the deterministic logic together with modifications was introduced as follows:

Anterior Myocardial Infarction:

(1) If the existing deterministic criteria for anterior, anteroseptal or septal MI are met, then ignore the neural network result. If not, consider the neural network output as follows:

(2) If the neural network output is true, then
if there are Q waves in V2,V3 and V4,
set "Q waves in anterior lead" true,
else
if there are no Q waves in V2, V3, and V4,
set reason statement "low R wave" true in V3,
else
if |QV4| > 0.1 mV or |QV3| > 0.1 mV
set atypical Q = true in V3
else
set neural network output = false.

(3) If "low R wave" reason statement is true from (2)
and {(RV2>0.2 mV and QV2=0)
or (RV3 > 0.2 mV and QV3=0)} is true
then
set "low R wave" reason statement = false.

(4) Continue with existing deterministic logic to determine age of infarct.

Inferior Myocardial Infarction:

(1) If the output of the neural network is true
and QaVF > 0 mV
and |Q/R| aVF > 17.5%
and age > 35 years
then
set probable IMI true
else set IMI false.

(2) Run the deterministic logic and obtain the results, i.e. definite, probable, possible IMI or no IMI.

(3) Merge results of both deterministic logic and neural network by selecting the highest probability of IMI classification.

(4) Use the remainder of existing logic to determine age of MI if IMI = true.

6.4.1.2 RESULTS

The results for both anterior and inferior myocardial infarction using M1 and M2 separately on the second test set ECGs are shown in Table 6-12.

	GP	GP+NN+M1	GP+NN+M2
AMI Se	76%	78%	78%
IMI Se	69%	85%	88%
N Sp	100%	100%	100%
L Sp	93%	80%	85%
O Sp	99%	92%	97%

Table 6-12 Results of comparison on the second test set between the original Glasgow Program (GP) and the anterior (42v-20-1) & inferior (18v-80-1) myocardial infarction neural networks implanted in the deterministic program with modifications M1 or modifications M2.

From Table 6-12, it can be seen that the modified neural network implanted program had a 16% to 19% improvement in sensitivity of diagnosing inferior myocardial infarction and a 2% improvement in sensitivity of diagnosing anterior myocardial infarction. There was a 7% and 2% decrease in overall specificity for M1 and M2, respectively. This decrease of specificity was caused by the false positive diagnosis of myocardial infarction in the left ventricular hypertrophy cases of this test set. There was no false positive diagnosis of myocardial infarction in the normal controls. There was also an 5% improvement in the specificity of

myocardial infarction diagnosis in left ventricular hypertrophy cases by adopting the second modification as compared to the use of the first modification.

In general, the neural network implanted deterministic program was better than the deterministic program with respect to sensitivity in myocardial infarction detection but worse with respect to specificity in left ventricular hypertrophy cases.

6.4.2 TEST ON COMMON STANDARDS FOR QUANTITATIVE ELECTROCARDIOGRAPHY (CSE) DATABASE

6.4.2.1 MATERIALS & METHODS

The main aim of the Common Standards for Quantitative Electrocardiography (CSE) project was the development of standardised program testing methods based upon broad international expert consensus. This project also resulted in the establishment of ECG databases for testing measurement and diagnostic accuracy of computer-assisted ECG interpretation programs (Willems et al, 1990).

The case selection of the CSE database was based upon ECG independent clinical information (Willems et al, 1991). One of the most important aspects of this database is that the classification of individual cases has NEVER been disclosed. It is therefore a truly independent test set on which to assess any new developments. It consists of 1220 ECGs from adult Caucasians, 831 males and 389 females.

There were 382 normal individuals consisting of 286 ambulatory normals with normal medical history, physical examination and Chest X-ray, some of whom had normal echocardiograms. In addition, there were 96 catheterised normals with normal coronary arteriograms admitted for the investigation of atypical chest pain or ECG ST-T abnormalities at rest or after exercise.

There were 183 ECGs from clinically validated left ventricular hypertrophy cases. The diagnosis was based on the findings at cardiac catheterisation and/or echocardiographic examination. The group included patients suffering from longstanding hypertension or left sided valvular heart disease, either congenital or acquired. Left ventricular hypertrophy was diagnosed primarily on left ventricular pressure (peak aortic ventricular gradient no less than 50 mmHg) or volume overload in 115 cases. Echocardiograms were recorded in 85% of these cases. In the remaining 68 cases of hypertension, there was M-mode echocardiographic evidence of increased left mass index (LVMI). The criteria used were LVMI > 110 g/m^2 for females, and LVMI > 134 g/m^2 for males.

The ECGs collected from myocardial infarction patients consisted of 170 from patients with anterior myocardial infarction, 273 from inferior myocardial infarction patients and 73 cases of combined infarction. Akinesia or dyskinesia in 7 different segments of the left ventriculogram was used in the classification of infarct locations. Cases with hypokinesia in one or more wall segments together with an occluded or recanalised supplying vessel and typical history of infarction were also included. All patients had their left ventriculograms recorded in two orthogonal projections.

Cases with akinesia or dyskinesia in the anterolateral, anterobasal or septal region were classified as anterior infarction. Cases with akinesia or dyskinesia in the inferobasal, inferolateral, or true posterior segment were coded as inferior infarction. Patients meeting criteria both for anterior and inferior infarction or isolated apical infarction were classified as combined infarction. There were also 31 cases in which both infarction and hypertrophy coexisted.

All conventional 12 ECG leads and three orthogonal X, Y, Z leads were recorded simultaneously with a sampling rate of 500 samples/second on digital tape. The data were collected from 5 European centres (Dublin, Glasgow, Leiden, Leuven and Louvain). A review board comprised three cardiologists who checked the clinical information using a consensus approach for the case selection. Many cases initially accepted from each centre were ultimately not included, particularly normals, to avoid bias in case selection.

The 1220 ECGs of the CSE database were interpreted using the final version of the modified neural network implanted Glasgow Program. The interpretations were stored in a file and retrieved for copying to a tape with the appropriate format. Subsequently, the tape was forwarded to Leuven, Belgium for assessment. There, standard statistical formulae were used to calculate the sensitivity and specificity of the artificial neural network implanted Glasgow Program with modifications (M2) evaluated on the CSE ECGs.

The European Conformance Testing Services for Computerised Electrocardiography (CTS ECG, 1989-1992) have proposed the establishment of two pilot test centres for comprehensive testing of

computer electrocardiographs based upon harmonised international standards (Zywietz & Willems, 1993). CTS ECG recommended the minimum accuracy limits for diagnostic interpretation as shown in Table 6-13.

Category	Sensitivity	PPV
Normal	85%	60%
LVH	50%	60%
RVH	45%	60%
Anterior MI	65%	75%
Inferior MI	60%	70%
Mixed Infarction	65%	80%

Table 6-13 Recommended minimum accuracy limits for diagnostic interpretation. PPV= positive predictive value (*adapted from Zywietz & Willems: 1993 European Conformance Testing Services*).

6.4.2.2 RESULTS

The results of evaluating the M2 program on the 1220 CSE ECGs in late August 1994 are shown in Table 6-14. It can be seen from Table 6-14 that, compared to the 1992 version of the Glasgow deterministic program, which was also assessed in Leuven, the neural network implanted deterministic program together with the modification of the logic had a 10.4% and 0.3% improvement in the sensitivity of diagnosing inferior and anterior myocardial infarction, respectively, while the total accuracy of the new program also increased significantly by 2.6%. There was also a 1% gain in the sensitivity of detecting mixed

infarction. On the other hand, there was a decrease of 1.6% in the specificity of normals (=sensitivity of normals in the CSE study).

Sensitivity	1990 CP	1990 GP	1992 GP	1994 MGP
Normal	96.7%	94.0%	96.9%	95.3%
LVH	67.9%	54.1%	50.5%	53.3%
RVH	40.6%	46.4%	45.5%	39.1%
BVH	47.3%	30.2%	30.2%	38.7%
AMI	79.6%	77.9%	72.6%	72.9%
IMI	68.8%	58.8%	65.8%	76.2%
MIX	66.2%	61.0%	68.2%	69.2%
TA	76.3%	69.7%	71.4%	74.0%

Table 6-14 Comparison of the sensitivity in the interpretation of the CSE database (1220 ECGs) among the 1990 version of 9 combined computer assisted ECG interpretation programs (1990 CP) (including statistical and deterministic programs), two earlier versions of the Glasgow deterministic program (1990 GP, 1992 GP) and the neural network implanted deterministic logic version (1994 MGP). The figure are percentage of correct diagnosis versus true clinical diagnosis. (LVH: left ventricular hypertrophy, RVH: right ventricular hypertrophy, BVH: biventricular hypertrophy, : AMI: anterior myocardial infarction, IMI: inferior myocardial infarction, MIX: combined infarction, VH+MI: hypertrophy and myocardial infarction, TA: total accuracy.)

The results of the positive predictive value of the modified neural network implanted Glasgow Program are shown in Table 6-15. In comparison with the recommended minimum accuracy limits for diagnostic interpretation from the European Conformance Testing Services for Computerised Electrocardiography (Table 6-13), the positive

predictive values and sensitivity of all classifications were all well above the recommended level, except for the sensitivity of diagnosing right ventricular hypertrophy.

As shown in Table 6-14, in comparison with the 1990 combined program (Willems et al, 1991), i.e. the combination of all the 9 computer programs tested (including programs using deterministic logic or statistical techniques), the new program was 7.4% and 3% more sensitive in detecting inferior and mixed myocardial infarction, respectively, although it was 6.7% less sensitive in diagnosing anterior myocardial infarction.

Category	Sensitivity	PPV
Normal	95.3%	97.9%
LVH	53.3%	89.6%
RVH	39.1%	78.3%
Anterior MI	72.9%	96.6%
Inferior MI	76.2%	93.1%
Mixed Infarction	69.2%	96.2%

Table 6-15 Results of the neural networks implanted Glasgow Program together with modifications tested on the CSE database. (PPV= positive predictive value)

It was also of interest to compare the specificities of diagnosing anterior and inferior myocardial infarction in the left ventricular hypertrophy cases using the different programs. The results are shown in Table 6-16. The new program was 1% more specific in diagnosing anterior myocardial infarction and 2.5% less specific in diagnosing inferior myocardial infarction in left ventricular hypertrophy cases compared to

the 1992 version of the deterministic program. With respect to the diagnosis of anterior myocardial infarction, there was 0.5% increase of specificity in normals but a 1.1% decrease in specificity of reporting inferior myocardial infarction in normals compared to the 1992 version of the Glasgow deterministic program.

	1990	1992	1994
AMI LSp	95.1%	95.4%	96.4%
AMI NSp	96.3%	98.2%	98.7%
IMI LSp	95.1%	94.0%	91.5%
IMI NSp	99.7%	99.5%	98.4%

Table 6-16 Results of the comparisons of the specificities of diagnosing anterior and inferior myocardial infarction in the left ventricular hypertrophy cases (LSp) and normals (NSp) using the different versions of the Glasgow Program (1990, 1992 and 1994).

6.5 TEST ON ECGs FROM CHINESE NORMALS

The final modified neural network implanted Glasgow Program (M2) was also assessed on the 503 Chinese ECGs described in detail in Chapter 3.

The specificity of the diagnosis of anterior and inferior myocardial infarction was 99% in both cases. This compares with 98.7% and 98.4% respectively for the 382 Caucasian ECGs in the CSE study analysed by M2.

6.6 DISCUSSION

6.6.1 Value of Implanting a Neural Network into the Deterministic Glasgow Program

As far as is known, this is the first report concerning the implantation of neural networks into a deterministic program together with modification of the existing logic for the computer-assisted ECG diagnosis of myocardial infarction.

There was a clear benefit from using this approach in the diagnosis of inferior myocardial infarction, whereas no benefit was obtained in the diagnosis of anterior myocardial infarction.

The major drawback of incorporating a network was shown to be the sacrifice of the high specificity of reporting inferior myocardial infarction in left ventricular hypertrophy cases using the original deterministic program. Thus, additional modifications were subsequently introduced in order to improve the specificity of reporting myocardial infarction in left ventricular hypertrophy cases.

The reason why a neural network can not perform better than the deterministic program with respect to the specificity of diagnosing myocardial infarction in left ventricular hypertrophy cases may be ascribed to the fact that left ventricular hypertrophy shows a continuum of different age, sex, and amplitude dependence. The diagnosis of left ventricular hypertrophy by the experienced electrocardiographer is also usually made on the consideration of the age, sex and other clinical information, e.g. Hypertension, Aortic valvular disease, etc. Under this

circumstance, an artificial neural network using only ECG parameters as input variables normalised solely by amplitude and duration and not stratified by age and sex factors may not have enough information to handle this complex situation. Perhaps the only solution for improving the diagnostic specificity in left ventricular hypertrophy cases would be to train multiple neural networks corresponding to each age and sex group normalised by the measurements inside each group. In the time available for the present study, it was not feasible to collect such an enormous training material and to train so many networks.

On balance, however, a comparison of the implanted neural network versus isolated neural network versus Glasgow Program indicates that in future, the section of the Glasgow Program dealing with inferior myocardial infarction should contain an implanted neural network. The situation as regards anterior myocardial infarction is currently less clear. All attempts to increase performance have not been successful. It might be worth implanting a neural network for the diagnosis of anterior myocardial infarction into the Glasgow Program if new parameters could be found, because it takes longer to develop a section of deterministic logic than to train a neural network.

It seemed that the performance of various neural networks was not significantly different if the same set of input measurements was used. The input parameters used were also a limitation of this software based neural network approach. Essentially, the neural network uses the same information collected from the signal processing part of the existing deterministic logic. This may be an inherent limitation that the information supplied by the deterministic logic has a plateau and the performance of the neural network can only be developed to this level. It

might be possible that training a neural network with more complete ECG wave forms rather than maximum amplitudes could further enhance its diagnostic power in the future. The neural networks required less time to be developed than a deterministic program and might still be further refined by new learning algorithms or new input parameters. It might also be possible to design a more sophisticated network structure to perform this task.

6.6.2 Theoretical Consideration

Kolmogorov's theorem has already been mentioned. It was of interest that the best neural networks for the diagnosis of both anterior and inferior myocardial infarction had only a single hidden layer. In anterior myocardial infarction, with 42 inputs, the number of hidden neurons (20) was well within the limit (2n+1) of the theorem. Paradoxically, in inferior myocardial infarction with 18 inputs to the network there were 80 hidden neurons and a significant improvement confirmed by the independent CSE database. Therefore, theory and practice are two different matters.

6.6.3 Assessment On CSE Database

The results of the assessment of the final version of the program on the CSE database showed that the sensitivity for each diagnosis was above the recommended minimum accuracy limits except in the diagnosis of right ventricular hypertrophy (Tables 6-13 & 6-15). In this case, the fall in sensitivity of reporting right ventricular hypertrophy from 45% to 39.1% was probably related to the introduction of new continuous equations for the upper normal limits of amplitude. These equations

were developed for neonates, children and adults and probably require adjustment but as their development does not form part of this thesis, no further comment is offered.

The total accuracy of the program also improved from 71.4% with the 1992 version to 74.0% with the 1994 version, a significant improvement. In particular, the new 1994 program had a pronounced improvement in the diagnosis of inferior myocardial infarction and a small gain in the sensitivity of detecting anterior infarct and mixed infarct. The new program M2 had already been shown to be better in the detection of inferior myocardial infarction on the locally collected second test set. This assessment on the CSE database further confirmed that the new program is more sensitive in the diagnosis of inferior myocardial infarction, and suggested that all of the tests undertaken on the locally collected data were indeed valid because of the mirroring of trends in each test set.

In the diagnosis of inferior myocardial infarction, there was a decrease of specificity in both left ventricular hypertrophy cases and normals. In the diagnosis of anterior myocardial infarction, there was an increase of specificity in the left ventricular hypertrophy cases and in normals. This was concordant with the result on the locally collected second test set where the overall specificity of reporting myocardial infarction decreased, particularly in left ventricular hypertrophy cases. The difference in the trend was that the specificity in the normals of the local test set was 100% for both programs. This might be because the 230 normals collected locally were all ambulatory normals, while 96 cases inside the 382 normals in the CSE database were in fact patients with normal coronary arteriograms admitted for the investigation of atypical

chest pain or ECG ST-T abnormalities at rest or after exercise. These cases might not be free from microvascular disease, and might have been reported as myocardial infarction by either the deterministic logic or the neural network implanted deterministic program.

The sensitivity of diagnosing anterior myocardial infarction did not change significantly, but the sensitivity of detecting inferior myocardial infarction had a 10.4% increase on the CSE database and a 19% increase on the locally collected second test set. This increase was overdue because the performance of diagnosing inferior myocardial infarction by the 1992 version of the Glasgow Program (65.8%) was below the 1990 combined program (68.8%). As regards the diagnosis of anterior myocardial infarction, every computer program with higher sensitivity than the Glasgow Program had lower specificity in normals in the 1990 CSE study (Willems et al, 1991). The 1992 version of the Glasgow Program (72.6%) was slightly less sensitive than the average (74.5%) in the diagnosis of anterior myocardial infarction.

The neural network technique considerably increases the sensitivity in diagnosing inferior myocardial infarction but not anterior myocardial infarction. This may be because the deterministic Glasgow program diagnosis of anterior infarct already has reached a diagnostic plateau and can not be improved without sacrificing the specificity in normals as shown by comparing the 1990 CSE study results (Willems et al, 1991) with the 1992 results. The 94.0% specificity for normals in 1990 was caused by low R waves in V2, V3 being misinterpreted as anteroseptal infarction. When this was corrected in 1992 by adding the need for inverted T waves to accompany low R waves for the diagnosis of anterior myocardial infarction, the specificity in normals increased to almost 97%

but the sensitivity in the CSE classification of anterior myocardial infarction fell from 77.9% to 72.6%, showing the direct link between sensitivity and specificity.

On the other hand, in inferior myocardial infarction, the 1992 version gave a sensitivity of 65.8% which approached the sensitivity of the 9 combined programs (Table 6-14) and yet the use of the network produced a significant improvement. In this case, the benefit of using a neural network is clear.

6.6.4 Assessment On Chinese Normals

The neural networks implanted in the modified deterministic program were also assessed on the 503 ECGs from the Chinese normals used in the chapter 3. The result showed 99% specificity in normals for both anterior and inferior myocardial infarction which was lower than in the locally collected 230 Caucasian normals (100% for both anterior and inferior myocardial infarction) but higher than in the CSE database 382 Caucasian normals (98.7% for anterior myocardial infarction and 98.4% for inferior myocardial infarction). The reason for this discordance in the results might be based on the fact that the way of collecting Chinese normals was not the same as that of collecting the CSE normals, i.e. there were no patients with chest pain or abnormal ST-T changes included in the Chinese cohort.

The results of Chapter 3 and Chapter 6 raise the question of whether the training of neural networks should be race dependent. On the one hand, the 99% specificity of the neural networks developed in Caucasians and tested in Chinese suggests that the approach need not

be race dependent. On the other hand, the results may disguise racial dependence because normals should not exhibit significant Q waves in the inferior or anterior leads in any event, while the modifications to the logic also inhibited the false positive diagnosis of myocardial infarction in the presence of small R waves. On balance, because there is known racial variation in these ECG parameters, especially in amplitude measurements, between Caucasians and Chinese in each corresponding age group as demonstrated in Chapter 3 and a previous study (Macfarlane, Chen & Chiang, 1988), logically, the neural network for the diagnosis of myocardial infarction should also be trained with materials from different races. Thus, in future, it might be valuable to recruit myocardial infarction and left ventricular hypertrophy cases from the Chinese population for further study of neural networks.

6.6.5 Advantages of Neural Network Implantation

The present study has shown that the incorporation of an artificial neural network in a deterministic program has the following advantages over using either alone:

(1) There is improved control of specificity in diagnosing myocardial infarction in left ventricular hypertrophy cases.

(2) Reasons can be retained together with diagnostic statements, in order to provide more support for the diagnosis of certain abnormalities, e.g. "T inversion is also present" provides supplementary evidence of reporlarisation abnormalities in the diagnosis of myocardial infarction. In addition, this can also provide more information and reassurance to the less experienced physicians in clinical decision making.

(3) More control is obtained compared to using a single neural network in isolation for diagnosing all major abnormalities.

(4) It is possible to preserve the timing of events, e.g. acute, old, etc. which can not be obtained from the use of single neural network in isolation unless it is trained with large numbers of suitably classified cases and has multiple output neurons as appropriate. Alternatively, if neural networks are used in isolation, it would be necessary to train one neural network for each classification, i.e. old, acute etc. and so there would be multiple neural networks involved for the diagnosis. Another consideration is that the collection of the training materials and the time needed for training is not economical compared to the incorporation of the neural networks into an existing deterministic program.

6.7 CONCLUSIONS

The implantation of a neural network into the deterministic program together with modification of the logic improved the sensitivity but decreased the overall specificity in normals for the diagnosis of either anterior or inferior myocardial infarction compared to the original deterministic program. The enhanced sensitivity was more pronounced in inferior myocardial infarction which was the same as demonstrated in the experiments using neural networks in isolation.

It was learned that various methods of implanting a neural network into the logic with different modifications can lead to different results. Thus, the selection of the optimal point inside the deterministic logic for

incorporation and the choice of a suitable method of modification has an important influence on the performance of the program.

In conclusion, the neural network implanted deterministic program has an improved performance and clearly can be of use in the computer-assisted ECG diagnosis of myocardial infarction. Further improvements, however, can still be expected.

CHAPTER 7

CONCLUSIONS

7.1 GENERAL CONCLUSIONS

Several conclusions can be drawn from the current experience based on the study of the normal derived vectorcardiograms and the use of neural networks for the diagnosis of atrial fibrillation and myocardial infarction.

(1) The artificial neural network approach can be beneficial if used selectively for certain diagnoses, e.g. it has been shown in the present study that a neural network implanted into deterministic logic together with some modification of the existing logic can improve performance dramatically in the diagnosis of inferior myocardial infarction (10.4% increase of sensitivity on the CSE database). On the contrary, there was no such improvement in the diagnosis of anterior myocardial infarction (0.3% increase of sensitivity on the CSE database), underlining the selective nature of the benefits of neural networks.

(2) It is also demonstrated from the present study that the neural network implanted deterministic program (using only twelve-lead ECG measurements) can produce the equivalent diagnostic accuracy (74%) on the CSE database as the Modular ECG Analysis System program (Kors et al, 1992) which (i) combined 12-lead ECG and derived VCG interpretations to obtain 73.6% accuracy, and (ii) combined 12-lead ECG and Frank VCG interpretations to obtain 74.2% accuracy. The accuracy of this 12-lead ECG program alone

was 69.8%. However, this thesis has demonstrated that if the appropriate 12-lead ECG measurements are used, there appears to be no gain from adding VCG parameters to ECG measurements for automated ECG analysis.

(3) This study has shown that it takes considerably less time and experience to develop a well trained neural network than to produce a specific section of deterministic logic for the diagnosis of atrial fibrillation or myocardial infarction in order to achieve at least an equivalent performance. An artificial neural network may even be considerably superior to the original well developed deterministic logic as exemplified in this study with respect to the detection of inferior myocardial infarction.

(4) Artificial neural networks can be integrated with an existing deterministic program to assist its interpreting capability. The neural network integrated deterministic program can have the advantage of preserving the diagnostic reason statements and the timing descriptors (old, acute) for myocardial infarction statements which can further clarify the diagnosis for physicians.

(5) In some situations, the choice of diagnostic parameters appears to be more important than using either a neural network or deterministic logic as shown in the study of atrial fibrillation.

(6) The use of the artificial neural network approach is unlikely to provide the solution to all the diagnostic classification problems in computer-assisted ECG interpretation, e.g. the age, sex and racial variations of the twelve-lead ECG reported previously (Macfarlane,

Chen & Chiang, 1988) and of the derived VCG measurements as demonstrated in the present study (Chapter 3; Yang & Macfarlane, 1994 b&c) cannot easily be dealt with by a single neural network. Although the use of multiple neural networks trained according to age, sex and race could be a solution, it is impractical to obtain such training materials without an extremely large scale international effort.

(7) Although "Kolmogorov's theorem" (i.e. Twice the number of input neurons + 1 = optimum number of hidden neurons) predicts the sufficient number of hidden neurons for computing any arbitrary continuous function, in this study neural networks with a sufficient number of hidden neurons did not always perform better than other neural networks either in the detection of atrial fibrillation (chapter 4) or in the diagnosis of myocardial infarction (chapter 5). The experience of the present study suggests that there is no means of predicting the optimum topology or the performance of a neural network before it is assessed by the test set. Performance in the present study was evaluated by the Association Index, with emphasis being placed on obtaining a high specificity for a particular diagnosis.

7.2 LIMITATIONS OF THE CURRENT STUDY

There are so many waveform measurements in electrocardiograms that the possibility of training a single neural network to interpret all ECGs accurately with high specificity is limited. Indeed, a patient with an old myocardial infarction may well have a normal ECG, so that myocardial infarction can not be diagnosed either by the deterministic or neural

network approach, or even by the experienced electrocardiographer (Figs 7-1 & 7-2). In other words, the limitations of artificial neural networks are closely related to the inherent limitations of electrocardiography itself in the diagnosis of myocardial infarction, because the gold standard used is the clinical diagnosis, whereas in the diagnosis of atrial fibrillation, there is no such clinically related problem, because the gold standard of the rhythm diagnosis is the ECG wave form itself. On the other hand, the number of ECGs collected is never sufficient for use in the training set for a neural network. The more varieties of ECGs of certain classifications used in the training process, the better will be the performance of the neural network in diagnosing that abnormality. Therefore, it is still possible that more training material will be collected for the development of neural networks in the future. In addition, this study has used a limited number of measurements for each lead. Perhaps newer approaches combining detailed classification of ST-T segments (Edenbrandt, Devine & Macfarlane, 1992) with separate Q wave classification networks will offer further improvement.

7.3 ARTIFICIAL NEURAL NETWORKS versus DETERMINISTIC CRITERIA

Deterministic methods appeal more to the cardiologist than neural network approaches because they are more readily understood. Deterministic criteria may also be selected on the basis of a knowledge of well-established electrophysiological processes. This approach is also reasonably flexible so that alterations to existing criteria can be implemented, and new diagnostic categories easily added (Bailey & Horton, 1977; Kors & van Bemmel, 1990). However, the deterministic approach sets sharp thresholds for decision making. If, for example, a

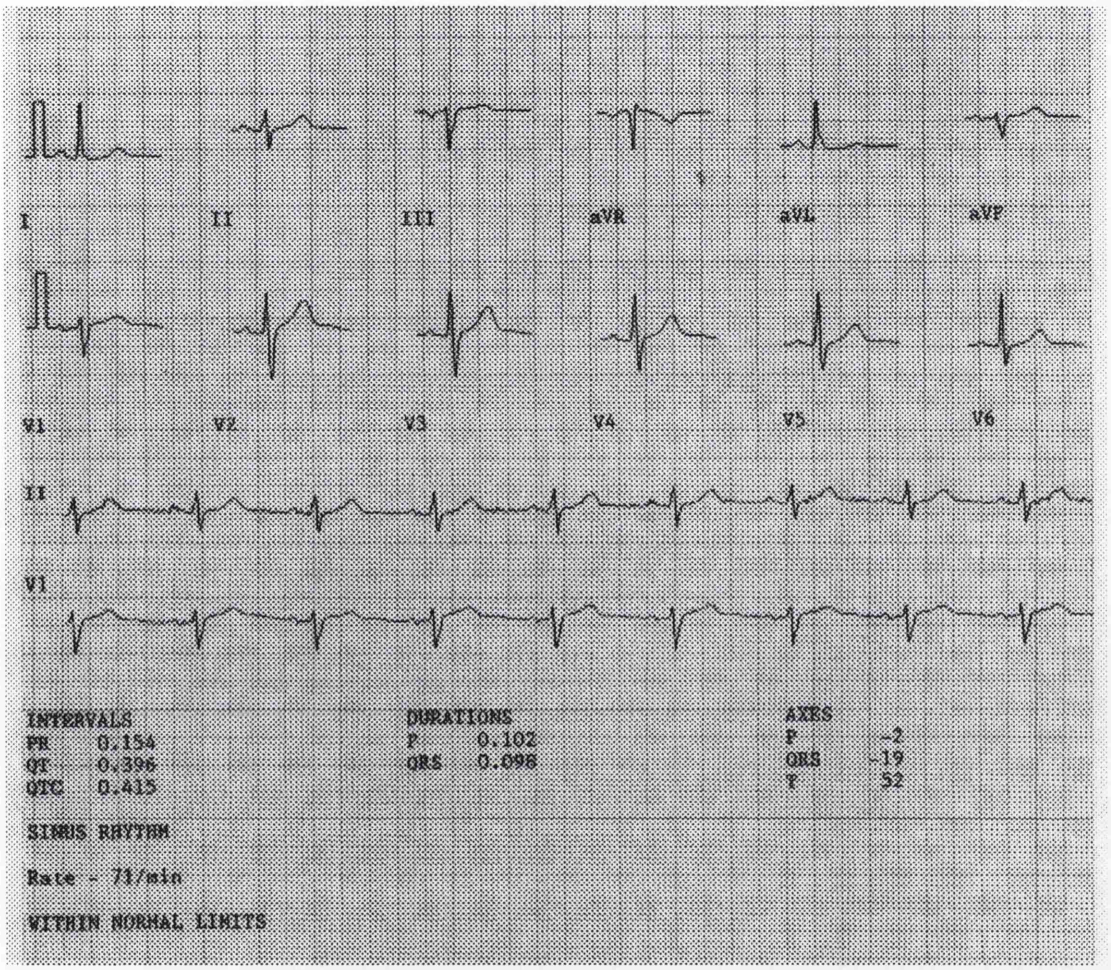


Fig.7-1 A normal ECG recorded from a patient with a proven inferior myocardial infarction.

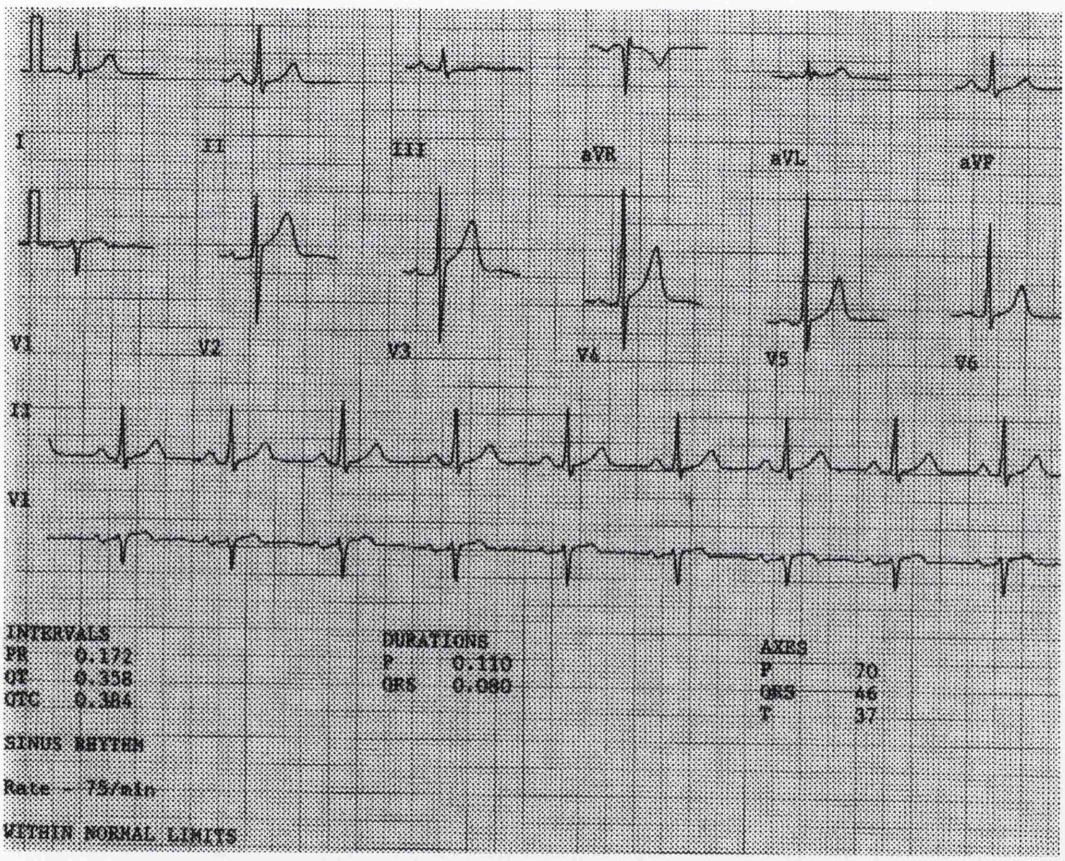


Fig.7-2 A normal ECG recorded from a patient with a proven anterior myocardial infarction.

30 msec Q wave is required for the diagnosis of myocardial infarction, then any value less than this is not considered to represent infarction, even 29 msec, but in the clinical situation, a cardiologist will not discriminate to such a level and will interpret the pattern as a whole. Artificial neural networks have a mathematically more complex way of associating an ECG pattern with a diagnosis which is completely different from and more flexible than the deterministic logic.

In making an ECG interpretation, an experienced electrocardiographer considers other wave forms, e.g. ST-T changes and demographic characteristics, namely, age, sex, race, symptoms and signs as well as clinical history. In essence, electrocardiographers follow Bayesian probability theory in ECG diagnosis, and subconsciously use more flexible thresholds for the actual ECG measurements.

7.4 THE FUTURE OF ARTIFICIAL NEURAL NETWORKS IN ELECTROCARDIOGRAPHY

The use of neural networks as well as deterministic and statistical approaches for ECG analysis requires a substantial amount of data from different populations since it is known that ECG and derived VCG parameters vary according to age, sex, and race. The present study has underlined these differences by looking at the derived VCG in Caucasians and Chinese but it was not feasible to pursue these differences by obtaining training cases of myocardial infarction, etc., in Chinese patients. Therefore, it is also reasonable to predict that in the future, race-related training sets will be used to design neural networks to enhance computerised ECG interpretations.

It would certainly be possible to train multiple neural networks each using only a single lead measurements (e.g. three neural networks, which use V2, V3 and V4 measurements as inputs respectively, but run together for the diagnosis of anteroseptal myocardial infarction) and to introduce new diagnostic parameters as input variables in order to improve performance. It is also feasible to classify QRS and ST-T segments separately. This could lead to the incorporation of multiple artificial neural networks at particular points of the deterministic logic. In other words, within a deterministic program, it is quite conceivable that a whole series of neural networks could be invoked to determine whether or not a set of ECG wave measurements was suggestive of atrial fibrillation, anterior or inferior myocardial infarction, myocardial ischaemia etc. The present study therefore represents a step forward in utilising artificial neural networks in a limited but realistic manner in automated ECG interpretation.

APPENDIX 1

SUPPLEMENTARY TABLES FOR CHAPTER 3

Table 3-6 Magnitude of the projection of the maximal QRS vector onto the frontal(F), horizontal(H), and right sagittal(R) planes in 1555 Caucasians [microVolts]

	Age	<30	30-39	40-49	50-59	60+
F male	M±SD	2039±525	1802±489	1586±462	1503±424	1401±373
	Range	1083-3146	874-2936	728-2493	780-2485	848-2081
F female	M±SD	1538±404	1512±420	1289±369	1288±334	1268±524
	Range	722-2462	740-2384	743-2067	677-1961	552-2309
H male	M±SD	2088±515	1850±475	1640±427	1536±390	1417±363
	Range	1074-3205	1004-2891	858-2586	897-2335	861-2146
H female	M±SD	1506±392	1504±400	1306±355	1312±312	1270±475
	Range	696-2422	881-2380	759-2008	691-2014	799-2240
R male	M±SD	1748±605	1485±580	1226±452	1070±402	1118±406
	Range	796-3315	561-3075	515-2221	491-1862	528-1918
R female	M±SD	1278±442	1254±432	1018±349	1021±362	820±298
	Range	519-2277	643-2209	395-1777	421-1857	591-1281

Table 3-7 Magnitude of the projection of the maximal QRS vector onto the frontal(F), horizontal(H), and right sagittal(R) planes in 503 Chinese [microVolts]

	Age	<30	30-39	40-49	50-59	60+
F male	M±SD	1891±468	1746±464	1652±473	1650±412	1662±489
	Range	1061-2961	904-2622	711-2660	1027-2520	794-2465
F female	M±SD	1497±350	1496±319	1371±342	1423±351	1406±394
	Range	872-2011	903-2010	657-2247	854-2072	944-2363
H male	M±SD	1876±475	1705±390	1649±419	1592±411	1608±467
	Range	1174-2852	1132-2359	934-2459	962-2431	732-2452
H female	M±SD	1311±326	1361±296	1342±342	1388±338	1356±365
	Range	756-2122	884-1998	951-2310	818-2083	774-2266
R male	M±SD	1512±559	1347±497	1194±406	1079±521	1086±430
	Range	754-2792	596-2388	586-2085	503-2795	532-1872
R female	M±SD	1164±395	1082±344	1040±371	952±308	873±328
	Range	496-2055	481-2058	520-1917	426-1585	429-1488

Table 3-8 Magnitude of the projection of the maximal T vector onto the frontal(F), horizontal(H), and right sagittal(R) planes in 1555 Caucasians [microVolts]

	Age	<30	30-39	40-49	50-59	60+
F male	M±SD	527±199	472±174	424±177	392±161	430±190
	Range	197-1002	159-855	158-838	129-783	111-708
F female	M±SD	396±130	377±135	324±122	319±124	305±151
	Range	170-681	161-704	90-581	104-591	79-691
H male	M±SD	649±196	587±176	543±173	521±179	497±201
	Range	302-1088	280-992	257-945	213-911	146-968
H female	M±SD	403±136	394±130	338±122	331±114	309±147
	Range	172-730	191-729	96-577	99-625	162-614
R male	M±SD	515±169	461±154	430±146	414±148	403±165
	Range	212-895	216-830	176-758	159-714	132-762
R female	M±SD	297±110	286±97	254±90	255±83	246±104
	Range	110-595	110-491	93-461	110-415	157-474

Table 3-9 Magnitude of the projection of the maximal T vector onto the frontal(F), horizontal(H), and right sagittal(R) planes in 503 Chinese [microVolts]

	Age	<30	30-39	40-49	50-59	60+
F male	M±SD	530±166	470±196	416±156	432±196	377±155
	Range	209-836	119-893	127-758	117-767	152-715
F female	M±SD	406±111	377±133	325±107	287±136	285±127
	Range	205-600	129-625	149-499	77-579	123-579
H male	M±SD	639±176	573±184	503±173	506±189	432±141
	Range	319-963	290-895	224-879	155-848	220-790
H female	M±SD	384±113	372±146	319±124	293±141	275±129
	Range	203-557	131-680	136-653	81-613	124-558
R male	M±SD	473±141	433±141	354±146	360±129	325±123
	Range	221-777	213-733	127-690	121-621	154-598
R female	M±SD	258±98	245±107	218±85	198±80	183±73
	Range	130-448	87-574	103-376	55-339	77-334

Table 3-10 96-percentile ranges of angles of the projection of maximal QRS vector onto the frontal(F), horizontal(H), and right sagittal(R) planes in 1555 Caucasians [degrees]

Age	<30	30-39	40-49	50-59	60+
F male	12-63	0 - +57	-3 - +54	-1 - +43	-6 - +58
F female	13-59	16-54	6-48	6-53	0-47
H male	-105 - +14	-100 - +22	-104 - +32	-116 - +25	-109 - +25
H female	-85 - +4	-97 - -5	-109 - +13	-108 - +16	-107 - +27
R male	33-190	16-212	22-207	25-204	115-180
R female	70-199	101-193	64-200	38-195	83-183

Table 3-11 96-percentile ranges of angles of the projection of maximal QRS vector onto the frontal(F), horizontal(H), and right sagittal(R) planes in 503 Chinese [degrees]

Age	<30	30-39	40-49	50-59	60+
F male	7-60	6-51	9-57	5-66	3-56
F female	16-67	8-52	9-49	1-45	5-43
H male	-69 - +20	-77 - +46	-71 - +21	-113 - +17	-106 - +29
H female	-101 - +8	-66 - +26	-91 - +19	-66 - +32	-48 - +23
R male	92-197	11-186	87-183	11-187	36-180
R female	80-178	54-194	27-184	21-221	27-204

Table 3-12 96-percentile ranges of angles of the projection of maximal T vector onto the frontal(F), horizontal(H), and right sagittal(R) planes in 1555 Caucasians [degrees]

Age	<30	30-39	40-49	50-59	60+
F male	1 - 60	6 - 53	3 - 56	3-51	10-81
F female	8-50	9-48	9-56	7-53	0-153
H male	13-81	5 - 79	6-77	2-78	2-65
H female	-15 - +62	-15 - +68	10-77	-18 - +77	-6-+105
R male	-4 - +66	-1 - +84	-1 - +69	-2 - +89	4-85
R female	0 - +110	-3 - +100	-5 - +82	-12 - +103	0-102

Table 3-13 96-percentile ranges of angles of the projection of maximal T vector onto the frontal(F), horizontal(H), and right sagittal(R) planes in 503 Chinese [degrees]

Age	<30	30-39	40-49	50-59	60+
F male	5-43	1-46	5-43	-4-44	2-46
F female	6-47	-2 - +65	10-58	-1 - +66	-19 - +59
H male	20-62	5-73	11-67	13-69	-1 - +60
H female	-22 - +55	-50 - +52	-47 - +50	-31 - +67	-34 - +48
R male	-1 - +62	-4 - +69	2-56	-5 - +62	0-86
R female	-55 - +109	-41 - +174	-85 - +138	-40 - +129	-26 - +120

Table 3-14 Magnitude of the projection of the initial 20 mS QRS vector onto the frontal(F), horizontal(H), and right sagittal(R) planes in 1555 Caucasians [microVolts]

	Age	<30	30-39	40-49	50-59	60+
F male	M±SD	165±128	192±140	193±187	240±167	210±159
	Range	20-468	20-549	17-766	32-575	39-482
F female	M±SD	183±140	165±100	210±154	207±136	217±160
	Range	21-554	28-399	10-475	19-675	87-667
H male	M±SD	450±190	385±168	357±154	347±153	316±146
	Range	143-898	132-786	105-809	105-684	69-562
H female	M±SD	347±147	329±134	295±113	330±156	294±179
	Range	127-665	121-641	126-531	109-749	97-731
R male	M±SD	424±188	339±164	299±128	285±118	266±145
	Range	94-876	99-776	85-629	98-563	85-586
R female	M±SD	312±144	293±138	238±106	251±126	236±135
	Range	107-645	99-776	43-479	73-702	32-486

Table 3-15 Magnitude of the projection of the initial 20 mS QRS vector onto the frontal(F), horizontal(H), and right sagittal(R) planes in 503 Chinese [microVolts]

	Age	<30	30-39	40-49	50-59	60
F male	M±SD	241±158	260±172	320±200	241±220	347±213
	Range	47-579	56-597	81-593	21-686	33-724
F female	M±SD	219±205	266±183	274±190	303±200	364±290
	Range	14-506	29-600	56-657	48-657	28-861
H male	M±SD	491±220	435±200	463±206	394±189	431±208
	Range	170-815	163-845	154-719	157-804	132-782
H female	M±SD	470±226	427±191	431±215	395±186	446±257
	Range	139-955	189-758	188-711	85-729	82-865
R male	M±SD	428±215	330±196	343±146	315±124	287±131
	Range	139-730	87-739	129-583	101-519	111-520
R female	M±SD	416±206	339±157	361±193	274±135	268±149
	Range	122-823	163-696	140-540	87-594	68-525

Table 3-16 Magnitude of the projection of the initial 30 mS QRS vector onto the frontal(F), horizontal(H), and right sagittal(R) planes in 1555 Caucasians [microVolts]

	Age	<30	30-39	40-49	50-59	60+
F male	M±SD	810±374	868±373	882±349	902±348	885±236
	Range	210-1610	259-1828	288-1860	223-1717	370-1199
F female	M±SD	874±322	819±286	824±265	889±281	876±464
	Range	267-1600	274-1318	397-1551	280-1359	228-2219
H male	M±SD	860±325	898±340	905±330	909±327	874±242
	Range	333-1564	379-1784	399-1788	342-1706	466-1252
H female	M±SD	817±301	760±259	793±24	871±268	851±443
	Range	286-1594	320-1241	424-1443	384-1377	245-2122
R male	M±SD	643±276	563±294	514±238	459±197	458±230
	Range	237-1400	178-1219	191-1174	195-907	186-966
R female	M±SD	607±262	536±198	478±202	516±256	378±170
	Range	229-1333	213-967	174-973	194-1416	139-658

Table 3-17 Magnitude of the projection of the initial 30 mS QRS vector onto the frontal(F), horizontal(H), and right sagittal(R) planes in 503 Chinese [microVolts]

	Age	<30	30-39	40-49	50-59	60+
F male	M±SD	962±487	929±457	1031±368	1021±343	1107±395
	Range	194-1667	134-1690	526-1726	458-1557	491-1669
F female	M±SD	952±348	918±358	921±394	1013±333	999±471
	Range	431-1519	201-1675	123-1473	495-1474	376-1695
H male	M±SD	998±393	936±394	1048±316	1013±301	1098±373
	Range	352-1498	310-1570	556-1650	604-1466	607-1703
H female	M±SD	866±314	853±302	909±366	1006±305	1018±368
	Range	397-1526	447-1532	279-1514	506-1455	468-1582
R male	M±SD	706±270	647±307	600±228	545±240	540±178
	Range	305-1292	284-1247	257-949	248-951	303-797
R female	M±SD	681±291	588±229	523±287	499±207	485±212
	Range	296-1248	302-1029	163-957	220-827	187-869

Table 3-18 Angle of the projection of the initial 20 mS QRS vector onto the frontal(F), horizontal(H), and right sagittal(R) planes in 1555 Caucasians [degrees]

	Age	<30	30-39	40-49	50-59	60+
F male	Mean	-27	-12	0	+10	0
	Range	-177 - +162	-169 - +157	-164 - +160	-160 - +159	-123 - +95
F female	Mean	+6	-7	+16	+9	+25
	Range	-170 - +163	-171 - +173	-136 - +163	-143 - +108	-142 - +161
H male	Mean	+83	+70	+63	+63	+58
	Range	+19 - +127	0 - +131	-9 - +117	+12 - +122	+1 - +109
H female	Mean	+64	+69	+56	+47	+58
	Range	-35 - +123	+10 - +123	-8 - +111	-24 - +110	0 - +178
R male	Mean	+2	+6	+13	+13	+13
	Range	-26 - +80	-33 - +89	-23 - +112	-18 - +77	-35 - +81
R female	Mean	+17	+10	+23	+27	+22
	Range	-25 - +142	-24 - +69	-15 - +107	-23 - +119	-11 - +84

Table 3-19 Angle of the projection of the initial 20 mS QRS vector onto the frontal(F), horizontal(H), and right sagittal(R) planes in 503 Chinese [degrees]

	Age	<30	30-39	40-49	50-59	60+
F male	Mean	-13	-16	0	0	+7
	Range	-154 - +162	-150 - +155	-73 - +55	-136 - +83	-77 - +76
F female	Mean	+10	+5	-5	+6	+3
	Range	-132 - +93	-160 - +105	-122 - +50	-134 - +69	-72 - +58
H male	Mean	+71	+71	+51	+59	+53
	Range	+9 - +118	+21 - +115	0 - +92	+19 - +105	+19 - +99
H female	Mean	+58	+63	+49	+48	+45
	Range	+5 - +105	+2 - +113	-1 - +110	+8 - +128	+10 - +94
R male	Mean	+5	+5	+12	+12	+12
	Range	-24 - +40	-23 - +50	-29 - +84	-14 - +55	-19 - +48
R female	Mean	+19	+16	+17	+17	+16
	Range	-20 - +68	-15 - +82	-37 - +90	-31 - +75	-30 - +68

Table 3-20 P wave amplitude in three orthogonal leads in 1555 Caucasians [microVolts]

	Age	<30	30-39	40-49	50-59	60+
X male	M±SD	71±22	74±21	73±22	74±21	68±25
	Range	31-119	32-118	36-132	43-123	24-119
X female	M±SD	69±20	76±22	76±23	78±21	75±26
	Range	33-122	36-122	41-124	37-128	30-138
Y male	M±SD	111±51	110±44	109±42	105±41	109±41
	Range	21-236	35-215	39-212	36-198	49-178
Y female	M±SD	110±48	119±51	108±44	113±42	116±60
	Range	23-223	29-235	34-214	42-198	26-233
Z male	M±SD	51±21	46±21	42±18	41±20	45±25
	Range	20-95	13-111	13-84	13-80	14-93
Z female	M±SD	44±17	42±17	40±17	38±16	35±13
	Range	14-81	14-85	12-82	11-65	10-60

Table 3-21 P wave amplitude in three orthogonal leads in 503 Chinese [microVolts]

	Age	<30	30-39	40-49	50-59	60+
X male	M±SD	57±17	61±22	66±15	69±18	65±16
	Range	29-95	22-93	41-100	40-102	36-92
X female	M±SD	59±17	64±16	68±17	69±28	80±21
	Range	35-94	32-90	38-99	37-156	46-117
Y male	M±SD	111±56	122±45	102±43	112±48	120±40
	Range	25-214	42-218	34-164	33-193	50-190
Y female	M±SD	105±51	110±59	112±29	122±68	120±37
	Range	26-203	31-239	62-167	43-199	40-184
Z male	M±SD	45±19	37±14	34±14	34±12	31±14
	Range	12-93	15-69	13-66	16-61	12-67
Z female	M±SD	40±13	40±14	41±15	37±15	37±15
	Range	16-61	19-76	19-74	10-60	14-63

Table 3-22 P wave duration in three orthogonal leads in 1555 Caucasians [mS]

Age	<30	30-39	40-49	50-59	60+
Male M±SD	98±14	103±11	107±11	108±11	110±12
Range	66-122	78-124	86-128	80-128	90-130
Female M±SD	97±11	100±10	104±10	104±13	104±12
Range	72-116	78-116	82-124	64-128	88-130

Table 3-23 P wave duration in three orthogonal leads in 503 Chinese [mS]

Age	<30	30-39	40-49	50-59	60+
Male M±SD	103±16	106±13	105±12	107±11	109±11
Range	66-134	86-132	76-124	84-124	82-122
Female M±SD	96±13	104±19	109±19	106±13	112±16
Range	64-118	70-148	86-148	80-132	78-140

Table 3-24 R wave amplitude in three orthogonal leads in 1555 Caucasians [microVolts]

	Age	<30	30-39	40-49	50-59	60+
X male	M±SD	1662±457	1533±431	1403±430	1375±380	1200±362
	Range	937-2800	686-2445	662-2417	720-2014	744-1932
X female	M±SD	1226±339	1226±365	1106±323	1118±285	1278±464
	Range	565-1961	579-1970	620-1749	592-1761	539-2225
Y male	M±SD	1161±458	909±455	686±392	577±338	646±392
	Range	327-2238	138-2125	40-1508	109-1460	84-1502
Y female	M±SD	906±343	873±317	655±259	614±308	549±212
	Range	319-1749	297-1481	237-1163	135-1309	245-1155
Z male	M±SD	519±227	424±243	366±186	357±168	311±187
	Range	154-1022	91-1067	49-868	103-706	56-669
Z female	M±SD	332±149	317±146	255±118	263±144	319±143
	Range	74-700	103-667	45-490	60-725	68-563

Table 3-25 R wave amplitude in three orthogonal leads in 503 Chinese [microVolts]

	Age	<30	30-39	40-49	50-59	60+
X male	M±SD	1576±400	1461±368	1436±430	1444±408	1485±484
	Range	883-2571	826-2148	681-2266	717-2218	734-2298
X female	M±SD	1134±327	1202±297	1173±311	1284±336	1239±405
	Range	660-1755	717-1742	510-1995	640-1839	700-2191
Y male	M±SD	966±502	891±481	765±395	744±439	720±442
	Range	79-1935	38-2073	151-1538	192-1766	170-1615
Y female	M±SD	916±332	869±277	712±267	634±287	585±311
	Range	348-1634	313-1386	191-1187	196-1220	120-1180
Z male	M±SD	526±250	502±232	397±227	407±171	424±186
	Range	136-969	210-1136	60-927	155-754	118-720
Z female	M±SD	392±239	369±129	325±146	364±181	358±206
	Range	97-1165	131-600	114-644	100-757	100-881

Table 3-26 S wave amplitude in three orthogonal leads in 1555 Caucasians [microVolts]

	Age	<30	30-39	40-49	50-59	60+
X male	M±SD	[-270]±175	[-260]±197	[-289]±176	[-253]±166	[-345]±181
	Range	[-640]-[-61]	[-650]-[-66]	[-568]-[-99]	[-582]-[-91]	[-622]-[-94]
X female	M±SD	[-168]±100	[-199]±110	[-181]±140	[-196]±135	[-164]±98
	Range	[-364]-[-52]	[-337]-[-64]	[-341]-[-42]	[-416]-[-61]	[-344]-[-42]
Y male	M±SD	[-157]±93	[-184]±161	[-172]±127	[-174]±155	[-178]±140
	Range	[-306]-[-36]	[-419]-[-42]	[-402]-[-35]	[-372]-[-49]	[-446]-[-48]
Y female	M±SD	[-145]±88	[-139]±117	[-136]±108	[-161]±94	[-181]±104
	Range	[-269]-[-46]	[-292]-[-53]	[-251]-[-37]	[-286]-[-59]	[-408]-[-102]
Z male	M±SD	[-1383]±471	[-1049]±372	[-989]±353	[-865]±322	[-811]±307
	Range	-2056- -633	-1676 - -534	-1519 - -475	-1361 - -373	-1285 - -439
Z female	M±SD	[-916]±328	[-990]±388	[-792]±270	[-744]±301	[-658]±253
	Range	-1503 - -423	-1530 - -528	-1253 - -436	-1315 - -338	-1097 - -271

Table 3-27 S wave amplitude in three orthogonal leads in 503 Chinese [microVolts]

	Age	<30	30-39	40-49	50-59	60+
X male	M±SD	[-264]±214	[-328]±238	[-309]±163	[-260]±196	[-342]±193
	Range	[-35]-[-695]	[-63]-[-743]	[-65]-[-589]	[-62]-[-495]	[-102]-[-612]
X female	M±SD	[-210]±161	[-171]±123	[-212]±93	[-226]±121	[-282]±194
	Range	[-43]-[-481]	[-66]-[-342]	[-94]-[-321]	[-92]-[-357]	[-82]-[-657]
Y male	M±SD	[-175]±129	[-115]±85	[-174]±100	[-136]±158	[-89]±52
	Range	[-42]-[-428]	[-32]-[-269]	[-33]-[-403]	[-30]-[-239]	[-35]-[-159]
Y female	M±SD	[-142]±92	[-127]±67	[-144]±90	[-164]±141	[-171]±77
	Range	[-54]-[-322]	[-60]-[-233]	[-62]-[-224]	[-44]-[-325]	[-72]-[-287]
Z male	M±SD	[-1140]±422	[-946]±280	[-844]±257	[-794]±369	[-755]±280
	Range	-575 - -1908	-652 - -1507	-501 - -1190	-443 - -1252	-282 - -1133
Z female	M±SD	[-768]±275	[-767]±276	[-735]±244	[-692]±240	[-609]±308
	Range	-424 - -1179	-455 - -1292	-324 - -1080	-390 - -1079	-115 - -1149

Table 3-28 T wave amplitude in three orthogonal leads in 1555 Caucasians [microVolts]

	Age	<30	30-39	40-49	50-59	60+
X male	M±SD	462±186	418±162	375±167	352±155	364±182
	Range	145-934	122-801	89-823	101-716	60-660
X female	M±SD	339±117	325±125	276±113	270±112	261±140
	Range	137-604	119-644	54-513	70-559	31-619
Y male	M±SD	247±111	212±100	189±97	166±79	209±118
	Range	44-529	45-464	41-416	57-366	72-470
Y female	M±SD	202±78	184±77	162±68	167±73	170±79
	Range	65-388	62-352	47-326	33-324	42-329
Z male	M±SD	468±179	420±158	392±151	382±160	341±166
	Range	138-862	130-803	116-705	94-701	109-730
Z female	M±SD	228±114	226±104	202±92	199±86	209±82
	Range	43-555	52-429	55-425	28-62	38-369

Table 3-29 T wave amplitude in three orthogonal leads in 503 Chinese [microVolts]

	Age	<30	30-39	40-49	50-59	60+
X male	M±SD	483±159	423±184	383±144	387±189	330±133
	Range	199-810	114-808	104-688	101-718	140-636
X female	M±SD	351±98	331±127	286±106	254±131	259±117
	Range	189-517	108-615	116-491	58-524	116-505
Y male	M±SD	220±88	205±103	156±82	178±98	175±103
	Range	87-414	59-406	56-349	24-416	48-444
Y female	M±SD	199±83	179±83	155±54	135±66	135±62
	Range	46-372	60-317	47-251	31-294	51-267
Z male	M±SD	442±143	393±142	328±149	322±121	284±115
	Range	209-710	191-718	121-690	135-618	72-498
Z female	M±SD	172±100	168±118	152±97	152±87	138±67
	Range	47-390	49-509	40-351	32-332	42-277

APPENDIX 2
SUPPLEMENTARY TABLES FOR CHAPTER 5

Se=Sensitivity, Sp=Specificity, LVH Sp=Sp of MI in the ECGs from LVH cases, AI=Association Index=Se+Sp-100

Table 5-8a, b & c The performance of the 55 neural networks using 46 input variables (QRS+ST-T+dVCG) with only a single output neuron for the diagnosis of anterior myocardial infarction

Table 5-8a The performance of the 15 artificial neural networks with single hidden layer on the test set ECGs for the diagnosis of anterior myocardial infarction.

Number of neurons in the hidden layers	Se	Sp	AI	LVH Sp
5	93	100	93	48
10	93	100	93	46
15	93	100	93	45
20	92	100	92	49
25	92	100	92	50
30	92	100	92	45
35	92	100	92	50
40	94	100	94	43
45	92	100	92	44
50	92	100	92	49
60	93	100	93	48
70	93	100	93	49
80	93	100	93	46
90	92	100	92	50
100	92	100	92	45

Table 5-8b The performance of the 16 artificial neural networks with two hidden layers on the test set ECGs for the diagnosis of anterior myocardial infarction.

Number of neurons in the hidden layers	Se	Sp	AI	LVH Sp
5-5	97	99	96	50
10-10	93	100	93	49
15-15	92	100	92	48
25-25	92	100	92	49
30-30	93	100	93	45
35-35	92	100	92	51
40-40	92	100	92	44
45-45	92	100	92	49
50-50	92	100	92	46
60-60	93	100	93	45
70-70	92	100	92	44
70-90	92	100	92	46
80-80	92	100	92	50
80-90	92	100	92	50
90-90	92	100	92	50
100-100	93	100	93	48

Table 5-8c The performance of the 24 artificial neural networks with three hidden layer on the test set ECGs for the diagnosis of anterior myocardial infarction.

Number of neurons in the hidden layers	Se	Sp	AI	LVH Sp
5-5-5	95	99	94	49
10-10-10	94	100	94	50
15-15-15	94	99	93	46
20-20-20	93	100	93	54
20-20-100	92	100	92	48
25-25-25	92	100	92	46
30-30-30	92	100	92	48
30-30-100	93	100	93	52
35-35-35	92	100	92	50
40-40-40	92	100	92	45
45-45-45	92	100	92	48
45-45-100	92	100	92	52
50-50-50	92	100	92	46
50-50-100	94	100	94	52
60-60-60	93	100	93	46
60-60-100	92	100	92	49
70-70-70	93	100	93	45
70-70-90	93	100	93	45
80-80-80	92	100	92	45
80-80-100	93	100	93	46
90-90-90	93	99	92	44
90-90-100	92	100	92	45
90-100-100	93	100	93	45
100-100-100	93	100	93	49

Table 5-9a, b & c The performance of 52 neural networks using 42 input variables (QRS+ST-T) with only a single output neuron for the diagnosis of anterior myocardial infarction

Table 5-9a The performance of the 14 artificial neural networks with a single hidden layer on the test set ECGs for the diagnosis of anterior myocardial infarction.

Hidden layers	Se	Sp	AI	LVH Sp
5	94	99	93	50
10	94	100	94	48
15	94	98	92	45
20	94	100	94	46
25	94	100	94	48
30	94	100	94	48
35	93	100	93	48
40	93	100	93	46
45	94	99	93	49
50	94	100	94	46
60	93	100	93	48
70	94	100	94	49
80	94	100	94	46
90	94	100	94	48

Table 5-9b The performance of the 19 artificial neural networks with two hidden layers on the test set ECGs for the diagnosis of anterior myocardial infarction.

Number of neurons in the hidden layers	Se	Sp	AI	LVH Sp
5-5	93	100	93	48
10-10	94	100	94	49
15-15	94	100	94	46
20-20	94	100	94	46
25-25	94	100	94	48
30-30	94	100	94	48
35-35	94	100	94	48
40-40	94	100	94	49
45-45	94	100	94	48
50-50	94	100	94	48
50-60	95	100	95	44
50-70	94	100	94	48
60-60	94	100	94	48
70-70	94	100	94	48
70-90	93	100	93	48
80-80	95	100	95	42
80-90	95	100	95	49
90-90	94	100	94	48
100-100	93	100	93	48

Table 5-9c The performance of the 19 artificial neural networks with three hidden layers on the test set ECGs for the diagnosis of anterior myocardial infarction.

Number of neurons in the hidden layers	Se	Sp	AI	LVH Sp
5-5-5	94	99	93	40
10-10-10	94	100	94	48
15-15-15	94	100	94	44
20-20-20	94	100	94	46
25-25-25	93	100	93	46
30-30-30	94	100	94	48
35-35-35	93	100	93	45
40-40-40	94	100	94	44
45-45-45	93	100	93	42
50-50-50	93	100	93	49
50-50-100	94	100	94	48
60-60-60	95	100	95	42
70-70-70	93	100	93	45
70-70-100	93	100	93	45
80-80-80	94	100	94	49
80-80-100	94	100	94	48
90-90-90	93	100	93	46
90-90-100	93	100	93	44
100-100-100	94	100	94	45

Table 5-10 The performance of the 30 neural networks using 25 input variables (QRS + dVCG) for the diagnosis of anterior myocardial infarction

Number of neurons in the hidden layers	Se	Sp	AI	LVH Sp
10	78	99	77	77
15	80	96	76	71
20	84	97	81	74
30	81	99	80	79
40	72	100	72	76
50	81	96	77	74
60	77	96	73	77
70	79	98	77	79
80	75	97	72	74
90	76	98	74	81
10-10	83	99	82	70
15-15	82	100	82	75
20-20	85	99	84	75
30-30	84	99	83	74
40-40	84	98	82	74
50-50	82	99	81	71
60-60	85	98	83	74
70-70	80	100	80	77
80-80	86	98	84	77
90-90	91	97	88	77
10-10-10	80	98	78	81
15-15-15	82	98	80	75
20-20-20	85	98	83	74
30-30-30	81	99	80	80
40-40-40	82	98	80	76
50-50-50	84	99	83	80
60-60-60	84	96	80	69
70-70-70	83	97	80	74
80-80-80	88	97	85	79
90-90-90	83	99	82	79

Table 5-11 The performance of the 30 neural networks using 21 input variables (QRS) for the diagnosis of anterior myocardial infarction

Number of neurons in the hidden layers	Se	Sp	AI	LVH Sp
10	81	99	80	73
15	77	100	77	79
20	78	99	77	75
30	74	99	73	86
40	79	100	79	77
50	76	100	76	80
60	74	99	73	82
70	75	99	74	76
80	72	100	72	82
90	71	99	70	81
10-10	78	98	76	74
15-15	89	98	87	79
20-20	83	99	82	77
30-30	88	99	87	80
40-40	85	99	84	81
50-50	83	100	83	80
60-60	84	100	84	82
70-70	81	100	81	83
80-80	82	100	82	76
90-90	81	100	81	79
10-10-10	79	100	79	80
15-15-15	84	98	82	83
20-20-20	84	99	83	79
30-30-30	81	100	81	82
40-40-40	84	100	84	85
50-50-50	82	99	81	83
60-60-60	88	99	87	80
70-70-70	86	99	85	81
80-80-80	81	99	80	80
90-90-90	82	100	82	80

Table 5-12a, b, & c The performance of the 45 neural networks using 21 input variables (QRS + ST-T + dVCG) and only a single output for the diagnosis of inferior myocardial infarction.

Table 5-12a The performance of the 12 neural networks with a single hidden layer for the diagnosis of inferior myocardial infarction.

Number of neurons in the hidden layer	Se	Sp	AI	LVH Sp
5	86	100	86	76
10	87	100	87	76
15	85	100	85	75
20	87	100	87	75
30	86	100	86	74
40	86	100	86	71
50	86	100	86	70
60	88	100	88	75
70	86	100	86	75
80	86	100	86	71
90	87	100	87	71
100	85	100	85	74

Table 5-12b The performance of the 12 neural networks with two hidden layers for the diagnosis of inferior myocardial infarction

Number of neurons in the hidden layers	Se	Sp	AI	LVH Sp
5-5	87	100	87	76
10-10	85	100	85	77
15-15	87	100	87	77
20-20	87	100	87	79
30-30	87	100	87	75
40-40	87	100	87	76
50-50	87	100	87	76
60-60	87	100	87	75
70-70	87	100	87	76
80-80	86	100	86	77
90-90	84	100	84	73
100-100	87	100	87	77

Table 5-12c The performance of the 21 neural networks with three hidden layers for the diagnosis of inferior myocardial infarction

Number of neurons in the hidden layers	Se	Sp	AI	LVH Sp
5-5-5	86	100	86	79
10-10-10	87	100	87	76
15-15-15	86	100	86	77
20-20-20	87	100	87	77
20-30-40	87	100	87	79
25-50-100	86	100	86	79
30-30-30	87	100	87	77
30-60-90	87	100	87	77
40-40-40	85	100	85	79
50-50-50	87	100	87	76
50-50-100	86	100	86	77
60-60-60	87	100	87	79
70-70-70	87	100	87	76
70-70-100	86	100	86	80
70-80-90	86	100	86	80
80-80-80	87	100	87	77
80-80-100	87	100	87	77
90-90-90	88	100	88	76
90-90-100	86	100	86	76
90-100-100	87	100	87	77
100-100-100	87	100	87	77

Table 5-13a, b, & c The performance of the 36 neural networks using 18 input variables (QRS + ST-T) and only a single output for the diagnosis of inferior myocardial infarction.

Table 5-13 a The performance of 12 neural networks with a single hidden layer for the diagnosis of inferior myocardial infarction

Number of neurons in the hidden layer	Se	Sp	AI	LVH Sp
5	85	100	85	71
10	85	100	85	73
15	85	100	85	73
20	88	100	88	68
30	85	100	85	70
40	85	100	85	71
50	86	100	86	71
60	87	100	87	71
70	84	100	84	71
80	87	100	87	73
90	86	100	86	71
100	86	100	86	69

Table 5-13b The performance of 12 neural networks with two hidden layers for the diagnosis of inferior myocardial infarction

Number of neurons in the hidden layers	Se	Sp	AI	LVH Sp
5-5	84	100	84	75
10-10	85	100	85	75
15-15	84	100	84	75
20-20	83	100	83	75
30-30	84	100	84	75
40-40	84	100	84	75
50-50	82	100	82	77
60-60	84	100	84	75
70-70	84	100	84	75
80-80	83	100	83	75
90-90	84	100	84	74
100-100	84	100	84	75

Table 5-13c The performance of 12 neural networks with three hidden layers for the diagnosis of inferior myocardial infarction

Number of neurons in the hidden layers	Se	Sp	AI	LVH Sp
10-10-10	84	100	84	71
20-20-20	84	100	84	76
30-30-30	84	100	84	70
40-40-40	84	100	84	75
50-50-50	84	100	84	74
50-50-100	84	100	84	75
60-60-60	83	100	83	75
70-70-70	84	100	84	75
80-80-80	85	100	85	75
80-80-100	84	100	84	75
90-90-90	84	100	84	76
100-100-100	83	100	83	77

Table 5-14 The performance of the 30 neural networks using 12 input variables (QRS + dVCG) for the diagnosis of inferior myocardial infarction.

Number of neurons in the hidden layers	Se	Sp	AI	LVH Sp
10	72	96	68	83
15	74	96	70	85
20	73	96	69	83
30	74	98	72	83
40	74	98	72	83
50	75	97	72	85
60	74	95	69	85
70	76	96	72	85
80	75	96	71	86
90	73	97	70	85
10-10	76	95	71	85
15-15	79	93	72	85
20-20	78	95	73	85
30-30	76	95	71	85
40-40	75	96	71	85
50-50	74	96	70	85
60-60	77	96	73	85
70-70	77	93	70	87
80-80	82	95	77	83
90-90	81	93	74	83
10-10-10	78	95	73	85
15-15-15	78	94	72	83
20-20-20	80	94	74	83
30-30-30	77	94	71	85
40-40-40	81	92	73	81
50-50-50	80	95	75	82
60-60-60	81	93	74	81
70-70-70	84	94	78	82
80-80-80	82	94	76	82
90-90-90	84	94	78	82

Table 5-15 The performance of the 30 neural networks using 9 input variables (QRS) for the diagnosis of inferior myocardial infarction.

Number of neurons in the hidden layers	Se	Sp	AI	LVH Sp
10	77	96	73	85
15	76	96	72	85
20	76	96	72	85
30	77	98	75	85
40	75	97	72	85
50	77	95	72	83
60	75	97	72	86
70	78	96	74	85
80	76	95	71	83
90	76	95	71	85
10-10	75	96	71	85
15-15	75	95	70	85
20-20	77	95	72	85
30-30	77	95	72	85
40-40	76	97	73	83
50-50	77	96	73	83
60-60	78	96	74	85
70-70	73	96	69	85
80-80	79	96	75	83
90-90	77	96	73	83
10-10-10	80	96	76	89
15-15-15	77	96	73	83
20-20-20	78	95	73	86
30-30-30	76	97	73	83
40-40-40	77	95	72	85
50-50-50	73	95	68	85
60-60-60	76	98	74	83
70-70-70	73	97	70	85
80-80-80	74	97	71	85
90-90-90	74	96	70	85

Table 5-20 Results of neural networks using 46 input variables (QRS+ST-T+dVCG) and only a single output neuron for the diagnosis of anterior myocardial infarction.

Number of neurons in the hidden layers	AI	Sensitivity	Specificity
20	66	73	93
30	81	85	96
40	80	85	95
50	80	84	96
20-20	72	80	92
30-30	81	85	96
*40-40	83	87	96
50-50	82	88	94
30-30-30	79	86	93

Table 5-21 Results of neural networks using 42 input variables (QRS+ST-T) and only a single output neuron for the diagnosis of anterior myocardial infarction.

Number of neurons in the hidden layers	AI	Sensitivity	Specificity
*20	82	88	94
30	79	86	93
40	77	85	92
50	78	86	92
20-20	79	87	92
30-30	77	86	91
40-40	78	85	93
50-50	78	85	93
30-30-30	78	85	93

Table 5-22 Results of neural networks using 25 input variables (QRS+dVCG) and only a single output neuron for the diagnosis of anterior myocardial infarction.

Number of neurons in the hidden layers	AI	Sensitivity	Specificity
20	66	73	93
30	68	74	94
40	69	75	94
50	71	77	94
20-20	72	80	92
30-30*	81	85	96
40-40	75	81	94
50-50	80	86	94
30-30-30	79	85	94

Table 5-23 Results of neural networks using 21 input variables (QRS) and only a single output neuron for the diagnosis of anterior myocardial infarction.

Number of neurons in the hidden layers	AI	Sensitivity	Specificity
20	72	78	94
30	75	77	98
40	73	79	94
50	73	75	98
20-20	76	79	97
30-30	75	81	94
40-40*	78	82	96
50-50	77	81	96
30-30-30	73	79	94

Table 5-25 Results of neural networks using 21 input variables (QRS+ST-T+dVCG) and only a single output neuron for the diagnosis of inferior myocardial infarction.

Number of neurons in the hidden layers	AI	Se	Sp
10	69	75	94
20*	70	76	94
30	69	75	94
40	69	75	94
50	69	74	95
10-10	68	74	94
20-20	67	73	94
30-30	68	73	95
40-40	67	73	94
50-50	67	73	94

Table 5-26 Results of neural networks using 18 input variables (QRS+ST-T) and only a single output neuron for the diagnosis of inferior myocardial infarction.

Number of neurons in the hidden layers	AI	Se	Sp
10	72	78	94
20	70	76	94
30	69	76	93
40	70	76	94
50	70	76	94
10-10*	73	79	94
20-20	70	76	94
30-30	70	76	94
40-40	72	78	94
50-50	70	76	94

Table 5-27 Results of neural networks using 12 input variables (QRS+dVCG) and only a single output neuron for the diagnosis of inferior myocardial infarction.

Number of neurons in the hidden layers	AI	Se	Sp
10	53	61	92
20	51	64	87
30	65	75	90
40	58	68	90
50	60	68	92
10-10	56	64	92
20-20	60	69	91
30-30	56	65	91
40-40	61	70	91
50-50*	62	70	92

Table 5-28 Results of neural networks using 9 input variables (QRS) and only a single output neuron for the diagnosis of inferior myocardial infarction.

Number of neurons in the hidden layers	AI	Se	Sp
10	63	70	93
20	62	70	92
30	61	70	91
40	61	69	92
50	61	70	91
10-10	58	65	93
20-20	57	65	92
30-30	64	70	94
40-40	62	69	93
50-50*	67	74	93

Table 5-30 Results of neural networks using 46 input variables (QRS+ST-T+dVCG) and three output neurons for the diagnosis of anterior myocardial infarction.

	AMI		LVH		Normal
	Se	Sp	Se	Sp	Sp
10	86	95	71	88	89
*20	84	96	76	88	89
30	84	96	69	88	90
40	86	96	69	88	89
50	86	96	69	88	88
20-20	88	95	74	89	87
30-30	87	94	64	91	92
40-40	86	95	64	91	94
50-50	85	95	74	87	85
30-30-30	88	92	69	90	87

Table 5-31 Results of neural networks using 42 input variables (QRS+ST-T) and three output neurons for the diagnosis of anterior myocardial infarction.

	AMI		LVH		Normal
	Se	Sp	Se	Sp	Sp
10	87	92	55	91	94
20	85	94	64	89	90
30	86	91	64	91	89
40	88	93	60	92	93
50	88	92	60	92	93
20-20	85	94	67	91	93
30-30	88	95	69	91	92
40-40	90	92	62	93	93
*50-50	86	93	67	92	93
30-30-30	90	92	50	95	95

Table 5-32 Results of neural networks using 25 input variables (QRS+dVCG) and three output neurons for the diagnosis of anterior myocardial infarction.

	AMI		LVH		Normal
	Se	Sp	Se	Sp	Sp
*10	84	96	71	85	74
20	78	95	69	83	73
30	78	93	57	86	75
40	79	95	64	85	78
50	81	94	62	83	72
20-20	83	93	60	83	70
30-30	83	96	67	84	75
40-40	82	94	57	84	70
50-50	78	94	52	78	68
30-30-30	83	94	62	88	79

Table 5-33 Results of neural networks using 21 input variables (QRS) and three output neurons for the diagnosis of anterior myocardial infarction.

	AMI		LVH		Normal
	Se	Sp	Se	Sp	Sp
10	77	96	50	79	64
20	75	94	52	87	79
30	75	98	57	81	72
40	74	94	55	80	73
50	75	97	62	85	78
*20-20	83	94	45	87	81
30-30	77	95	64	76	60
40-40	74	93	48	79	66
50-50	82	94	50	82	69
30-30-30	80	96	67	75	56

Table 5-35 Results of neural networks using 21 input variables (QRS+ST-T+dVCG) and three output neurons for the diagnosis of inferior myocardial infarction.

	IMI		LVH		Normal
	Se	Sp	Se	Sp	Sp
10	79	94	64	86	84
20	75	94	67	91	93
30	78	94	64	91	92
40	78	95	64	88	90
50*	75	94	67	93	94
10-10	74	94	62	90	94
20-20	78	95	64	89	92
30-30	75	96	71	88	90
40-40	74	94	62	87	88
50-50	70	95	64	86	93

Table 5-36 Results of neural networks using 18 input variables (QRS+ST-T) and three output neurons for the diagnosis of inferior myocardial infarction.

	IMI		LVH		Normal
	Se	Sp	Se	Sp	Sp
10	78	94	55	93	95
20	78	93	60	87	86
30	78	94	60	92	93
40	80	94	57	93	93
50	76	94	60	93	95
10-10	69	94	64	91	98
20-20	74	95	64	86	87
30-30	75	95	67	91	92
40-40	70	95	64	89	90
50-50*	79	95	71	89	90

Table 5-37 Results of neural networks using 12 input variables (QRS+dVCG) and three output neurons for the diagnosis of inferior myocardial infarction.

	IMI		LVH		Normal
	Se	Sp	Se	Sp	Sp
10	73	89	38	86	74
20	68	89	19	84	73
30	71	91	43	83	74
40	68	90	43	86	76
50	65	92	43	78	70
10-10*	69	94	64	91	98
20-20	66	91	45	79	70
30-30	68	90	45	78	67
40-40	73	91	38	78	65
50-50	65	93	40	84	82

Table 5-38 Results of neural networks using 9 input variables (QRS) and three output neurons for the diagnosis of inferior myocardial infarction.

	IMI		LVH		Normal
	Se	Sp	Se	Sp	Sp
10	70	89	21	91	76
20	70	92	33	83	71
30*	70	94	36	88	84
40	66	90	31	87	76
50	61	92	31	83	74
10-10	68	94	38	81	76
20-20	65	91	31	81	69
30-30	66	93	38	84	78
40-40	71	92	31	86	76
50-50	74	93	43	82	75

APPENDIX 3
SUPPLEMENTARY TABLES FOR CHAPTER 6

Table 6-2 The deterministic criteria for the diagnosis of Q waves in ANTEROSEPTAL (V2, V3, V4), ANTERIOR (V3, V4), OR SEPTAL (V2) leads in the original Glasgow Program. (VQ1= definitive Q criteria, VQ2= Atypical Q criteria, VQ3= low R wave criteria, VQ4= Reversed R wave progression criteria, QRVH= right ventricular hypertrophy criteria, PRWP= poor R wave progression criteria)

VQ1	(a)	(1) |Q| > 0.2 mV
	and	(2) Q > 0.03 secs
	and	(3) QRS peak to peak amp>0.2mV
or	(b)	(1) R amp= 0
	and	(2) |S| > 0.2 mV
	and	(3) S > 0.03 secs
	and	(4) QRS peak to peak amp>0.2mV
VQ2	(a)	(1) |Q| > 0.10 mV
	and	(2) Q > 0.015 secs
	and	(3) |Q/R| > 1/4 or R=0
	and	(4) QRS peak to peak amp>0.2mV
or	(b)	(1) R < 0.065 mV
	and	(2) |S| > 0.14 mV
	and	(3) S > 0.015 secs
	and	(4) |S/R'| > 1/4
VQ3	(a)	(1) R < 0.11 mV
	and	(2) R' < 2R amp, or RBBB, or IVCD of RBBB type is present.
	and	(3) |R/S| < 0.125
	and	(4) QRS peak to peak amp>0.2mV
	and	(5) RVH is not present
VQ4		(1) RV(n)-RV(n+1)>0.02mV
		(2) R<0.3mV in those 2 leads
		(3) R' < R in those 2 leads

QRVH	(a)	(1) R>0.3mV with S=0 or
		R<0.1mV with R' > 0.3mV
	and	(2) RBBB or IVCD or IVCD
		of RBBB type are not present
	and	(3) STV2<0.12mV or ST<1/2 T+
or	(b)	(1) R<0.3mV or S is not 0mV
	and	(2) S or S'<-0.5 mV in lead I
	and	(3) there is a clinical classification of
		congenital heart disease --- etc.
	and	(4) RBBB or IVCD of RBBB type
		not present.
PRWP		(1) RV3<0.3mV & R'V3<0.3mV
	and	(2) None of VQ1-VQ4 is true.

Table 6-5 Traditional deterministic Q wave criteria for the diagnosis of inferior myocardial infarction inside the original Glasgow Program. (All the criteria are definite Q waves)

Q1	a	(1) Q>35mS, and |Q/R|>20%
		or(2) Q>40 mS
and	b	|Q|>0.09 mV
and	c	peak-peak QRS>0.15mV
Q2	a	(1) Q>35mS, with |Q/R|>20%
		or(2) Q>30mS, with |Q/R|>1/3
and	b	|Q|>0.2mV
and	c	peak-peak QRS>0.15mV
Q3	a	Q>26mS or |Q/R|>20%
and	b	|Q|>0.14 mV
and	c	peak-peak QRS>0.15mV

Q4	a	(1) Q>30mS and T- <-0.1mV
		or(2) |Q/R|>1/3 and Q>20mS
and	b	|Q|>0.075mV
and	c	peak to peak QRS>0.2mV
and	d	(1) T- < -0.05 mV
		or(2) ST > 0.06 mV

Table 6-6 Non-Q criteria for the diagnosis of inferior myocardial infarction preserved inside the original deterministic Glasgow Program

A.			(1) |Q/R|>25% in lead II and |Q|>0.1 mV,
	and		(2) Frontal QRS axis<0
	and		(3) Age > 20 years old
B.	(1)		(i)R amplitude II < R amplitude III
		and	(ii) QRS axis not greater than -30°
		and	(iii) R amplitude III < 0.20 mV
or	(2)		(i) Q > 0.015 seconds & R<0.1mV & S>0.02 seconds in aVF
		and	(ii) peak to peak QRS > 0.15 mV in aVF
C.			(1) T axis < -10° in frontal plane
			(2) R in lead II < 0.90 mV
			(3) |Q/R| > 25% in any 2 of II, III, and aVF

REFERENCES

Abildskov, J.A. (1987)
Prediction of ventricular arrhythmias from ECG waveforms.
Journal of Electrocardiology, 20(suppl), 97-101.

AIMS Trial Study Group (1988)
Effect of intravenous APSAC on mortality after acute myocardial infarction: preliminary report of a placebo-controlled clinical trial.
Lancet, 1, 545-549.

Aleksander, I. & Burnett, P. (1987)
Thinking Machines. The search for artificial intelligence.
Oxford University Press.

American Heart Association Committee on Electrocardiography (1975)
Recommendation for standardization of leads and of specifications for instruments in electrocardiography and vectorcardiography.
Circulation, 52, 11-31.

Bailey, J.J. & Horton, M.R. (1977)
Advantages of automation of ECG analysis with conventional (heuristic) criteria.
In: Trends in Computer-Processed Electrocardiograms: ed. van Bemmel, J.H. Willems, J.L. pp 221-228. Amsterdam, North Holland.

Baxt, W.G.(1991a)
Use of an artificial neural network for the diagnosis of myocardial infarction.
Annals of Internal Medicine, 115, 843-848.

Baxt, W.G.(1991b)
Use of an artificial neural network for data analysis in clinical decision making: The diagnosis of acute coronary occlusion.
Neural Computing, 2, 480-489.

Bjerle, P. & Arvedson, O.(1986)
Comparison of Frank vectorcardiogram with two different vectorcardiograms derived from conventional ECG leads .
In, Proceedings of Engineering Foundation Conference, 11, 13-26.

Bonner, R.E., Crevasse, L., Ferrer, M.I. & Greenfield, J.C. Jr. (1972)
A new computer program for analysis of scalar electrocardiograms.
Computers in Biomedical Research, 5, 629-653.

Boone, J.M., Gross, G.W.& Greco-Hunt, V. (1990)
Neural networks in radiologic diagnosis.
Investigational Radiology, 25, 1012-1023.

Bortolan, G., Degani, R. & Willems, J.L. (1991)
Neural networks for ECG classification.
In Computers in Cardiology 1990, ed. Ripley, K.L. & Murray, A. pp269-272. Long Beach, CA: IEEE Computer society.

Bortolan, G., Degani, R. & Willems, J.L. (1992)
ECG classification with neural networks and cluster analysis..
In Computers in Cardiology 1991, ed. Ripley, K.L. & Murray, A. pp177-180. Long Beach, CA: IEEE Computer society.

Branscombe, M. (1990)
Modelling the brain.
Program now, June, 28-30.

Brohet, C.R., Robert, A., Derwael, C., Fesler, R., Stijn, M., Vliers, A. & Braasseur, L.A. (1984)
Computer interpretation of Pediatric orthogonal electrocardiograms: statistical and deterministic classification methods.
Circulation, 70, 255-263.

Brohet, C.R. (1991)
Special value of the vectorcardiogram in pediatric cardiology.
Journal of Electrocardiology , 23, supplement, 58-62.

Burch, G.E. (1985)
The history of vectorcardiography.
Medical History , 5, 103-131.

Burchell, H.B. (1987)
A centennial note on Waller and the first human electrocardiogram.
American Journal of Cardiology, 59, 979-983.

Burdon Sanderson, J. & Page, F.J.M. (1878)
Experimental results relating to the rhythmical and excitatory motions of the ventricle of the heart of the frog and of the electrical phenomena which accompany them.
Proceedings of Royal Society of London, 27, 410-414.

Burger, H.C., van Brummelen, A.G.W. & van Herpen, G. (1962)
Compromise in vectorcardiography. Alteration of coefficients as a means of adapting one lead system to another.
American Heart Journal , 63, 666-678.

Caceres, C.A., Steinberg, C.A. & Abraham, S. (1962)
Computer extraction of electrocardiographic parameters.
Circulation, 25, 356-362.

Cady, L.D.Jr., Woodbury, M.A., Tick, L.J. & Gertler, M. (1961)
A method for electrocardiogram wave pattern estimation.
Circulation Research, 9, 1078-1082.

Casaleggio, A., Morando, M. & Ridella, S. (1991)
Neural network for automated anomalous QRS complexes detection.
In Ripley, K. L. & Murray, A. (Eds.) Computers in Cardiology 1990, 553-556. IEEE Computer Society, Washington.

Cerebral Embolism Task Force (1986)
Cardiogenic brain embolism.
Archive of Neurology, 43, 71-84.

Cerebral Embolism Task Force (1989)
Cardiogenic brain embolism.The second report.
Archive of Neurology, 46, 727-743.

Chen, C.Y., Chiang, B.N. & Macfarlene, P.W. (1989)
Normal limits of electrocardiogram in a Chinese population.
Journal of Electrocardiology, 22, 1-15.

Chou, T.C. & Helm, R.A. (1967)
Clinical vectorcardiography.
Grunne and Stratton Inc: New York and London.

Chou, T.C. (1986)
When is the vectorcardiogram superior to the scalar electrocardiogram ?
Journal of American College of Cardiology, 8, 791-799.

Churchland, P.M. & Churchland, P.S. (1990)
Could a machine think?
Scientific American, 262, 32-37.

Cooper, J.K. (1986)
Electrocardiography 100 years ago.
New England Journal of Medicine , 315, 461-464.

The Common Standards for Quantitative Electrocardiography Working Party (1985)
Recommendations for measurement standards in quantitative electrocardiography.
European Heart Journal, 6, 815-825.

Cubanski, D., Cyganski, D., Hubelbank, M., Antman, E. & Feldman, C. (1993)
A neural network system for automatic detection of atrial fibrillation.
Journal of American College of Cardiology, 21, 182A (Abstract)

Dassen, W.R.M., Mulleneers, R.G.A., den Dulk, K., Smeets, J.R.L.M., Cruz, F., Penn, O.C.K.M. & Wellens, H.J.J.(1990)
An artificial neural network to localize atrioventricular accessory pathways in patients suffering from the Wolff-Parkinson-White syndrome.
PACE, 13, 1792-1796.

Dassen, W.R.M., Mulleneers, R.G.A., den Dulk, K. & Wellens, H.J.J. (1993)
The influence of initial parameters on the learning behavior and performance of an artificial neural network.
Journal of Electrocardiology, 25(supp), 26.

Degani, R. & Bortolan, G. (1986)
Combining measurement precision and fuzzy diagnostic criteria.
In Computer ECG analysis: Towards standardisation, ed. Willems, J.L. van Bemmel, J.H. Zywietz, C. pp177-182. Amsterdam, North Holland.

DePace, N.L., Colby, J., Hakki, A., Manno, B., Horowitz, L.N. & Iskandrian, A.S. (1983)
Poor R wave progression in the precordial leads: Clinical implication for the diagnosis of myocardial infarction.
Journal of American College of Cardiology, 2, 1073-1079.

Devine, B. (1990)
Neural Networks in Electrocardiography.
MSc thesis.University of Glasgow.

Devine, B. & Macfarlane, P.W. (1993)
Detection of Electrocardiographic "left ventricular strain" using neural nets.
Medical & Biological Engineering & Computing, 31, 343-348.

Dower, G.E. (1968)
A lead synthesizer for the Frank system to simulate the standard 12-lead electrocardiogram.
Journal of Electrocardiology, 1,101-116.

Dower, G.E., Machado, H.B. & Osborne, J.A. (1980)
On deriving the electrocardiogram from vectorcardiographic leads.
Clinical Cardiology, 3,87-95.

Draper, H.W., Peffer, C.J., Stallmann, F.W., Littmann, D. & Pipberger, H.V. (1964)
The corrected orthogonal electrocardiogram and vectorcardiogram in 510 normal men (Frank lead system).
Circulation, 30, 853-864.

Drazen, E., Mann, N., Borun, R., Laks, M. & Berson, A. (1988)
Survey of computer assisted electrocardiography in the United States.
Journal of Electrocardiology, 21 (supplement), 98-104.

Dunn, M., Alexander, J., de Silva, R. & Hildner ,F. (1989)
Antithrombotic therapy in atrial fibrillation.
Chest, 95, 118S-127S.

Dytch, H.E. & Wied, G.L. (1990)
Artificial neural networks and their use in quantitative pathology.
Analytic Quantitative Cytology and Histology, 12, 379-393.

Edenbrandt, L., Johnson, B., Lundh, B. & Pahlm, O. (1987)
Sex and age related normal limits for QRS complex in vectorcardiography.
Clinical Physiology, 7,525-536.

Edenbrandt, L. & Pahlm, O. (1988a)
Comparison of various methods for synthesizing Frank-like vectorcardiograms from the conventional 12-lead ECG.
In Proceedings of computers in cardiology, Leuven 1987. pp71-74. Washington:IEEE computer society.

Edenbrandt, L. & Pahlm, O. (1988b)
Vectorcardiogram synthesized from a 12-lead ECG : superiority of the inverse Dower matrix.
Journal of Electrocardiology, 21, 361-367.

Edenbrandt, L., Ek, A., Lundh, B. & Pahlm O. (1989)
Vectorcardiographic bites.
Journal of Electrocardiology, 22, 325-331.

Edenbrandt, L., Pahlm, O., Lyttkens, K. & Albrechtsson, U. (1990)
Improved ECG interpretation using synthesized VCG for the diagnosis of inferior myocardial infarction.
Journal of Electrocardiology, 23, 207-211.

Edenbrandt, L., Devine, B. & Macfarlane, P. W.(1992)
Neural networks for classification of ECG ST-T segments.
Journal of Electrocardiology, 25, 167-173.

Einthoven, W. (1903)
Die galvanometrische Registrirung des menschlichen Elektrokardiogramms, zugleich eine Beurtheilung der Anwendung des Capilar-Elektrometers in der Physiologie.
Pfluegers Arch, 99, 472-480.

Einthoven, W. (1912)
The different forms of the human electrocardiogram and their signification.
The Lancet, 853-861.

Einthoven, W., Fahr, G. & de Waart, A. (1913)
Uber die Richtung und die manifeste Grosse der Potentialschwankungen im menschlichen Herzen und uber den Einfluss der Herzlage auf die Form des ElektroKardiogramms.
Pfluegers Arch., 150, 275-315.
(Translation: Hoff, H.E.& Sekelj, E.(1950) A.H.J. 40, 163-194)

Erb, R.J. (1993)
Introduction to backpropagation neural network computation.
Pharmaceutical Research, 10, 165-170.

Fisch, C. (1980)
The clinical electrocardiogram: A classic.
Circulation , 62, 1-4.

Frank, E. (1954)
The image surface of a homogeneous torso.
American Heart Journal , 47, 757-768.

Frank, E. (1956)
An accurate clinically practical system for spatial vectorcardiography.
Circulation , 13 , 737-749.

Geselowitz, D.B. (1960)
Multipole representation for an equivalent cardiac generator.
Proceedings of IRE, 48, 75-79.

Goldman, L., Weisberg, M., Olshen, R., Cook, E.F., Sargent, R.K., Lamas, G.A., Dennis, C., Wilson, C., Deckelbaum, L., Fineberg, H., Stiratelli, R. & the Medical House Staffs at Yale-New Haven Hospital and Brigham and Woman's Hospital. (1982)
A computer-derived protocol to aid in the diagnosis of emergency room patients with acute chest pain.
New England Journal of Medicine, 318, 797-803.

Gruppo Italiano per lo Studio della Streptochinasi nell'infarcto Miocardico (GISSI)(1987)
Long-term effects of intravenous thrombolysis in acute myocardial infarction: final report of the GISSI study.
Lancet, 2, 871-874.

Goldberger, E. (1942)
A simple, indifferent, electrocardiographic electrode of zero potential and a technique of obtaining augmented, unipolar, extremity leads.
American Heart Journal, 23,483-492.

Heden, B., Edenbrandt, L., Haisty, W.K. & Pahlm, O. (1993)
Neural networks for ECG diagnosis of myocardial infarction.
Computers in Cardiology 1993, 96 (abstract)

Hecht-Nielsen, R. (1987)
Komogorov's mapping neural network existence theorem.
In Proceedings of First IEEE International Joint Conference of Neural Networks, San Diego, CA, July 21-24, PP III-11-III-14.

Hecht-Nielsen, R. (1990)
Neurocomputing.
Addison-Wesley, Reading, MA.

Hoffman, I. & Hamby, R.I. (1971)
Vectorcardiography 2.
North-Holland Publishing Company: Amsterdam & London.

Hoffman, I. & Hamby, R.I. (1976)
Vectorcardiography 3.
North-Holland Publishing Company: Amsterdam & Oxford.

Hollman, A. (1981)
Thomas Lewis- The early years.
British Heart Journal, 46, 233-244.
Hornik, K., Stinchombe, M. & White, H. (1989)
Multi-layer feedforward networks are universal approximators.
Neural Networks, 2, 359-366.

Hurd, P.H., Sterling, M.R., Crawford, M.H., Dlabal, P.W. & O'Rourke, R.A. (1981)
Comprarative accuracy of electrocardiographic and vectorcardiograpfic criteria for inferior myocardial infarction.
Circulation, 63, 1025-1029.

Huwez, F.U. (1990)
Electrocardiography of the Left Ventricle in Coronary Artery Disease and Hypertrophy.
PhD Thesis, University of Glasgow.

ISIS-2 (Second International Study of Infarct Survival) Collaborative Group (1988)
Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17187 cases of suspected acute myocardial infarction.
Lancet, 2, 349-360.

Iwata, A., Nagasaka, Y. & Suzumura, N. (1990)
Data compression of the ECG using neural network for digital Holter monitor.
IEEE Engineering in Medicine and Biology Megazine, 9, 53-57.

Jones, W.P. & Hoskins, J. (1987)
Back propagation: a generalized delta learning rule.
Byte, Oct, 155-162.
Kannel, W.B., Abbott, R.D., Savage, D.D. & McNamara, P.M. (1982)
Epidemiologic features of chronic atrial fibrillation: the Framingham study.
New England Journal of Medicine, 306, 1018-1022.

Kannel, W.B. & Abbott, R.D. (1984)
Incidence and prognosis of unrecognised myocardial infarction: an update on the Framingham Study.
New England Journal of Medicine, 311, 1144-1147.

Kitzman, W.D., Scholz, D.G., Hagen, P.T., Ilstrup, D.M. & Edwards, W.D. (1988)
Age-related changes in normal human hearts during the first 10 decades of life. Part II (Maturity): A quantitative anatomic study of 765 specimens from subjects 20-99 years old. Mayo Clin Proc, 63, 137-146.

Kolen, J.F. & Pollack, J.B. (1990)
Back propagation is sensitive to initial conditions.
Complex Systems, 4, 269-280.

Kolliker, A. & Muller, H. (1856)
Nachweis der negative Schwankung des Muskelstroms am naturlich sich contrahirenden Muskel.
Verh. Phys. Med. Ges. Wurzburg, 6, 528-533.

Kors, J.A., Talmon, J.L. & van Bemmel, J.H. (1986)
Multilead ECG analysis.
Computers in Biomedical Research, 19, 28-46.

Kors, J.A. & van Bemmel, J.H. (1990a)
Classification methods for computerised interpretation of the electrocardiogram.
Methods of Information in Medicine, 29, 330-336.

Kors, J.A., van Herpen, G., Sittig, A.C. & van Bemmel, J.H. (1990b)
Reconstruction of the Frank vectorcardiogram from standard electrocardiographic leads : diagnostic comparison of different methods.
European Heart Journal , 11, 1083-1092.

Kors, J.A., van Herpen, G., Willems, J.L. & van Bemmel, J.H. (1992)
Improvement of automated electrocardiographic diagnosis by combination of computer interpretations of the electrocardiogram and vectorcardiogram.
American Journal of Cardiology, 70, 96-99.

Kossmann, C.E. & Johnston, F.D. (1935)
The precordial electrocardiogram.
American Heart Journal, 10, 925-941.

Laks, M.M., Morady, F. & Swan, H.J. (1969)
Canine right and left ventricular cell and sarcomere lengths after banding the pulmonary artery.
Circulation Research, 24, 705-710.

Le Cun, Y. (1985)
Une procedure d'apprentissage pour reseau a seuil assymetrique.
in Proceedings of cognitiva, Paris.

Lempert, G.L. (1976)
From electrocardiography to electrocardiology.
Advances in Cardiology, 16, 57-61.

Lewis, T. (1916)
The spread of the excitatory process in the vertebrate heart.
Phil. Transaction of Royal Society of London, Ser. B, 207, 221-310.

Liebman, J., Lee, M.H., Rao, P.S. & Mackay, W. (1973)
Quantitation of the normal Frank and McFee-Parungao orthogonal electrocardiogram in the adolescent.
Circulation, 48, 735-752.

Lippmann, R.P. (1987)
An introduction to computing with NeuralNets.
IEEE ASSP Megazine, Apr, 4-22.

Lui, C.Y., Ornato, J.P., Buell, J.C., Grissom, R.L. & Ericksn, C.C. (1987)
Lack of superiority of the vectorcardiogram over the electrocardiogram in detecting inferior wall myocardial infarction regardless of time since infarction.
Journal of Electrocardiology, 20, 241-246.

Lyon, A.F. & Belletti, D.A. (1968)
The Frank vectorcardiogram in normal men.
British Heart Journal, 1968,30,172-181.

Macfarlane, P.W., Lorimer, A.R. & Lawrie, T.D.V. (1971a)
Three and twelve lead electrocardiogram interpretation by computer.
A compaison on 1093 patients.
British Heart Journal, 33, 266.

Macfarlane, P.W. (1971b)
ECG waveform identification by digital computer.
Cardiovascular Research, 5, 141-146.

Macfarlane, P.W. & Lawrie, T.D.V. (1974)
An introduction to automated electrocardiogram interpretation.
London: Butterworth & Co.

Macfarlane, P.W. (1979)
A hybrid lead system for routine electrocardiography.
In Optimization of computer ECG processing,
eds. Wolf, H.K. & Macfarlane, P.W. pp1-5. Tunbridge Wells: Pitman Medical.

Macfarlane, P.W., Watts, M.P., Podolski, M., Shoat, D. & Lawrie, T.D.V. (1986a)
The new Glasgow system.
In Computer ECG analysis - toward standardization,
eds. Willems, J.L. van Bemmel, J.H. & Zywietz, C. pp31-36. Amsterdam: North - Holland.

Macfarlane, P.W. (1986b)
Computer interpretation of cardiac rhythm.
In Computer ECG analysis - toward standardization. eds. Willems, J.L., van Bemmel, J.H. & Zywietz, C. pp279-284. Amsterdam: North - Holland.

Macfarlane, P.W., Chen, C.Y. & Chiang, B.N. (1988)
Comparison of the ECG in apparently healthy Chinese and Caucasians.
In Computers in Cardiology 1988, pp143-146. IEEE Proceedings.

Macfarlane, P.W. & Lawrie, T.D.V. (1989)
Comprehensive Electrocardiology, vol 1, 2, &3.
Oxford: Academic Press.

Macfarlane, P.W. (1990a)
A brief history of computer-assisted electrocardiography.
Methods of Information in Medicine , 20,272-281.

Macfarlane, P.W., Devine, B., Latif, S., McLaughlin, S., Shoat, D.B. & Watts, M.P. (1990b)
Methodology of ECG interpretation in the Glasgow Program.
Methods of Information in Medicine ,29, 354-361.

Macfarlane, P.W. & Edenbrandt, L. (1992)
12-lead vectorcardiography in ischaemic heart disease.
Journal of Electrocardiology,24, supplement, 188-193.

Mann, H. (1920)
A method of analyzing the electrocardiogram.
Archive of Internal Medicine, 25, 283-294.

Mann, H. (1938)
The monocardiograph.
American Heart Journal, 5, 681-699.

Maren, A.J., Harston, C.T. & Pap, R.M. (1991)
Handbook of neural computing applications.
Academic Press, San Diego, CA.

Marey, E.J. (1876)
Des variations electriques des muscles et du coeur enparticulier etudies au moyen de l'electrometre de M. Lippmann.
C. R. Academy Science, 82, 975-977.

Meijler, F.L. (1983)
Atrial fibrillation: a new look at an old arrhythmia.
Journal of American College of Cardiology, 2, 391-393.

Minsky, M.L. & Papert, S.A. (1969)
Perceptron: an introduction to computational geometry.
Cambridge,Massachusetts: MIT Press.

Mizuno, Y. (1966)
Normal limit and variability of electrocardiographic items of the Japanese.
Japanese Circulation Journal, 30,357-360.

Morikawa, J., Kitamura, K., Habuchi, Y., Tanaka, N., Nishimoto, Y., Hirano, M., Tsujimura, Y., Hamamoto, H. & Takanashi, T. (1987)
Three-dimensional vectorcardiography (3-D VCG) by computer graphics in old myocardial infarction.
Angiology, 38, 449.

Mulsant, B.H. (1990)
A neural network as an approach to clinical diagnosis.
MD Computing, 7, 25-36.

Nemati, M., Doyle, J.T., McCaughan, D., Dunn, R.A. & Pipberger, H.V. (1978)
The orthogonal electrocardiogram in normal women. Implication of sex differences in diagnostic electrocardiography.
American Heart Journal,95,12-21.

Ostrander, L.D., Brandt, R.L., Kjelsberg, M.O. & Epstein, F.H. (1965)
Electrocardiographic findings among the adult population of a total natural community, Tecumseh, Michigan.
Circulation, 31, 889

Pahlm, O., Edenbrandt, L., Heden, B. & Sornmo, L. (1994)
Derived Vectorcardiogram - A useful adjunct to the conventional 12-lead ECG (Abstract).
XXI International Congress on Electrocardiology, Yokohama, Japan.

Parker, D.B. (1985)
Learning-logic, TR-47, Center for computational research in economics and management science, MIT.
Patrick, E.A., Margolin, G., Sanghvi, V. & Uthurusamy, R. (1976)
Pattern recognition applied to early diagnosis of heart attacks.
In Proceedings of the IEEE 1976 Systems, Man, and Cybernetics Conference, Washington D.C., November 1-3, pp 403-406.

Patrick, E.A., Margolin, G., Sanghvi, V. & Uthurusamy, R. (1977)
Pattern recognition applied to early diagnosis of heart attacks.
In Proceedings of the 1977 International Medical Information Processing Conference (MEDINFO), Toronto, August 9-12, pp 203-207.

Pipberger, H.V., Freis, E.D., Taback, L. & Mason, H.L. (1960)
Preparation of electrocardiographic data for analysis by digital electronic computer.
Circulation, 21, 413-418.

Pipberger, H.V., Goldman, M.J., Littmann, D., Murphy, G.P., Cosma, J. & Snyder, J.R. (1967)
Correlations of the orthogonal electrocadiogram and vectorcardiogram with constitutional variables in 518 normal men.
Circulation, 35, 536-551.

Pipberger H.V. (1968)
New quantitative vectorcardiographic criteria.
American Heart Journal, 76, 717.

Pipberger, H.V., McCaughan, D., Littmann, D., Pipberger, H.A., Cornfield, J., Dunn, R.A., Batchlor, C.D. & Berson, A.S. (1975)
Clinical application of a second generation electrocardiographic computer program.
American Journal of Cardiology, 35, 597-608.

Pipberger, H.V., Simonson, E. & Lopez, E.A.Jr. (1982)
The electrocardiogram in epidemiologic investigation: a new classification system.
Circulation, 65, 1456.

Pitts, W. & McCulloch, W.S. (1943)
A logical calculus of the ideas imminent in nervous activity.
Bulletin of Mathematical Biophysics , 5, 115-133.

Pozen, M.W., Stechmiller, J.K. & Voigt, G.C. (1977)
Prognostic efficacy of early clinical categorization of myocardial infarction patients.
Circulation, 56, 816-819.

Pozen, M.W., D'Agostino, R.B., Mitchell, J.B., Rosenfeld, D.M., Guglielmino, J.M., Schwartz, M.L., Teebagy, N., Valentine, J.M. & Hood, W.B. (1980)
The usefulness of a predictive instrument to reduce inappropriate admissions to the coronary care unit.
Annals of Internal Medicine, 92, 238-242.

Pringle, S.D. (1990)
An Investigation into the Mechanisms Responsible for Sudden Death in Hypertensive Patients with the Electrocardiographic Pattern of Left Ventricular Hypertrophy and Strain.
MD Thesis, University of Dundee.

Rahimtoola, S.H. (1989)
The hibernating myocardium.
American Heart Journal, 117, 211-221.

Recke, S. (1986)
Diagnostic implications of figure-of-eight and clockwise QRS loop rotation on the horizontal vectorcardiogram in chronic aortic valve disease.
Journal of Electrocardiology, 19, 123-130.

Reddy, M.R.S., Edenbrandt, L., Svensson, J., Haisty, W.K. & Pahlm, O. (1993)
Neural network versus Electrocardiographer and conventional computer criteria in diagnosing anterior infarct from the ECG.
Computers in Cardiology 1993.

Riff E.R. & Riff K.M. (1974)
Abnormalities of myocardial depolarization in overt, subclinical and prediabetes. A vectorcardographic study.
Diabetes, 23, 572.

Rijlant, P. (1936)
Introduction a l'etude de la distribution spatiale des variations de potentiel produites par le coeur chez l'homme.
C. R. Seances Soc. Biol. 121, 1358-1361.

Rijlant, P. (1980)
The coming of age of electrophysiology and electrocardiography.
NTM-Schriftenr. Gesch. Naturwiss. Tech. Med., 17, 108-123.

Rosenblatt, F. (1958)
The perceptron: A probabilistic model for information storage and organization in the brain.
Psychological Review , 65, 386-408.

Rubel, P., Benhadid,I. & Fayn, J. (1992)
Quantitative assessment of eight different methods for synthesizing Frank VCGs from simultaneously recorded standard ECG leads.
Journal of Electrocardiology, 24, supplement, 197-202.

Rumelhart, D.E., Hinton, G.E. & Williams, R.J. (1986)
Learning internal representation by error propagation.
In Parallel distributed processing: Explorations in the microstructures of cognition. ed. Rumelhart, D.E. & McClelland, J.L. Vol. 1. pp 318-362. Cambridge,Massachusetts: MIT Press.

Schellong, F (1936)
Elektrographische diagnostik der herzmuskelerkrankungen.
Verh. Dtsch. Ges. Inn. Med. 48, 288-310.

Searle, J. (1990)
Is the brain's mind a computer program?
Scientific American, 262, 26-31.

Selvester, R.H., Rubin, H.B., Hamlin, J.A. & Pote, W.W. (1968)
New quantitative vectorcardiographic criteria for the detection of unsuspected myocardial infarction in diabetics.
American Heart Journal, 75, 335-348.

Selvester, R.H., Palmersheim, J. & Pearson, R.B. (1971)
VCG inverse model for the prediction of myocardial disease.
In Hoffman I ed. Vectorcardiography 2, 54-65. North Holland, Amsterdam.

Simonson, E., Nakagawa, K. & Schmitt, O.H. (1957)
Respiratory changes of the spatial vectorcardiogram recorded with different lead systems.
Amercan Heart Journal, 54, 919-939.

Simonson, E., Blackburn, H.Jr., Puchner, T.C., Ribeiro, F. & Meja, M. (1960)
Sex differences in the electrocardiogram.
Circulation, 22,598.

Simonson, E. (1961)
Differentiation between normal and abnormal in electrocardiography.
St. Louis : C.V.Mosby Co.

Smets, Ph., Kornreich, F., Block, P., Bernard, R. & Vainsel, H. (1977)
Fuzzy diagnostic, degree of belief and utility.
In Trends in Computer-Processed Electrocardiograms, ed. van Bemmel, J.H. Willems, J.L. pp 257-259. Amsterdam, North Holland.

Snellen, H.A. (1981)
Thomas Lewis (1881-1945) and cardiology in Europe.
British Heart Journal, 46, 121-125.

Snellen, H.A. (1984)
A history of cardiology.
Rotterdam:Donker.

Sotobata, I., Richman, H. & Simonson, E. (1968)
Sex differences in the vectorcardiogram.
Circulation, 37, 438-448.

Soucek, B. (1989)
Neural and concurrent real-time systems: the sixth generation.
Chichester: Wiley & son.

Spitzer, A.R., Hassoun, M., Wang, C. & Bearden, F. (1990)
Signal decomposition and diagnostic classification of the electromyogram using a novel neural network technique.
Proceedings of the 14th Annual Symposium on Computer Application in Medical Care, 552-556.

Starr, J.W., Wagner, G.S., Behar, V.S., Walston, A. & Greenfield, J.C.Jr. (1974)
Vectorcardiographic criteria for the diagnosis of inferior myocardial infarction.
Circulation, 49, 829-836.

Starr, J.W., Wagner, G.S., Draffin, R.M., Reed, J.B., Walston, A. & Behar, V.S. (1976)
Vectorcardiographic criteria for the diagnosis of anterior myocardial infarction.
Circulation, 53, 229-234.

Startt/Selvester, R. (1986)
Comprehensive multilead normal ECG database for the next few decades: a position paper. p221. in Willems, J., van Bemmel, J.H. & Zywietz, C (eds): Computer ECG analysis: towards standardization. North Holland, Amsterdam.

Stevenson, M., Winter, R. & Widrow, B. (1990)
Sensitivity of feedfoward neural networks to weight error.
IEEE Transaction on Neural Networks, 1, 71-80.

Takatsu, F., Osugi, J. & Nagaya, T. (1986)
Is it possible to rule out extensive anterior myocardial infarction in the absence of abnormal Q waves in lead I and aVL ? Effect of infero-apical extension of infarction over apex.
Japanese Circulation Journal, 50, 601-606.

Taylor, T.P. (1973)
Computer analysis of cardiac arrhythmias.
PhD Theiss, University of Glasgow.

Touretzky, D.S. & Pomerleau, D.A. (1989)
What's hidden in the hidden layers?
Byte, 227-233.

Turing, A.M. (1950)
Computing machinery and intelligence.
Mind, 59, 433-460.

van Bemmel, J.H., Duisterhout, J.S., van Herpen, G., Bierwolf, L.G., Hengeveld, S.J. & Versteeg, B. (1971)
Statistical processing methods for recognition and classification of vectorcardiograms.
In Vectorcardiography 2, ed. Hoffman, I. Hamby, R.I. Glassman, E. Proceedings of 11th International Vectorcardiography Symposium, pp207-215, Amsterdam, North Holland.

Vitolo, E., Madoi, S., Colli, A.M., Roveda, M., La Rovere, M.T., Obbiassi, M. & Labianca, R. (1982)
The meaning of bites on the vectorcardiogram: study in adriamycin cardiomyopathy.
Journal of Electrocardiology, 15, 265.

Waller, A.D. (1887)
A demonstration on man of electromotive changes accompanying the heart's beat.
Journal of Physiology, 8, 229-234.

Walston, A., Harley, A. & Pipberger, H.V. (1974)
Computer analysis of the orthogonal electrocardiogram and vectorcardiogram in mitral stenosis.
Circulation, 50, 472-477.

Warner, R., Hill N.E., Sheehe, P.R., Mookherjee, S., Fruehan, T. & Smulyan, H. (1982)
Improved electrocardiographic criteria for the diagnosis of inferior myocardial infarction.
Circulation, 66, 422-428.

Watts, M.P. & Shoat, D.B. (1987)
Trends in electrocardiographic design.
Journal of the Institution of Electronic and Radio Engineers, 57, 140-150.

Wellens, H.J.J. (1986)
The electrocardiography 80 years after Einthoven.
Journal of American College of Cardiology , 7, 484-491.

Widrow, G. & Hoff, M.E. (1960)
Adaptive Switching Circuits.
Institute of Radio Engineers, Western Electronic Show and Convention Record Part 4, 96-104.

Willems, J.L. (1977)
Introduction to multivariate and conventional computer ECG analysis: pro's and contra's.
In Trends in Computer-Processed Electrocardiograms, ed. van Bemmel, J.H. Willems, J.L. pp213-220. Amsterdam, North Holland.

Willems, J.L., Lesaffre, E., Pardaens, J. & de Schreye, D. (1986)
Multigroup logistic classification of the standard 12- and 3-lead ECG.
In Computer ECG analysis: Toward Standardisation, ed. Willems, J.L., van Bemmel, & J.H. Zywietz, C. pp203-210. Amsterdam, North Holland.

Willems, J.L., Arnaud, P., van Bemmel, J.H., Degani, R., Macfarlane, P.W. & Zywietz, C. (1990)
Common standards for quantitative electrocardiography: goals and main results.
Methods of Information in Medicine, 29, 263-271.

Willems, J.L., Abreu-Lima, P., Arnaud, P., van Bemmel, J.H., Brohet, C., Degani, R., Denis, B., Gehring, J., Graham, I., van Herpen, G., Machado, H., Macfarlane P.W., Michaelis, J., Moulopoulos, S.D., Rubel, P. & Zywietz, C (1991)
The diagnostic performance of computer programs for the interpretation of electrocardiograms.
New England Journal of Medicine, 352, 1767-1773.

Williams, H.B. (1914)
On the cause of the phase difference frequently observed between homonymous peaks of the electrocardiogram.
American Journal of Physiology, 35, 292-300.

Wilson, F.N., Johnston, F.D., MacLeod, A.G. & Barker, P.S. (1934)
Electrocardiograms that represent the potential variations of a single electrode.
American Heart Journal, 9, 447-458.

Wilson, F.N. & Johnston, F.D. (1938)
The vectorcardiogram.
American Heart Journal, 16, 14-28.

Witham, A.C. & Lahman, J.E. (1970)
Vectorcardiogram past forty.
American Heart Journal, 79, 149-159.

Wolf, H.K., Rautaharju, P.M., Unite, V.C. & Stewart, J. (1976)
Evaluation of synthesized standard 12 leads and Frank vector leads.
Advances in Cardiology ,16 , 87-97.

Yang, T.F., Chen, C.Y., Chiang, B.N. & Macfarlane, P.W. (1993a)
Normal limits of derived vectorcardiogram in Chinese.
Journal of Electrocardiology, 26, 97-106.

Yang, T.F., Devine, B. & Macfarlane, P.W. (1993b)
Artificial neural networks for the diagnosis of atrial fibrillation.
p325-328 in ed. Macfarlane, P.W. & de Padua, F.: Electrocardiology '92
London, World Scientific Publishing Co.

Yang, T.F., Devine, B. & Macfarlane, P.W. (1993c)
Deterministic logic versus software based artificial neural networks in the diagnosis of atrial fibrillation.
Journal of Electrocardiology, 26,supplement, 90-94

Yang, T.F., Devine, B. & Macfarlane, P.W. (1994a)
Artificial neural networks for the diagnosis of atrial fibrillation.
Journal of Biomedical Engineering and computing, 1994 (in press)

Yang, T.F. & Macfarlane, P.W.(1994b)
Comparison of the Derived Vectorcardiogram in Apparently Healthy Caucasians and Chinese.
Chest, 1994 (in press)

Yang, T.F. & Macfarlane, P.W. (1994c)
Normal limits of the Derived Vectorcardiogram in Caucasians.
Clinical Physiology, 1994 (in press)

Yang, T.F., Devine, B. & Macfarlane, P.W. (1994d)
Use of Artificial Neural Networks Within Deterministic Logic for the Computer ECG Diagnosis of Inferior Myocardial Infarction.
Journal of Electrocardiology 1994, 27, supplement (in press)

Yasui, S., Whipple, G.H. & Stark, L. (1964)
Comparison of human and computer electrocardiographic wave form classification and identification.
American Heart Journal, 68, 236-242.

Yoon, Y.O., Brobst, R.W., Berstresser, P.R. & Peterson, L.L. (1989)
A desktop neural network for dermatology diagnois.
Journal of Neural Network & Computation, Summer, 43-52.

Zema, M.J. & Kligfield, P. (1979)
Electrocardiographic poor R wave progression. I: Correlation with the Frank vectorcardiogram.
Journal of Electrocardiology, 12, 3-10.

Zema, M.J. & Kligfield, P. (1979)
Electrocardiographic poor R wave progression. II: Correlation with angiography.
Journal of Electrocardiology, 12, 11-15.

Zema, M.J., Luminais, S.K., Chiaramida, S., Goldman, M. & Kligfield, P. (1980)
Electrocardiographic poor R wave progression. III. The normal variant.
Journal of Electrocardiology, 13, 135-142.

Zema, M.J., Collins, M., Alonso, D.R. & Kligfield, P. (1981)
Electrocardiographic poor R-wave progression. Correlation with postmortem findings.
Chest, 79, 195-200.

Zema, M.J. & Kligfield, P. (1982)
ECG poor R-wave progression. Review and synthesis.
Archive of Internal Medicine, 142, 1145-1148.

Zema, M.J. (1990)
Electrocardiographic tall R waves in the right precordial leads. Comparison of recently proposed ECG and VCG criteria for distinguishing posterolateral myocardial infarction from prominent anterior forces in normal subjects.
Journal of Electrocardiology, 23, 147-156.

Zoneraich, S. & Zoneraich, O. (1971)
Atrial tachycardia with high degree of intra-atrial block.
Journal of Electrocardiology , 4 , 369.

Zoneraich, O. & Zoneraich, S. (1977)
A vectorcardiographic study of spatial P, QRS and T loops in diabetic patients without clinically apparent heart disease.
Journal of Electrocardiology , 10, 207.

Zywietz, C., Borovsky, D., Faltinat, D., Klusmeir, S. & Schiemann, W. (1977)
The Hannover EKG system HES.
In: Van Bemmel, J.H. & Willems, J.L. (eds) Trends in Computer-Processed Electrocardiograms. Amsterdam, North Holland, pp159-166.

Zywietz, C. & Willems, J.L. (1993)
European conformance testing services for computerised electrocardiography.
New procedures and standards.
Journal of Electrocardiology, 26 (supplement), 137-145.