Intra-individual Variability of the Electrocardiogram

Assessment and exploitation in computerized ECG analysis

Bob J.A. Schijvenaars

Acknowledgements

Financial support by Cardio Control B.V. for the publication of this thesis is gratefully acknowledged.

Schijvenaars, R.J.A.
Intra-individual Variability of the Electrocardiogram.
Assessment and exploitation in computerized ECG analysis
Thesis Erasmus University Rotterdam — with summary in Dutch.

ISBN: 90-9013853-6

This thesis was typeset by the author with LaTeX 2_ε.
Cover design: JP Witteman

Intra-individual Variability of the Electrocardiogram

Assessment and exploitation in computerized ECG analysis

Intra-individuele variabiliteit van het electrocardiogram

Bepaling en exploitatie in geautomatiseerde ECG analyse

Proefschrift

ter verkrijging van de graad van doctor
aan de Erasmus Universiteit Rotterdam
op gezag van de Rector Magnificus
Prof. dr. ir. J.H. van Bemmel
en volgens besluit van het College voor Promoties

de openbare verdediging zal plaatsvinden op
woensdag 20 september 2000 om 13:45 uur

door

Robertus Johannes Adrianus Schijvenaars
geboren te Breda

Promotiecommissie

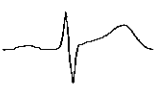
Promotor:	Prof. dr. ir. J.H. van Bemmel
Overige leden:	Prof. dr. J.D.F. Habbema Prof. dr. A. van Oosterom Prof. dr. M.L. Simoons
Co-promotor:	Dr. ir. J.A. Kors

Financial support by the Netherlands Heart Foundation for the publication of this thesis is gratefully acknowledged.

Voor Marjan en Sofie
Voor Papa en Mama

Contents

1 General Introduction



1.1 Computerized Interpretation of the Electrocardiogram

Computer interpretation of the electrocardiogram (ECG) was one of the first applications of computers in health care. The first systems were developed in the early sixties by Pipberger [1] and Caceres [2]. In the last decades, computerized ECG analysis has become one of the most widespread computer applications for decision support in health care. For example, over 50 million ECGs were analyzed by computer in the United States already in 1988 [3].

Large research efforts have been made to bring the performance of these systems to a level acceptable for routine use. In an international cooperative study, called 'Common Standards for Quantitative Electrocardiography' (CSE), most present-day ECG computer programs were evaluated [4]. This evaluation was done by comparing measurements and interpretations computed by these systems with those of a panel of experienced cardiologists on the one hand, and with a diagnosis based on clinical evidence, not involving the ECG, on the other. It was concluded that ECG analysis programs can assist clinicians in achieving more uniform and consistent interpretations of ECGs. However, continued testing and refinement was deemed necessary to enhance the performance of these systems.

Despite their good diagnostic performance, ECG analysis programs suffer from a number of drawbacks that limit their practical utility [5,6].

One of the most important shortcomings is the vulnerability for individual ECG variability [5,7]. Identical ECG signals will result in identical measurements and interpretations, but small (and diagnostic inconsequential) differences between signals may result in an entirely different diagnostic interpretation [7–9]. This variability can already be a problem when, e.g., multiple ECGs of the same patient are interpreted that have been recorded only minutes apart. Instable interpretations will not only considerably diminish the practical use of ECG computer programs, users will also lose confidence in their performance. Furthermore, variation in interpretations of similar ECGs suggests that the accuracy of programs can still be improved.

1.2 Aims and Scope of this Study

The sensitivity of ECG computer programs to intra-individual variability in the ECG is the subject of this study. The aim of the study is to determine the sources of variability in the ECG and their effect on the interpretation produced

by an ECG computer program in particular our program MEANS (Modular ECG Analysis System) [10], and to find methods to diminish this interpretation variability. The questions addressed in this study are presented in the following sections.

Question 1:

What are the most prominent sources of intra-individual ECG variability?

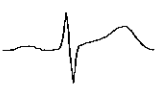
A literature study was undertaken the results of which are presented in Chapter 2. Since no comparative studies had been undertaken, no conclusive evidence about the relative importance of sources can be given. Among the most important sources of intra-individual ECG variability are chest electrode position variability and respiration. The former causes variation between ECGs of the same individual, while the latter causes variability within a particular ECG. Noise may also cause intra-individual variability, but has the advantage that it is generally clearly visible on the ECG, and that its effects on the ECG are well known [7]. We decided to investigate the effect of chest electrode position changes on the ECG in more detail since previous investigations on this subject had their limitations: the number of study participants was small, only the effect on measurements, not diagnostic interpretation, was assessed, or sources of variation other than electrode position changes were not excluded. We therefore posed the following research question:

Question 2:

What is the effect of chest electrode position changes on ECG measurements and interpretations?

In Chapters 3 and 4 we present the results of our investigations to assess the effect of chest electrode position changes on ECG analysis results. To exclude other forms of intra- individual variability when assessing the effect of chest electrode position changes, body surface potential maps (BSPMs) were used. These maps, consisting of 117 chest electrodes distributed over the entire thorax, needed to be interpolated to enable the extraction of signals from thorax locations that had not been sampled. In Chapter 3, several interpolation methods are compared. Chapter 4 presents the actual assessment of the effects of various electrode position changes on the ECG.

One of the results of the research presented in Chapter 4 is that the magnitude of the effect of chest electrode position changes is ECG dependent. Some ECGs show large variability in diagnostic interpretation when electrodes are shifted, some show no variability at all. In daily practice BSPMs are not recorded, which



makes it impossible to assess the effect on a particular 12-lead ECG in the same way. Furthermore, the chest electrode positions are generally unknown upon interpretation. Therefore, we investigated whether it was possible to use the results of the research in Chapter 4 when interpreting a standard 12-lead ECG.

Question 3:

Can intra-individual variability be assessed from the standard 12-lead ECG alone?

It is well-known that signals from electrode positions other than the ones used during the ECG recording can be approximated using a linear combination of the given ECG signals [11,12]. In Chapter 5 we followed a similar approach to simulate chest electrode position shifts. Once such shifts can be simulated, a standard 12-lead ECG can be used to generate a set of simulated ECGs from different chest electrode positions. It was assumed that processing these ECGs provides insight in the stability of the measurements and interpretation of the original ECG. In the same way, ECG variability caused by respiration can be determined by processing all individual representative beats of an ECG instead of one averaged beat. This also results in multiple interpretations and thus provides information about instability.

Question 4:

Can intra-individual variability be exploited to improve computerized ECG interpretation?

In Chapter 6 the issue is addressed whether there is a relationship between this stability information and the performance of the original interpretation. We expected that the interpretation performance expressed in terms of sensitivity and specificity on ECGs with low variability is higher than on ECGs with high variability. Furthermore, we intended to realize a practical application in our ECG program that makes use of this variability-performance relationship, to improve the efficiency of checking computer-generated ECG interpretations by hand.

In addition to providing information about the accuracy of the interpretation of a particular ECG, this stability information might also be useful in improving that interpretation. Chapter 7 addresses this issue. A combination of the interpretations of either the simulated ECGs or all individual ECG beats is expected to be more accurate than the original interpretation.

A discussion of these investigations and the conclusions are presented in Chapter 8.

1.3 References

[1] Pipberger H, Arms R, Stallmann F. Automatic screening of normal and abnormal electrocardiograms by means of a digital electronic computer. Proc Soc Exp Biol Med 1961;106:130–2.

[2] Caceres CA, Steinberg CA, Abraham S, Carbery WJ, McBride JM, Tolles WE, et al. Computer extraction of electrocardiographic parameters. Circulation 1962;25:356–62.

[3] Drazen E, Mann N, Borun R, Laks M, Bersen A. Survey of computer-assisted electrocardiography in the United States. J Electrocardiol 1988;21 Suppl:98–104.

[4] Willems JL, Abreu-Lima C, Arnaud P, Van Bemmel JH, Brohet C, Degani R, et al. The diagnostic performance of computer programs for the interpretation of electrocardiograms. N Engl J Med 1991;325:1767–73.

[5] Hurst JW, Treasure CB, Sathavorn CS. Computer errors in electrocardiography. Clin Cardiol 1996;19(7):580–6.

[6] RuDusky BM. Errors of computer electrocardiography. Angiology 1997; 48(12):1045–50.

[7] Zywietz C, Willems JL, Arnaud P, Van Bemmel JH, Degani R, Macfarlane PW. Stability of computer ECG amplitude measurements in the presence of noise. Comput Biomed Res 1990;23:10–31.

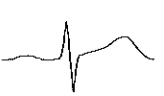
[8] Bailey JJ, Horton M, Itscoitz SB. A method for evaluating computer programs for electrocardiographic interpretation. III. Reproducibility testing and the sources of program errors. Circulation 1974;50:88–93.

[9] Willems JL, Zywietz C, Arnaud P, Van Bemmel JH, Degani R, Macfarlane PW. Influence of noise on wave boundary recognition by ECG measurement programs. Recommendations for preprocessing. Comput Biomed Res 1987;20(6):543–62.

[10] Van Bemmel JH, Kors JA, Van Herpen G. Methodology of the modular ECG analysis system MEANS. Methods Inf Med 1990;29(4):346–53.

[11] Dower GE, Machado HB, Osborne JA. On deriving the electrocardiogram from vectorcardiographic leads. Clin Cardiol 1980;3:87–95.

[12] Kors JA, Van Herpen G, Sittig AC, Van Bemmel JH. Reconstruction of the Frank vectorcardiogram from standard electrocardiographic leads: diagnostic comparison of different methods. Eur Heart J 1990;11:1083–92.



2 Intra-individual Variability in Electrocardiograms

R.J.A. Schijvenaars, J.A. Kors

Dept. of Medical Informatics, Erasmus University, Rotterdam, The Netherlands

Submitted for publication

2.1 Introduction

Human electrocardiography started with the discovery of Augustus D. Waller in 1887 that a Mercury capillary electrometer could be used to record the heart's electrical activity on the body surface and with the development of the string galvanometer by Willem Einthoven in 1903. The clinical and scientific achievements of electrocardiography since then have been manifold [1,2]. With the advent of computer programs to classify the ECG, its position as a diagnostic procedure has broadened. Electrocardiography can now easily be used in large population surveys [3–5]. In addition, recent technical developments have considerably brought down the cost of ECG recording equipment, enabling health-care providers outside the cardiological clinic to record ECGs. Thus, the importance of good-quality computer programs is further enhanced since the interpretative skills of these health-care providers often do not allow them to judge ECGs on their own [6].

In order to interpret ECGs, one must have knowledge about the distribution of ECG measurements in a normal population. A number of studies in the past have provided data on normal limits, e.g., [7,8], and numerous textbooks describe how pathological conditions become manifest in the ECG. The variability expressed by normal limits is based on normal ECGs taken from different individuals and is thus called inter-individual variability. Inter-individual variability should be distinguished from intra-individual variability, defined as variability between different ECGs from the same individual, or variability within one ECG (the latter type is called also beat-to-beat variability). The ECG can be contaminated by various sources of intra-individual variability obscuring the underlying cardiac condition and limiting the accuracy of ECG interpretation. Intra-individual variability also plays an important role in serial ECG analysis, where ECGs of the same individual but taken at different times are compared. This technique aims to benefit from the fact that intra-individual variability is smaller than inter-individual variability, as was shown by Simonson et al. in their classic paper [9]. However, it loses potential when diagnostically inconsequential manifestations of intra-individual variability are reported as serial changes.

Two diagnostically insignificant sources of intra-individual variability can be considered: variability related to the technical circumstances during ECG recording (technical sources) and non-pathological biologic variability (biological sources). This review will discuss these two sources of intra-individual variability and to what extent they affect the ECG. Inter-individual ECG variability will be briefly discussed to be able to put the importance of the two sources of intra-individual variability into context. We will concentrate on the ECG during rest, excluding intra-individual ECG variability in exercise ECGs [10], am-

bulatory monitoring [11], patient monitoring, and signal-averaged ECGs [12]. Finally, possible solutions or ways to deal with intra-individual variability are discussed.

2.2 Technical sources of ECG variability

In the past, considerable variation existed in the quality and type of recording equipment and in the lead system used. This variation resulted in ECGs that were incomparable when not recorded using the same type or even brand of equipment or the same lead system. Nowadays, this type of variation still exists although it is smaller for some sources. In the following, the technical sources that still can play a role in variability of current ECG recordings are discussed: electrode placement and reversal, electrode size and type, and skin preparation.

2.2.1 Electrode placement

Correct placement of electrodes is a crucial step for accurate ECG recording. The observation that chest electrode position variations affect ECG measurements or ECG classification was made already in 1958 by August et al. [13], who presented a few normal ECGs that showed signs of acute anterior myocardial infarction after slight electrode positional changes. These type of observations are still reported [14]. Several standards for electrode positioning have been defined and are used nowadays. In coronary care units (CCUs), leads for monitoring are often located at various places according to a local standard, often resulting in suboptimal lead configurations for event detection [15,16]. For the resting ECG, the 12-lead ECG system is most widely used. It uses three electrodes plus a reference electrode on the limbs and six electrodes on the chest [17].

The precordial electrodes require careful positioning by palpating the bony structures of the chest. Lack of observance of anatomical landmarks may lead to positioning errors. Several studies determined the occurrence rate and severity of these errors [15,18,19]. Common errors in placement of a chest electrode include: 1) placing an electrode directly on a rib, rather than on an intercostal space, 2) placing it on the wrong intercostal space, 3) placing V_1 or V_2 directly on the sternum, 4) placing V_1 to the left of the sternum. In studies about lead placement by nurses in critical care, accurate placement was performed by only 13% [15] to 25% [19] of the nurses during placement tests, regardless of their experience. In non-test situations, this percentage even decreased to 9% [19]. Wenger et al. [18] found that the average distance from the electrodes to the

prescribed locations was 2.9 cm. The right precordial electrodes V_1 and V_2 were often too high and too far apart: in 50% of the cases the upward deviation was larger than 1.6 cm. The left precordial electrodes V_4 to V_6 were commonly placed both too low and too much towards the back (in 30–50% of the cases the shift in this inferoposterior direction was larger than 1.6 cm). To improve accuracy devices have been constructed to facilitate and guide precordial lead placement [20–22] and intracutaneous dye injections have been used for serial ECG recording [9,23]. Willems et al. [23] found a reduction of 25% in day-to-day variability when marked electrode positions were used. Adequate instruction and proper supervision of the technicians and nurses are the best cure for these inadequacies. Rautaharju et al. [24] remarked that technicians are commonly trained to place the chest electrodes under the breast of women. In his study involving 6,814 women he found the effect of breast tissue on the ECG to be smaller than that of misplacement, and recommended placing the electrodes on the breast.

Chest electrode mispositioning

The question is, however, in how far chest electrode mispositioning results in changes in ECG measurements and diagnostic interpretation. The effects are large: two studies on a small number of subjects (N=13 and N=23) indicated that the majority of precordial ECG measurements show considerable changes after electrode displacement [20,25]. The effects on diagnostic interpretation seem smaller: in one study clinically significant changes in cardiologists' interpretations were found for two out of 13 subjects and in computerized interpretations for five cases [25]. In another study only clinically inconsequential changes were found [20].

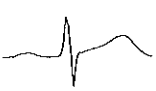
For 15 patients Herman et al. [22] compared ECGs recorded using his newly developed electrode-positioning device with ECGs obtained after deliberate electrode misplacement by 2 cm upward and downward. Considerable measurement variations were observed in all cases, while the interpretation changed in 9 and 10 cases for computer and human electrocardiographer, respectively. In another group of 80 patients, where device-aided placement was compared to routine placement, 60% (48) of the ECGs had a considerable change in at least one of 25 measurements. For instance, R amplitude changes exceeded 25% of the R amplitude in 15 cases, and ST amplitude differences larger than 100 μV were found in 10 cases. These changes are of the same order of magnitude as, e.g., the standard deviation in normal limits for male Caucasians aged 40–49 [7]: 516 μV for R amplitudes in V_3 (which is 48% of the mean R amplitude) and 50 μV for ST amplitude at the J-point. The measurement changes had clinical significance in 16% of the subjects when analyzed by an experienced reader,

and in 10% when interpreted by a computer program.

Variant limb electrode placement

The limb is in effect no more than an extension of the electrode cable, and so mistakes cannot be made in positioning these electrodes (lead reversal is discussed lateron). A number of researchers proposed to place the limb electrodes on the torso to reduce artifacts caused by patient movement and to shorten recording time in emergencies. In 1966, Mason et al. [26] proposed to place the arm electrodes in the infraclavicular fossae, 2 cm below the border of the clavicle and medial to the border of the deltoid muscle. The left leg electrode was to be placed on the anterior axillary line, midway between the rib margin and the iliac spine. This modified lead system (known as the Mason-Likar system) became widely used for exercise stress testing since it diminished artifacts due to limb movement. It has also become common practice to record a baseline (resting) ECG prior to the exercise testing using the same lead positions. Mason et al. reported differences in mean amplitude expressed as percentage of the standard ECG of 123% in the P in lead II, from 79% to 140% in RS ratio in leads I, II, III, V_3-V_6, and from 91% to 125% in the T in the same leads (N=54). They, and lateron also Diamond et al. [27], concluded that ECGs recorded with this system can be compared with resting ECGs recorded from normal electrode locations on the limb. Others, however, reported profound amplitude and waveform changes [28–31]. Therefore, ECGs recorded with the Mason-Likar lead system and ECGs recorded at the standard limb electrode positions appear not 'essentially identical' as was originally claimed by Mason et al.

An alternative lead system that Edenbrandt et al. [32] suggested entails electrodes on the lateral side of the arms at the level of the axillary fold and on the left iliac crest in the anterior axillary line. They reported a mean difference with the standard lead positions of approximately 100 μV in R wave amplitudes in leads II and aVF in 10 patients. The differences in R amplitude in other limb leads were smaller, while in all chest leads the mean difference was at most 40 μV. Lastly, Takuma et al. [33] recorded ECGs in 30 patients with arm electrodes on the anterior acromial region and leg electrodes on the anterior superior iliac spine, and found that only 0.4% of the relative differences in amplitude in all waves were larger than 5%. In an editorial, Fesmire [34] argued that the Takuma lead system allows ECGs to be recorded more quickly while minimizing motion artifacts and maximizing patient comfort.



Lead reversal

A different type of electrode misplacement is the reversal of electrodes. Some reversals are hard to detect and mimic a pathological condition [35,36]. A number of studies have investigated this problem [37–39]. Since there are 4 extremity electrodes the possible number of interchanges is $4! - 1 = 23$. For the chest electrodes, this number is 719, while for all electrodes the number of different interchanges is more than 3.5 million. Obviously, only the most common reversals have been considered. Interchanges between extremity and precordial electrodes are rare, since the precordial electrodes are often physically interconnected in modern ECG equipment. A commonly described and easily detected error is left/right arm electrode reversal [38]. Reversals involving the leg electrodes are more difficult to detect. Haisty et al. [39] reported that none of 25 cardiologists recognized a right arm/right leg (RA/RL) electrode switch in a normal ECG and presented a detection algorithm for it. Several others have also developed electrode reversal detection algorithms [40–45]. Reversal of precordial electrodes result in a break in the usual R-wave progression, for some of which also detection algorithms have been designed [42]. The best algorithms have a sensitivity for detection of electrode reversals of about 75% with a specificity of almost 100%.

Overall, unintentional mispositioning of precordial electrodes is a frequently occurring problem and may have considerable effects on the ECG. However, the studies that have tackled the problem systematically are few in number, involved often very small numbers of patients, assessed effects only qualitatively or targeted just one specific type of malpositioning. In addition, the effect on the ECG interpretation as opposed to ECG measurement has not been studied. Variant limb electrode placements were studied thoroughly to circumvent disturbance with the standard limb leads during exercise tests. It has been sufficiently documented that the de facto standard for exercise tests, the Mason-Likar system, causes large differences when compared to the standard limb electrode positions. It appears that carefully chosen alternative limb electrode positions cause only slight differences, although only small patient groups have been used to assess this. Lead reversals can mimic pathology but unlike mispositioning, they can be corrected when detected. Algorithms for detection of reversals have been developed, and show good performance, depending on which leads are reversed.

2.2.2 Recording procedure and equipment

Other sources of technical variability include size and type of electrodes used for the recording and pre-recording preparations such as skin preparation and

the application of electrode gel. Further technical sources of variability are the bandwidth and sampling frequency of recording devices and the amount of noise present on the ECG.

Electrode size

Ideally, an ECG lead records the signal between two specific points on the body surface. In practice, however, electrodes cover a finite area of the body, resulting in a potential averaged over that area. Berson et al. [46] studied the effect of two different electrode sizes (1 cm^2 and 7.5 cm^2) on the ECG in 20 subjects. They found differences in peaks of R or S amplitudes, especially in V_4 where differences ranged from -350 μV to $+220$ μV, with 8 subjects having a difference larger than 100 μV. Other precordial electrodes showed less variation. The larger electrode size resulted in systematically smaller amplitudes in all leads, although the standard deviations of the differences were about a factor 4 larger than the mean differences. Berson et al. argued that only a relatively small area close to the heart has voltage gradients high enough to be important for electrode size. When conducting electrode paste is used for ECG recording, the actual area over which potentials are averaged to be recorded by a specific electrode is not only determined by the size of the electrode itself, but also by the size of the area on which the paste is applied. If, for example, the paste is applied to the entire precordial area, the differences of the signals recorded by the chest electrodes are less pronounced (so-called gel short-circuiting) [8]. A disadvantage of smaller electrodes is that they result in higher electrode impedances, which might pose a problem for equipment with low input impedance. As remarked by Zywietz [47], care should be taken not to change parts of an ECG recording unit since their specifications may be carefully chosen to match those of other parts.

Electrode type

The type of electrode is also a possible cause of variability. Zywietz gives a review of possible electrode types [47]. Electrodes used widely nowadays are self-adhesive silver-silverchloride electrodes. Suction electrodes are another frequently used type (e.g., [48]). The electrical characteristics (impedance and offset potential) of various types of electrodes differ [49,50] but as long as they meet AHA standards [51], the effect on the ECG will be minimal.

Skin preparation and electrode gel

The ratio of skin-electrode contact impedance and recorder input impedance plays an important role in the quality of the recording. Zywietz [47] covers this aspect as well in his review. In case of unprepared skin, the skin-electrode impedance may vary between 50 to 200 kΩ at 60 Hz [52], decreasing with time [53,54]. A high input impedance of the recording device can minimize the effect of these impedance variations on the ECG [55]. The 1975 recommendations of the AHA [56] stated 5 MΩ for frequencies up to at least 60 Hz; modern ECG recorders have an input impedance in the order of 10 MΩ.

Skin-electrode impedance may be reduced by skin abrasion [57,58] or use of conductive electrode gel. Abrasion also reduces change in the skin-electrode potential when pressing the electrode [59] or moving it relative to the skin [60]. Since electrode impedance decreases with time, skin abrasion is not very useful for long-term recordings. It was therefore recommended to use at most light abrasion for long-term recordings since it may cause skin irritation and infections [61,62].

The effect of a conducting paste or gel between skin and electrode also reduces disturbance. Some considered electrode gel unnecessary for recording ECGs and achieved comparable quality by skin rubbing or using water or tubed materials such as handcream, toothpaste, mayonnaise, mustard, etc. [63,64]. Although these were no recent studies, probably any water-soluble lubricant containing a small quantity of free electrolytes and a consistency, adequate to maintain a thin aqueous film between skin and electrode, would suffice as electrode contact material. The controversy about whether or not to use water remains [65–67], partly because of the risk of gel short-circuiting.

Bandwidth and sampling frequency

A too low frequency response of the recording system can affect the ECG in the sense that amplitudes are reduced, as was already observed by Einthoven [68]. Wave durations are less affected. Using a 100 Hz bandwidth for adult rest ECGs, Berson et al. [69] showed that amplitude errors greater than 50 μV can occur in over 10% of the recordings. Therefore, the frequency response for adult rest ECGs recommended by the AHA in 1991 was at least 125 Hz [70].

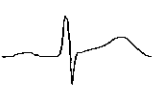
As a consequence of this minimum bandwidth, a sampling frequency of at least 250 Hz is required. Barr and Spach [71] performed tests to determine the minimum sampling frequency for ECG recording and found 500 Hz for adult rest ECGs to result in a mean reconstruction error of less than 1%. This value has also been adopted as a standard [70]. For pediatric ECGs, however, higher

sampling rates are necessary [71,72]. Computerized analysis of the ECG is also facilitated when a higher than the theoretically minimal sampling frequency of 250 Hz is used [73]. Conversely, a too low sampling frequency can have large effects on computer-generated measurements and interpretations [74]. Long-term storage, however, is often done after decreasing the sampling frequency to 250 Hz or by 'lossy' compression of the recorded signal, whereby information is lost (see [75] for a review on ECG compression), hampering the use of these ECGs in serial comparison.

Noise

Different noise sources may influence the recording: myoelectric activity, electromagnetic fields, or movement of the patient or the patient cables [47]. Electromyographic noise stems from muscular activity and is common in exercise and pediatric ECGs. The electromagnetic fields generated by AC wires cause interference of 50 or 60 Hz, or harmonics of these mains frequencies. Patient or patient cable movement often results in baseline wander [76]. This baseline wander can be gradual, consisting of frequencies not exceeding 1 Hz, or very abrupt (baseline shifts). Due to the discrete nature of diagnostic criteria used in computer programs, these disturbances can lead to very different diagnostic interpretations [14,77]. To suppress the different types of noise, dedicated filters have been developed [78–80]. Averaging individual beats is also used to decrease noise influence. However, noise can still affect the interpretation of the ECG as has been shown in the project 'Common Standards for Quantitative Electrocardiography' (CSE) [81,82]. The standard deviation of R amplitude differences after adding 35 μV RMS high frequency noise was 200 μV [82], indicating nonlinear effects (e.g., interchange of waveform labelling).

In conclusion, recent studies into ECG variability due to recording equipment are few in number and most are qualitative. From these studies it follows that the quality of present-day recording equipment reduces the occurrence of ECG variations due to skin-electrode impedance variations. Additionally, use of electrodes of identical composition as well as application of lubricant ensure minimal variability due to recording equipment. The specifications for procedures, devices and materials used for ECG recording nowadays are such that recording disturbances are well controlled. However, intra-individual variability may result when ECG quality varies. Noise can cause variation in interpretation due to the discrete nature of computerized ECG interpretation. For a reliable and repeatable computerized ECG interpretation very high quality recordings are required. To ensure accurate recording, the American Heart Association regularly publishes recommendations for instrument standardization [17,56,70,83,84].



2.3 Biological sources of ECG variability

Sources of ECG variability of biological origin are age, weight, heart position, respiration, physiological condition, pregnancy, meals, posture, and the 'Brody effect'. Environmental sources that have an effect on the human cardiovascular system such as emotional stress, temperature, and altitude are also discussed. Some biological sources of variability such as race [40,85–87], gender [88–91] and chest configuration [92], only affect inter-individual variability and will not be discussed here.

2.3.1 Constitutional sources

As already pointed out by Simonson [8], the constitutional variables age, weight, and heart position cannot be studied separately. In most humans, aging is accompanied by weight change that in turn influences heart position. This complicates a discussion on their effects on the ECG. The discussion below will attempt to differentiate between the various sources. In his classic book [8], Simonson discusses most of the items dealt with here. The discussion in this paper will, therefore, mainly focus on studies after the publication of Simonson's book in 1961.

Age

For a long time, age has been recognized as an important cause of intra-individual changes in the ECG. The investigation of intra-individual age trends is difficult, however, because of the long time needed for follow up and because factors such as body weight also change with age and contribute to the variability. Another difficulty is the increase in prevalence of pathology with increasing age, which makes the assessment of non-clinical age-related variability more difficult [93].

In the age group of 0–18 years, ECG changes are profound, and normal limits are dependent on the child's age [94,95]. In older age groups, the ECG changes are smaller and have a larger time scale. A review paper by Simonson [96] discusses the general age trends found among adults. Some of those are a general decrease of precordial amplitudes (e.g., the QRS spatial magnitude decreases with approximately 8% per decade), a leftward shift of the frontal plane axis (approximately 10° per decade), and a more anterior axis in the horizontal plane. Interval durations increase for PR and QT interval. According to Simonson, most trends flatten out after the age of 50. Others, however, focussing on the effects in elderly subjects aged 70 or older, still found a decrease

in QRS amplitudes with age. Although the above trends have been determined mainly in cross-sectional studies, they were confirmed in the few longitudinal studies addressing this aspect [40,97,98]. In these studies, a decrease in QRS duration was also found. Other factors like race, sex, and obesity influence the age effect [86,96,99]: for example, the R amplitude in V_5 decreases with age for normal-weight men, but increases for normal-weight women. Furthermore, the decrease of the mean R amplitude in V_5 in Hispanic men is about 40% of the decrease found in white and black men. These studies also suggested that in general obesity tends to reinforce the trends caused by age.

Comparing these effects with inter-individual variability (in the form of the mean and standard deviation of the distribution of measurements for caucasian male normals aged 40–49) shows that the intra-individual changes in QRS spatial amplitude become comparable to the standard deviation after three decades: 25% of the mean spatial amplitude [7]. This does not hold for all measurements: PR duration, for instance, increases with about 3 ms per decade, but the standard deviation of the normal distribution is 22 ms [7].

With respect to cardiac rhythm, there is a marked increase in the frequency of occurrence of supraventricular premature complexes (SVPC) and ventricular premature complexes (VPC), ranging from less than 1% in the age group 20–29 years to 5% (SVPC) and 10% (VPC) in the age group over 70 years [100].

Weight

Intra-individual variability due to changes in weight is mostly long-term variability that is hard to study. Most researchers determined correlations between ECG measurements and parameters describing body build. A number of studies showed that obesity accompanies a superior and anterior deviation of the QRS axis and a decrease in precordial voltages [8,86,101–103]. In a study on criteria for left ventricular hypertrophy in 550 subjects, Selzer et al. [104] already remarked that the majority of the false positives with high precordial amplitudes were emaciated. Others found a correlation between precordial amplitude measurements and weight or a weight-related index (e.g., ponderal index) [105,106]. However, in a study among 257 subjects, Dougherty et al. [92] corrected the frontal QRS axis for differences in heart position and observed no significant variations in the frontal axis attributable to weight. In a subsequent study (N=458), a correction for both frontal QRS axis and heart-electrode distance resulted in a reduction of variability in R amplitude in V_6 of 56% [107]. A smaller distance between heart and precordial electrode, e.g., induced by a methoxamine infusion or a left lateral patient position [108], causes larger R-wave amplitudes in the left precordial leads. Age alters weight-related electrocardiographic trends as pointed out by Simonson [8]. For example, the QRS axis

in obese subjects aged 20–29 years is more inferior than in non-obese subjects aged 50–59 years, although the general trend in obesity is a superior shift.

A study on intra-individual variability due to weight was performed by Brohet and Tuna [109], who studied the effects on the vectorcardiogram of weight reduction by surgical operation (jejuno-ileal bypass) in 37 chronic markedly obese patients. They observed in VCGs, recorded one year after surgery, a statistically significant decrease in Q and S amplitudes as well as in maximum P and QRS vector amplitudes. They concluded, however, that the magnitude of the effect was too small to be clinically significant: -70 μV for the maximum spatial QRS vector, for which the standard deviation of the normal distribution is 440 μV [7].

The obvious consequence of decreased precordial voltages is a decreased electrocardiographic sensitivity for LVH in obese subjects [110–113]. On the other hand, obesity has been demonstrated to be associated with the presence of LVH [114–116], possibly due to various causes [117]. Therefore, a number of suggestions have been made for weight-adjusted LVH criteria (e.g., [112]) that result in a higher sensitivity [118–120].

Heart position

The position of the heart in the thorax depends on a number of variables that individually contribute to ECG variability, such as obesity, respiration, posture, and pregnancy. A model study by Huiskamp et al. [121] showed that translational changes as small as 0.5 cm can reflect changes in QRS amplitude of the order of the observed standard deviation for large populations of normals. In a study on inter-individual variability of heart position and orientation using MRI images, Hoekema et al. [122] pointed out that variability in heart position can influence QRS amplitudes considerably. With respect to intra-individual variability in heart position, Dougherty studied changes in anatomic heart positions and consequential changes in electrical axes in individuals [123]. He showed that anatomic rotation in the frontal plane causes on average three times as large a shift in QRS axis than what is to be expected by the same rotation of a single dipole. This anatomical heart rotation is probably also the most important cause of upward directed left axis shifts in case of obesity and pregnancy. In addition, in postural changes heart position plays an important role: while standing up, the axis is more downward than in the supine position. Also switching from a supine to a left lateral position causes a smaller heart-thorax distance, which results in larger amplitudes in the left lateral leads V_5 and V_6 [108].

An illustration of the effect of the amount of tissue between heart and electrode is the study by LaMonte et al. [124], who assessed the effects of mastectomy

on the ECG in 39 subjects. They noted significantly increased amplitudes in the precordial leads in all waves. Remarkable, however, was the finding that there were no statistical differences between the changes in P or T amplitudes after right or left mastectomy. The QRS amplitude showed significant increases after right mastectomy in left precordial leads V_1, V_5 and V_6 (124 to 303 μV), while after left mastectomy QRS amplitudes in V_3R and V_1 to V_4 (66 μV up until 284 μV) increased. LaMonte and coworkers commented that these effects cannot be explained solely by reduced distance between myocardium and electrode, a factor also found by Horton et al. [125]. LaMonte et al. suggested a possible second, unknown, relationship between the precordial voltage and the altered thoracic volume conductor.

All variables discussed above have a considerable effect on the variability between ECGs. They are, however, often interdependent. The heart position becomes more horizontal when one gains weight, people generally gain some weight as years pass, etc. The general trend is a decrease in amplitudes and a left axis shift in frontal QRS axis with increasing age or weight. The amplitude decrease, however, will normally be gradual and thus is unlikely to be a major contributing factor in minute-to-minute or day-to-day variability. The left axis shift is probably mainly due to a horizontally directed change in heart position. Changes in heart position and thus the consequential ECG changes occur in other sources of intra-individual variability as well (pregnancy, posture, respiration, etc.).

2.3.2 Functional and physiologic sources

Pregnancy

The cardiovascular changes during normal pregnancy include sodium and water retention, increased blood volume and an increased resting cardiac output [126], resulting in an increase in resting heart rate [127,128]. Carruth et al. [128] studied the ECG in 102 normotensive pregnant women during pregnancy (first, second and third trimester), 1 to 3 days after delivery and 6 to 8 weeks postpartum. They found an increase in mean resting heart rate from 77 bpm in the first trimester of pregnancy to 87 in the third, while the postpartum heart rate was 66 bpm. They suggested that uterine growth causes a progressive upward pressure on the diaphragm, which in turn causes a progressive elevation and rotation of the heart. Like in obesity, this can be detected on the ECG as a leftward shift of the frontal QRS axis. A progressive leftward shift was seen in the T wave. However, in concordance with others [129,130] they stress the variability of changes in the QRS axis and in the T axis: the pattern of QRS change in an individual pregnant woman was not predictable. An early study even reported

opposite trends: a leftward QRS axis shift in early pregnancy and a rightward shift prior to delivery [131]. Changes in PR, QT and QTc durations were not clinically important, as was reported before [132].

Others studied the effect on the ECG during exercise and found ST depressions and T wave flattening on the maternal ECG [133–136]. An increased incidence of repolarization abnormalities in pregnant women could not be concluded, since the findings were also found postpartum [134]. The ECG during labor and delivery is very transient [137] and will not be discussed.

Brody effect

The short-circuiting effect of the relatively high conductive intracavitary blood-mass on the electrical signals has been presented in a theoretical analysis by Brody in 1956 [138]. In a simplified model study he showed that electrical field components generated by dipoles parallel to the heart model surface (a sphere) are reduced, while components from dipoles perpendicular to the heart surface are augmented. For example, this effect would cause V_6 to be less sensitive to signals originating from the anterior side of the heart. Since then, numerous researchers attributed observations on the ECG to this 'Brody-effect'. Others, however, showed that the magnitude of this effect is dependent on the measurement location on the body surface, and may even have an effect opposite to the one Brody presented [139,140]. Intra-individual variability may occur when the intracavitary blood volume changes, for instance due to profuse sweating after sauna bathing [141]. These changes will on the other hand also be accompanied by changes in heart volume. It has been shown that end-diastolic left ventricular volume is inversely correlated with maximum spatial QRS amplitude in patients without left ventricular hypertrophy [141,142]. In addition, a model-based comparison of the Brody effect and positional changes of the heart, when the heart volume changes, suggested that the latter has a more significant effect [143].

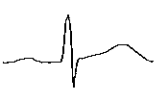
Respiration

Qualitative reports of ECG changes due to respiration have been published in the past (e.g., [13,144]) and consisted of sometimes large changes in QRS amplitudes, resulting in simulation of, for instance, an acute myocardial infarction. Systematic studies of respiratory effects indicated that deep inspiration caused more prominent changes than deep expiration. Deep inspiration produced orientation changes: the P vector shifted vertically, the QRS vector vertically and posteriorly and the T vector horizontally and anteriorly [145–147]. In addition,

amplitudes decreased significantly: a mean change of -299 μV in maximum QRS spatial amplitude between VCGs recorded during normal respiration and VCGs recorded during deep inspiration [146]. In the same study, a maximum QRS spatial amplitude decrease of larger than 10% was observed in 65% of the cases. Sutherland et al. [148] pointed out that resting tidal respiratory volume has a relatively small effect. In addition, Singh et al. [149] found that the magnitude of the shift in frontal QRS axis depends on the orientation during normal respiration: ECGs with more vertical QRS axes show larger shifts during inspiration. However, the above effects were not observed in all subjects and thus not predictable in individual cases [150]. Consequently, discrimination of pathological from non-pathological Q waves by way of recording at inspiration and expiration proved unsuccessful [151]. In a study in 194 patients [146], changes in diagnostic interpretation due to deep inspiration occurred in about 17% of the cases, as compared to 12% due to an electrode position shift of one intercostal space and 12% due to posture. It has been suggested that these changes are caused by the downward shift and clockwise rotation of the heart [152,153], by movement of the precordial electrodes relative to the heart [145], and by the decreased electrical conductivity of the lungs during inspiration [147,153]. To avoid respiration related ECG variability, it has been proposed to use voluntarily controlled respiration during ECG recording [154].

Typical rhythm changes due to respiration have also been observed. This phenomenon is called respiratory arrhythmia, and is probably caused by receptors that respond to changes in lung volume or respiratory blood pressure fluctuations (see, e.g., Hirsch et al. [155]). The phenomenon has even been used to estimate the respiration frequency [156] or the tidal lung volume [157] from the ECG. Absence of this phenomenon is considered abnormal and has been observed in alcoholic individuals [158].

Normal respiration is considered the main source of variability within one ECG, also called beat-to-beat variability. The magnitude of the latter was studied by Fischmann et al. [159], who assessed the beat-to-beat variability in 58 male subjects and reported a mean difference in wave durations of 8, 6, and 12 ms for the PR, QRS, and QT, respectively. The mean absolute differences in spatial amplitude were 30 μV for P and QRS, and 40 μV for T. These values are smaller than those of inter-individual variability. Standard deviations of the measurement distribution of white caucasian males aged 40–49 are: 21.8, 9.9 and 29.3 ms for PR, QRS and QT duration, respectively. For the maximal spatial amplitudes of P, QRS, and T the standard deviations of the normal limits are 60, 440 and 160 μV [7].



Physical training

The electrocardiographic changes due to physical training have been studied in the form of differences between normal individuals and athletes, i.e. inter-individual variability. The most prominent findings are bradycardia, longer ventricular conduction times, higher voltages in QRS complex, T wave and U wave, and ST elevation (see, e.g., [160–162]), consistent with LVH. Some of these differences are less pronounced in women [163]; one study even found cases of sinus tachycardia in female athletes [164]. Bjørnstad et al. performed a number of studies on ECG findings in athletes [162,165–169]. In addition to the differences mentioned, other typical findings were PQ prolongation associated with low heart rates and incomplete right bundle branch block. Since Bjørnstad et al. made a subdivision into top athletes and athletic students they were able to show that some measurements (heart rate, conduction times and, on a somewhat lower level of significance, indices of right and left ventricular hypertrophy) were related to the level of fitness [165,167,169]. The type of fitness also plays a role. Endurance athletes have significantly lower heart rates than, e.g., gymnasts or sprinters [167], which is probably due to aerobic training. It is likely that this inter-individual variability is transferable to intra-individual variability, since a number of ECG findings appeared reversible in athletes that were followed during detraining [161].

Meals

A meal can cause considerable ECG changes, both in normals and in cardiac patients [8,170,171]. In normals, the most prominent changes after a standard meal are: an increase of heart rate, a decrease of T-wave amplitude and QT interval, and small left axis shifts of the QRS and T axes [170]. In cardiac patients ST depressions and/or T wave inversions can be observed [171], and exercise tolerance decreases after a meal [172]. It was suggested that sympathetic stimulation might be a possible explanation for the observed changes [170].

Posture

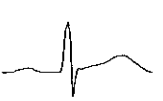
According to Simonson [8], the mean difference in amplitudes between supine and sitting position is small and statistically not significant. He concluded that the standards for chest leads in the supine position might be applicable to the sitting position. Later studies of effects on the VCG were in agreement with the conclusion that mean differences of orientation angle were not significant [30, 146,173]. However, these studies did reveal increases of QRS spatial amplitude and R amplitude in lead Z, and a decrease in Q amplitude in lead Y from supine

to a sitting or standing position. Less drastic posture changes, e.g., from supine to the 60° head-up position resulted in similar ECG changes [174].

However, posture changes also cause substantial orientation and magnitude changes in individuals without a discernable trend. Wave durations do not appear to change with posture [173]. According to Riekkinen et al. [146], changes in diagnostic interpretation due to posture occurred in 12% of the 144 cases studied, and appearance as well as disappearance of pathological patterns can occur [30].

Besides the difference between supine and sitting or standing, changes between different side-lying positions that frequently occur during continuous monitoring [175] have also been studied. If changes caused by changing side are not taken into account, they result in frequent false alarms in a CCU. An mean increase in left lateral R amplitudes of 700 μV when switching from supine to left lying position was found in a study by Feldman et al. [108]. In a study on 40 subjects [176], the mean QRS axis change when subjects moved from supine to right or left side was small: about +15°. The number of subjects with positional QRS changes in at least one lead was approximately 50% for both right as well as left lying position. Changes larger than 100 μV in the ST-segment were found in leads V_2 to V_6 in 6 subjects, in 5 of which the changes occurred while lying on the left side. Others also studied the effect of lying on different sides on the ST segment [175,177] and reported the most pronounced changes in the left lateral position in the lateral leads as well. The magnitude of the changes was small, however: Nørgaard et al. observed a largest average difference of 22 μV difference in ST segment amplitude in lead X in the left lateral position.

Heart position plays an important role in most of the sources discussed. Pregnancy and meals result in a more horizontal heart position, and thus in a leftward shift of the frontal QRS axis. Inspiration and a sitting or standing posture result in a more vertical heart position and thus a rightward axis shift. Secondary effects are, for instance, an increased lung volume during inspiration, causing lower amplitudes. The Brody effect normally causes no intra-individual variability and even in the cases that it does, the accompanying heart position changes have more effect on the ECG. Physical fitness results in changes similar to LVH and sometimes right bundle branch block. However, for all these effects the variability between individuals is large. This means that the trends observed cannot generally be applied to individuals.



2.3.3 Environmental sources

Emotional stress

It has been shown that emotional stress has a partly different effect on the ECG than physical stress [178,179]. ECGs of subjects under emotional stress show increased heart rates and minor ST changes [180]. In addition, changes in the R-wave amplitude in lead V_5 during mental arithmetic stress were studied by Doi et al. [181], who found a significant decrease. They attributed this decrease to an enhanced myocardial contractility and increased cardiac output, reducing the cardiac volume and increasing the distance between heart and anterior chest wall. Methods to induce mental stress are, however, difficult to standardize. The methods used include quizzes, mental arithmetic, stressful interviews, and the suggestion that faulty ECG equipment might cause electrocution [182]. Careful observation of the emotional response is important, however, since Huang et al. [183] studied the effect of stressful interviews on the length of the QT interval and found changes depending on the subject's reaction. Most subjects responded to the interview with anger and resentment and showed QT shortening. Two of 17 subjects, however, felt dejected and overwhelmed and showed QT lengthening.

Temperature

One study of the ECG in sauna-bathing subjects by Rautaharju et al. [184] revealed that high temperatures affect mainly the repolarization process, causing shorter ST-T segments and decreased T amplitudes. These repolarization changes were attributed to the temperature dependence of the single fiber action potential duration. In addition, the average heart rate increased considerably from 66 bpm to 109 bpm after 25 minutes of heat exposure. After a cool shower the heart rate dropped. According to the authors, this increase cannot be explained by temperature effects alone. They assume a sympathic effect as well, especially because the changes observed are similar to the ones caused by moderate exercise. The effects of low temperature, studied in subjects ingesting ice water also mainly effected the repolarization phase [185,186].

High altitude

A number of studies recorded serial ECGs at various altitudes (up to 8200 m) during expeditions or in hypobaric chambers (e.g., [187–190]). The overall effects observed are an increased P amplitude and a rightward shift of the frontal QRS axis resulting in a right axis deviation. In some cases a right bundle branch

block developed, while in others ST changes, T flattenings, or T inversions were observed. All studies attributed most of the changes at high altitude to the development of pulmonary hypertension resulting in a right ventricular strain pattern similar to acute cor pulmonale due to pulmonary embolism. This pulmonary hypertension was confirmed by measurement of pulmonary artery pressure in one of the studies [190]. Most changes disappeared again after return to normal pressure, but in some cases, signs of right ventricular hypertrophy persisted and were confirmed by echocardiography [191]. On the other hand, native people living at high altitudes show less ECG abnormalities than migrated people, even when migrated as a child [192].

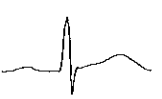
Vibrations

The effect of vibrations on the ECG seems to be temporary [193,194]. Vibrations at the resonance frequency of the heart (approximately 8 Hz) cause amplitude changes in the R wave and in the ST segment. These are attributed to ischemia, caused by inadequate oxygen supply. The changes disappear within a few minutes when the exposure to vibrations is stopped.

The environmental issues mentioned above can play a role in routine ECG recording, e.g., in anxious patients or individuals that have been staying at high altitudes for some time. Emotional stress produces mainly changes in the repolarization phase, partly similar to physical exercise. Repolarization is also affected by changes in temperature and the effect resembles changes due to exercise as well. High altitudes cause stress on the right ventricle and may induce RVH or sometimes RBBB.

2.4 Combined sources of intra-individual ECG variability

Intra-individual variability encountered in practice more likely than not is a combination of the sources discussed above. Fortunately, in most recording situations a number of these sources can be ruled out (e.g., vibrations or high altitude), while others are easily controlled (e.g., limb electrode position or skin preparation). The time interval between subsequent ECG recordings can also be used to rule out certain sources: age-related changes are not present in ECGs recorded a few weeks apart. Furthermore, intra-individual variability within one ECG (beat-to-beat variability) cannot be caused by electrode placement or weight changes. Respiration and noise are the main sources for this type of variability. Its magnitude is small, as discussed already.



Short-term intra-individual variability between ECGs recorded minutes apart or a few hours at most is called minute-to-minute variability. Sometimes electrodes remain in place, sometimes they are reapplied on marked locations. In addition to sources that cause beat-to-beat variability, other sources may contribute to the variability as well: posture, meals, or, in case of removed and reapplied chest electrodes, small position shifts may result in ECG changes. Minute-to-minute measurement changes were reported by De Bruyne et al. [195], who recorded two ECGs 30 minutes apart using marked electrode positions in 97 elderly subjects. They reported a coefficient of variation of 4.4%, and 6.5% for PR and QRS duration, respectively. The coefficient of variation for the Sokolow-Lyon voltage was 12.2%. Others reported smaller mean minute-to-minute variability of the Sokolow-Lyon voltage, with coefficients of variation of 3.1% [196] and 2.6% [197], but in these studies the electrodes remained in place between the recordings.

When the time interval between recordings spans one or more days, intra-individual variability is called day-to-day variability. Electrodes seldom remain in place but locations may be marked. Even then, however, larger changes than those encountered in minute-to-minute variability studies are found. Reported coefficients of variability for day-to-day changes of the Sokolow-Lyon voltage were 18.5% [197], 14.9% [195], and 10.0% [196]. When electrodes remained in place minute-to-minute variability is about a factor 2.5 times smaller than day-to-day variability [196,197], if electrodes are removed and reapplied, minute-to-minute variability is only 1.2 times smaller than day-to-day variability [195]. No position marking at all enlarges variability even more, as was shown by Willems et al. [23]. This demonstrates the importance of lead placement. These results confirm findings by Larkin et al. [198], who assessed minute-to-minute variability when leaving electrodes in place and when using marked electrode positions. They concluded that the former procedure reduced repeat variation by approximately 60%.

ECG recordings taken one or more years apart show largest intra-individual variability. Sources such as age, weight, or physical fitness then come into play, in addition to the sources already having effect on smaller time scales [3,97]. Some concluded that year-to-year variability is comparable to day-to-day variability [195,199]. Nowadays, normal limits are dependent on age [7], and age- and weight-dependent diagnostic criteria are used by some. Due to the different units of these two sources, their relative contribution to year-to-year variability is difficult to compare. Norman et al. corrects the Cornell Voltage-duration product using the formula $V + A \times (\text{age in years}) + B \times (\text{BMI in kg/m}^2)$, where $A \approx 3$ and $B \approx 20$ [119]. This indicates that a realistic weight change of 3 points in body mass index (about 10 kg for a person of average length) corresponds to an age change of 20 years. The overall magnitude of year-to-year variability is considerable. According to the results of Michealis

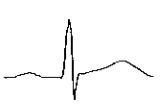
et al. [3], the 95% range of changes for adult males of age 36 to 45 for ECGs at most three years is 25 ms for QRS duration, and 700, 1550 μV for R amplitudes in aVF and V_5 respectively. Corresponding inter-individual variability values (twice the standard deviation of the normal limits) [7] are: 19.6 ms, and 906 and 1104 μV, respectively. This indicates that the magnitude of year-to-year variability is comparable to inter-individual variability.

2.5 Discussion

Intra-individual variability in the ECG can have many sources. Although most sources cause considerable effects on the ECG recorded during rest, only few sources play an important role in routine situations when adequate provisions are taken. The occurrence of sources of variability such as temperature, altitude, and pregnancy are either rare or easily detected and taken into account. Others can be controlled using a standard protocol for ECG recording, postprandial state, or posture. Finally, sources such as emotional status, physical fitness, weight, and age can also be taken into account, but are not so easily detected nor controlled. The most difficult sources to deal with are the ones that vary in a random fashion, like chest electrode positioning or that are not controlled in routine situations, like respiration.

Considering the relative magnitudes of intra-individual variability on different time scales, and the contribution to those variabilities by different sources, respiration appears to be the major contributing factor to beat-to-beat variability. For variability between recordings electrode position variation is a main source. For adults, effects caused by age and age-related factors come into play only when the time interval between recordings spans at least several years. Noise can play a large role in both types of variability, but has the advantage that it is clearly visible on the ECG.

Variability in the ECG manifests itself on three different levels: the signals, the measurements derived from those signals, and the diagnostic interpretation. Human electrocardiographers have less trouble discarding inconsequential signal or measurement changes than computer programs [14]. Although computerized ECG analysis eliminates measurement variability [200], and nowadays produces reliable measurements [81], these measurements may be influenced by small changes in the ECG signal [74,82,201]. Small changes in measurements can in turn considerably influence diagnostic interpretation. Methods to solve this problem include smoothing the discrete diagnostic criteria [202], or performing repeated measurements, which is common in other scientific areas. The latter method may be implemented in daily routine for beat-to-beat variability by measuring each dominant beat independently. Fischmann et al. [159]



suggested to use measurements averaged over all measured dominant beats for interpretation, instead of measurements of one averaged beat. The latter method smoothes out detail that may be clinically relevant. This method was taken one step further by Kors et al. [203], who reported improved interpretation accuracy when using a combination of the interpretations of individual beats compared to the interpretation of one representative beat. Others have tried different ways of correcting for beat-to-beat variability [204,205]. Some investigators tried to determine whether beat-to-beat variability contains diagnostic information. The results were mixed: ventricular arrhythmias seem to be preceded by beat-to-beat changes in complex morphology [206], but beat-to-beat QRS variability does not appear to have discriminative power for detection of coronary artery disease [207]. Another way to reduce the impact of variability is to avoid using measurements showing large variation when electrodes are shifted.

In conclusion, although most of the sources discussed cause intra-individual ECG variability smaller than inter-individual variability (the mean effect of sources on a population is often marginal), individual changes can be large. Even though the effects of these sources have been studied, the resulting knowledge is not always used in daily practice. Computer programs for ECG analysis, and to a lesser degree human interpreters, could benefit from taking this variability into account. Diagnostic criteria should be made dependent on those sources that are easily quantifiable before interpretation, e.g., age or obesity. Normal limits that are a function of these parameters already exist [7]. Other sources should be controlled as much as possible, e.g., chest electrode positions by thorough instruction to nurses. As a final measure, the interpretation process can be made less sensitive to variability by using less discrete thresholds (e.g., 'fuzzy logic') used in diagnostic criteria or performing repeated measurements by analyzing beats separately, possibly resulting in alternative diagnostic statements to consider. This last method could be taken one step further by modeling sources of intra-individual variability and simulating their effect on the ECG recorded. Interpretation of these simulated ECGs may provide insight into whether or not the diagnostic interpretation of the ECG at hand is sensitive to intra-individual variability.

2.6 References

[1] Burch GE, De Pasquale N. A history of electrocardiography. San Francisco: Normal Publishing, 1990.

[2] Fye B. A history of the origin, evolution, and impact of electrocardiography. Am J Cardiol 1994;73:937–49.

[3] Michaelis J, Lippold R, Gluck E, Schindler W, Scheidt E. Automated serial ECG-analysis within an epidemiological study. J Electrocardiol 1987;20 Suppl:34–6.

[4] Rautaharju P. Electrocardiography in epidemiology and clinical trials. In: Macfarlane P, Lawrie T, editors. Comprehensive electrocardiography. Theory and practice in health and disease, vol. 2, 1st ed. Oxford: Pergamon Press, 1989; pp. 1219–66.

[5] Hofman A, Grobbee DE, De Jong PVTM, Van den Ouweland FA. Determinants of disease and disability in the elderly: the Rotterdam elderly study. Eur J Epidemiol 1991;7:403–22.

[6] De Bruyne MC, Kors JA, Hoes AW, Kruijssen DA, Deckers JW, Grosfeld M, et al. Diagnostic interpretation of electrocardiograms in population-based research: computer program research physicians, or cardiologists? J Clin Epidemiol 1997;50(8):947–52.

[7] Macfarlane P, Lawrie T. Normal limits. In: Macfarlane P, Lawrie TDV, editors. Comprehensive Electrocardiology, vol. 3, 1st ed. New York: Pergamon Press, 1989; pp. 1442–526.

[8] Simonson E. Differentiation between normal and abnormal in electrocardiography. St Louis, Missouri: Mosby, 1961.

[9] Simonson E, Brozek J, Keys A. Variability of the electrocardiogram in normal young men. Am Heart J 1949;38:407–22.

[10] Okin PM, Kligfield P. Heart rate adjustment of ST segment depression and performance of the exercise electrocardiogram: a critical evaluation. J Am Coll Cardiol 1995;25(7):1726–35.

[11] Corday E. Historical vignette celebrating the 30th anniversary of diagnostic ambulatory electrocardiographic monitoring and data reduction systems. J Am Coll Cardiol 1991;17(1):286–92.

[12] Signal-averaged electrocardiography. J Am Coll Cardiol 1996;27(1):238–49.

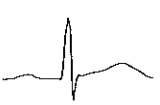
[13] August T, Mazzeleni A, Wolff L. Positional and respiratory changes in precordial lead patterns simulating acute myocardial infarction. Am Heart J 1958;55(5):706–14.

[14] Hurst JW, Treasure CB, Sathavorn CS. Computer errors in electrocardiography. Clin Cardiol 1996;19(7):580–6.

[15] Drew BJ, Ide B, Sparacino PS. Accuracy of bedside electrocardiographic monitoring: a report on current practices of critical care nurses. Heart Lung 1991; 20(6):597–607.

[16] Thomason TR, Riegel B, Carlson B, Gocka I. Monitoring electrocardiographic changes: results of a national survey. J Cardiovasc Nurs 1995;9(4):1–9.

[17] Sheffield LT, Berson A, Bragg-Remschel D, Gillette PC, Hermes RE, Hinkle L, et al. Recommendations for standards of instrumentation and practive in the use of ambulatory electrocardiography. The Task Force of the Committee on Electrocardiography and Cardiac Electrophysiology of the Council on Clinical Cardiology. Circulation 1985;71(3):626A–36A.



[18] Wenger W, Kligfield P. Variability of precordial electrode placement during routine electrocardiography. J Electrocardiol 1996;29(3):179–84.

[19] Bupp JE, Dinger M, Lawrence C, Wingate S. Placement of cardiac electrodes: written, simulated, and actual accuracy. Am J Crit Care 1997;6(6):457–62.

[20] Kerwin AJ, McLean R, Tegelaar H. A method for accurate placement of chest electrodes in the taking of serial electrocardiographic tracings. Canad Med Assoc J 1960;82:258.

[21] Rautaharju PM, Wolf HK, Eifler WJ, Blackburn H. A simple procedure for positioning precordial ECG and VCG electrodes using an electrode locator. J Electrocardiol 1976;9(1):35–40.

[22] Herman MV, Ingram DA, Levy JA, Cook JR, Athans RJ. Variability of electrocardiographic precordial lead placement: a method to improve accuracy and reliability. Clin Cardiol 1991;14(6):469–76.

[23] Willems JL, Poblete PF, Pipberger HV. Day-to-day variation of the normal orthogonal electrocardiogram and vectorcardiogram. Circulation 1972;45(5):1057–64.

[24] Rautaharju PM, Park L, Rautaharju FS, Crow R. A standardized procedure for locating and documenting ECG chest electrode positions: consideration of the effect of breast tissue on ECG amplitudes in women. J Electrocardiol 1998;31(1):17–29.

[25] Hill NE, Goodman JS. Importance of accurate placement of precordial leads in the 12-lead electrocardiogram. Heart Lung 1987;16(5):561–6.

[26] Mason RE, Likar I. A new system of multiple-lead exercise electrocardiography. Am Heart J 1966;71(2):196–205.

[27] Diamond D, Griffith DH, Greenberg ML, Carleton RA. Torso mounted electrocardiographic electrodes for routine clinical electrocardiography. J Electrocardiol 1979;12(4):403–6.

[28] Kleiner JP, Nelson WP, Boland MJ. The 12-lead electrocardiogram in exercise testing. A misleading baseline? Arch Intern Med 1978;138(10):1572–3.

[29] Rautaharju PM, Prineas RJ, Crow RS, Seale D, Furberg C. The effect of modified limb electrode positions on electrocardiographic wave amplitudes. J Electrocardiol 1980;13(2):109–13.

[30] Gamble P, McManus H, Jensen D, Froelicher V. A comparison of the standard 12-lead electrocardiogram to exercise electrode placements. Chest 1984;85(5):616–22.

[31] Papouchado M, Walker PR, James MA, Clarke LM. Fundamental differences between the standard 12-lead electrocardiograph and the modified (Mason-Likar) exercise lead system. Eur Heart J 1987;8(7):725–33.

[32] Edenbrandt L, Pahlm O, Sornmo L. An accurate exercise lead system for bicycle ergometer tests. Eur Heart J 1989;10(3):268–72.

[33] Takuma K, Hori S, Sasaki J, Shinozawa Y, Yoshikawa T, Handa S, et al. An alternative limb lead system for electrocardiographs in emergency patients. Am J Emerg Med 1995;13(5):514–7.

[34] Fesmire FM. Modified limb lead placement for rapid ECG. Am J Emerg Med 1995; 13(5):603–4.

[35] Chanarin N, Caplin J, Peacock A. "Pseudo reinfarction": a consequence of electrocardiogram lead transposition following myocardial infarction. Clin Cardiol 1990; 13(9):668–9.

[36] Guijarro-Morales A, Gil-Extremera B, Maldonado-Martin A. ECG diagnostic errors due to improper connection of the right arm and leg cables. Int J Cardiol 1991; 30(2):233–5.

[37] Castellanos A, Saoudi NC, Schwartz A, Sodi-Pallares D. Electrocardiographic patterns resulting from improper connections of the right leg (ground) cable. Pacing Clin Electrophysiol 1985;8(3 Pt 1):364–8.

[38] Peberdy MA, Ornato JP. Recognition of electrocardiographic lead misplacements. Am J Emerg Med 1993;11(4):403–5.

[39] Haisty W. K. J, Pahlm O, Edenbrandt L, Newman K. Recognition of electrocardiographic electrode misplacements involving the ground (right leg) electrode. Am J Cardiol 1993;71(16):1490–5.

[40] Macfarlane P, Lawrie T. The normal electrocardiogram and vectorcardiogram. In: Macfarlane P, Lawrie TDV, editors. Comprehensive Electrocardiology, vol. 1, 1st ed. New York: Pergamon Press, 1989; pp. 407–57.

[41] Hedén B, Ohlsson M, Edenbrandt L, Rittner R, Pahlm O, Peterson C. Artificial neural networks for recognition of electrocardiographic lead reversal. Am J Cardiol 1995;75(14):929–33.

[42] Hedén B, Ohlsson M, Holst H, Mjoman M, Rittner R, Pahlm O, et al. Detection of frequently overlooked electrocardiographic lead reversals using artificial neural networks. Am J Cardiol 1996;78(5):600–4.

[43] Abdollah H, Milliken JA. Recognition of electrocardiographic left arm/left leg lead reversal. Am J Cardiol 1997;80(9):1247–9.

[44] Cooper RA, Klatsky AL. Left arm/left leg electrocardiographic lead reversal: report of a "natural" experiment. Am J Cardiol 1998;81(6):808.

[45] Edenbrandt L, Rittner R. Recognition of lead reversals in pediatric electrocardiograms. Am J Cardiol 1998;82(10):1290–2, A10.

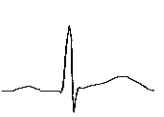
[46] Berson AS, Pipberger HV. Effects of electrode area on electrocardiographic voltages. J Electrocardiol 1978;11(2):147–50.

[47] Zywietz C. Technical aspects of electrocardiogram recording. In: Macfarlane P, Lawrie TDV, editors. Comprehensive electrocardiography. Theory and practice in health and disease, vol. 2, 1st ed. New York: Pergamon Press, 1989; pp. 353–404.

[48] Sippensgroenewegen A, Reek EJ, Spekhorst H, Broekhuijsen P, Smit HW, Grimbergen KA. A venturi suction electrode array for clinical body surface mapping. IEEE Trans Biomed Eng 1986;33(9):891–4.

[49] Patterson RP. The electrical characteristics of some commercial ECG electrodes. J Electrocardiol 1978;11(1):23–6.

[50] Clochesy JM, Ripich S, Hollowood D, Travis LL. Offset potential of commercial ECG electrodes. Heart Lung 1993;22(6):490–3.



[51] American Association for Medical Instrumentation. American National Standards for ECG Electrodes. (EC-12). Tech. rep., American Association for Medical Instrumentation, 1991.

[52] Berson AS, Pipberger HV. Skin-electrode impedance problems in electrocardiography. Am Heart J 1968;76(4):514–25.

[53] Almasi J, Schmitt O. Systematic and random variations of ECG electrode system impedance. Ann NY Acad Sci 1970;170:590–19.

[54] Olson WH, Schmincke DR, Henley BL. Time and frequency dependence of disposable ECG electrode-skin impedance. Med Instrum 1979;13(5):269–72.

[55] Huhta JC, Webster JG. 60-Hz interference in electrocardiography. IEEE Trans Biomed Eng 1973;20(2):91–101.

[56] Report of committee on electrocardiography, American Heart Association. Recommendations for standardization of leads and of specifications for instruments in electrocardiography and vectorcardiography. Circulation 1975;52:11–31.

[57] Klingler DR, Schoenberg AA, Worth NP, Egleston CF, Burkart JA. A comparison of gel-to-gel and skin measurements of electrode impedance. Med Instrum 1979; 13(5):266–8.

[58] Clochesy JM, Cifani L, Howe K. Electrode site preparation techniques: a follow-up study. Heart Lung 1991;20(1):27–30.

[59] Tam HW, Webster JG. Minimizing electrode motion artifact by skin abrasion. IEEE Trans Biomed Eng 1977;24(2):134–9.

[60] Ödman S, Öberg PA. Movement-induced potentials in surface electrodes. Med Biol Eng Comput 1982;20(2):159–66.

[61] Uter W, Schwanitz HJ. Contact dermatitis from propylene glycol in ECG electrode gel. Contact Dermatitis 1996;34(3):230–1.

[62] Corazza M, Maranini C, La Malfa W, Virgili A. Unusual suction-like contact dermatitis due to ECG electrodes. Acta Derm Venereol 1998;78(2):145.

[63] Maxwell J. Preparation of skin for electrocardiography. Br Med J 1957;1:41.

[64] Lewes D. Electrode jelly for electrocardiography. Br Heart J 1965;27:105–15.

[65] Martin A, Tiernan R, D'Arcy M, O'Brian E. Tap water instead of electrode jelly for electrocardiographic recording. Br Med J 1979;1(6161):454.

[66] Dehlin O, Hedenrud B, Lindberg B. Tap water instead of electrode jelly for electrocardiographic recording. Br Med J 1979;2:134.

[67] Birks M, Santamaria N, Thompson S, Amerena J. A clinical trial of the effectiveness of water as a conductive medium in electrocardiography. Aust J Adv Nurs 1992;10(2):10–3.

[68] Einthoven W. The different forms of the human electrocardiogram and their significance. Lancet 1912;1:853–61.

[69] Berson AS, Pipberger HV. Electrocardiographic distortions caused by inadequate high-frequency response of direct-writing electrocardiographs. Am Heart J 1967; 74(2):208–18.

[70] Bailey JJ, Berson AS, Garson A. J, Horan LG, Macfarlane PW, Mortara DW, et al. Recommendations for standardization and specifications in automated electrocardiography: bandwidth and digital signal processing. A report for health professionals by an ad hoc writing group of the Committee on Electrocardiography and Cardiac Electrophysiology of the Council on Clinical Cardiology, American Heart Association. Circulation 1990;81(2):730–9.

[71] Barr RC, Spach MS. Sampling rates required for digital recording of intracellular and extracellular cardiac potentials. Circulation 1977;55(1):40–8.

[72] Yamamoto H, Miyahara H, Domae A. Is a higher sampling rate desirable in the computer processing of the pediatric electrocardiogram? J Electrocardiol 1987; 20(5):321–8.

[73] Zywietz C, Spitzenberger CP, Wetjen A. A new approach to determine the sampling rate for ECGs. In: Ripley K, editor. Computers in Cardiology. Los Alamitos, CA: IEEE Comput Soc, 1983; pp. 261–4.

[74] Bailey JJ, Horton M, Itscoitz SB. A method for evaluating computer programs for electrocardiographic interpretation. III. Reproducibility testing and the sources of program errors. Circulation 1974;50:88–93.

[75] Zywietz C, Joseph G, Degani R. Data compression for computerized electrocardiography. In: Willems JL, editor. Proceedings of the first Working Conference of the SCP-ECG Project AIM 1015. Leuven: ACCO Press, 1990; pp. 95–136. Ee2161.

[76] Wan SW, Nguyen HT. 50Hz interference and noise in ECG recordings–a review. Australas Phys Eng Sci Med 1994;17(3):108–15.

[77] Zywietz C, Celikag D, Joseph G. Influence of ECG measurement accuracy on ECG diagnostic statements. J Electrocardiol 1996;29 Suppl:67–72.

[78] Mortara DW. Digital filters for ECG signals. In: Ostrow H, Ripley K, editors. Computers in Cardiology. Long Beach: IEEE Comput Soc, 1977; pp. 511–4.

[79] Mortara DW. Source consistency filtering. Application to resting ECGs. J Electrocardiol 1992;25 Suppl:200–6.

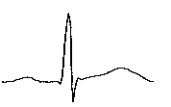
[80] Kors JA. Interference removal with an improved incremental estimation filter. Meth Inf Med 1993;33:15–8.

[81] Willems JL, Arnaud P, Van Bemmel JH, Degani R, Macfarlane PW, Zywietz C. Common standards for quantitative electrocardiography: goals and main results. CSE Working Party. Methods Inf Med 1990;29(4):263–71.

[82] Zywietz C, Willems JL, Arnaud P, Van Bemmel JH, Degani R, Macfarlane PW. Stability of computer ECG amplitude measurements in the presence of noise. The CSE Working Party. Comput Biomed Res 1990;23(1):10–31.

[83] Wilson FN, Kossmann CE, Burch GE, Goldberger E, Graybiel A, hecht HH, et al. Recommendations for standardization of electrocardiographic and vectorcardiographic leads. Circulation 1954;10:564.

[84] Report of committee on electrocardiography, American Heart Association. Recommendations for standardization of leads and of specifications for instruments in electrocardiography and vectorcardiography. Circulation 1967;35(3):583–602.



[85] Rao PS. Racial differences in electrocardiograms and vectorcardiograms between black and white adolescents. J Electrocardiol 1985;18(4):309–13.

[86] Rautaharju PM, Zhou SH, Calhoun HP. Ethnic differences in ECG amplitudes in North American white, black, and Hispanic men and women. Effect of obesity and age. J Electrocardiol 1994;27 Suppl:20–31.

[87] Chen CY, Chiang BN, Macfarlane PW. Normal limits of the electrocardiogram in a Chinese population. J Electrocardiol 1989;22(1):1–15.

[88] Gambill CL, Wilkins ML, Haisty W. K. J, Anderson ST, Maynard C, Wagner NB, et al. T wave amplitudes in normal populations. Variation with ECG lead, sex, and age. J Electrocardiol 1995;28(3):191–7.

[89] Raitt MH, Litwin PF, Martin JS, Weaver WD. ECG findings in acute myocardial infarction. Are there sex-related differences? J Electrocardiol 1995;28(1):13–6.

[90] Rautaharju PM, Zhou SH, Wong S, Calhoun HP, Berenson GS, Prineas R, et al. Sex differences in the evolution of the electrocardiographic QT interval with age. Can J Cardiol 1992;8(7):690–5.

[91] Durakovic Z, Mimica M. The electrocardiogram according to age and sex in a population. Indian Heart J 1982;34(4):221–4.

[92] Dougherty JD, Stoudt HW. The relation of frontal QRS axis to age and body build. J Electrocardiol 1970;3(3):285–98.

[93] Yasumura S, Shibata H. The effect of aging on the electrocardiographic findings in the elderly–a 10-year longitudinal study: the Koganei Study. Arch Gerontol Geriatr 1989;9(1):1–15.

[94] Davignon A, Rautaharju P, Boisselle E, Soumis F, Megelas M, Choguette A. Normal ECG standards for infants and children. Pediatr Cardiol 1979/80;1:123–31.

[95] Macfarlane PW, Coleman EN, Pomprhey EO, McLaughlin S, Houston A, Aitchison T. Normal limits of the high-fidelity pediatric ECG. J Electrocardiol 1989;22 Suppl:162–8.

[96] Simonson E. The effect of age on the electrocardiogram. Am J Cardiol 1972; 29(1):64–73.

[97] Harlan W. R. J, Graybiel A, Mitchell RE, Oberman A, Osborne RK. Serial electrocardiograms: their reliability and prognostic validity during a 24-yr. period. J Chronic Dis 1967;20(11):853–67.

[98] Bachman S, Sparrow D, Smith LK. Effect of aging on the electrocardiogram. Am J Cardiol 1981;48(3):513–6.

[99] Simonson E, Keys A. The effect of age and body weight on the electrocardiogram of healthy men. Circulation 1952;6:749–61.

[100] Okajima M, Scholmerich P, Simonson E. Frequency of premature beats. Minn Med 1960;43:751–3.

[101] Hiss R, Lamb L, Allen M. Electrocardiographic findings in 67,375 asymptomatic subjects. X. Normal values. Am J Cardiol 1960;6:200–31.

[102] Kilty SE, Lepeschkin E. Effect of body build on the QRS voltage of the electrocardiogram in normal men. Its significance in the diagnosis of left ventricular hypertrophy. Circulation 1965;31:77–84.

[103] Butrous GS, Baithun SI. Quantitative analysis of the electrocardiogram among healthy adult Iraqi males. Acta Cardiol 1982;37(6):387–99.

[104] Selzer A, Ebnother CL, Packard P, Stone AO, Quinn JE. Reliability of electrocardiographic diagnosis of left ventricular hypertrophy. Circulation 1958;17:255.

[105] Pipberger HV, Goldman MJ, Littmann D, Murphy GP, Cosma J, Snyder JR. Correlations of the orthogonal electrocardiogram and vectorcardiogram with consitutional variables in 518 normal men. Circulation 1967;35(3):536–51.

[106] Tuna N, Liedtke CE. Correlations between the electrocardiogram and age, height, body weight and chest circumference. Adv Cardiol 1977;19:17–20.

[107] Dougherty JD. The relation of QRS amplitude to the frontal QRS axis and the heart-electrode distance. J Electrocardiol 1971;4(3):249–60.

[108] Feldman T, Borow KM, Neumann A, Lang RM, Childers RW. Relation of electrocardiographic R-wave amplitude to changes in left ventricular chamber size and position in normal subjects. Am J Cardiol 1985;55(9):1168–74.

[109] Brohet CR, Tuna N. Quantative analysis of the vectorcardiogram in obesity. The effects of weight reduction. J Electrocardiol 1975;8(1):1–11.

[110] Devereux RB, Phillips MC, Casale PN, Eisenberg RR, Kligfield P. Geometric determinants of electrocardiographic left ventricular hypertrophy. Circulation 1983; 67(4):907–11.

[111] Levy D, Labib SB, Anderson KM, Christiansen JC, Kannel WB, Castelli WP. Determinants of sensitivity and specificity of electrocardiographic criteria for left ventricular hypertrophy. Circulation 1990;81(3):815–20.

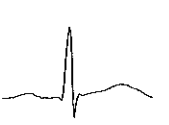
[112] Norman J. E. J, Levy D, Campbell G, Bailey JJ. Improved detection of echocardiographic left ventricular hypertrophy using a new electrocardiographic algorithm. J Am Coll Cardiol 1993;21(7):1680–6.

[113] Abergel E, Tase M, Menard J, Chatellier G. Influence of obesity on the diagnostic value of electrocardiographic criteria for detecting left ventricular hypertrophy. Am J Cardiol 1996;77(9):739–44.

[114] Levy D, Anderson KM, Savage DD, Kannel WB, Christiansen JC, Castelli WP. Echocardiographically detected left ventricular hypertrophy: prevalence and risk factors. The Framingham Heart Study. Ann Intern Med 1988;108(1):7–13.

[115] De Simone G, Daniels SR, Devereux RB, Meyer RA, Roman MJ, de Divitiis O, et al. Left ventricular mass and body size in normotensive children and adults: assessment of allometric relations and impact of overweight. J Am Coll Cardiol 1992;20(5):1251–60.

[116] De Simone G, Devereux RB, Roman MJ, Alderman MH, Laragh JH. Relation of obesity and gender to left ventricular hypertrophy in normotensive and hypertensive adults. Hypertension 1994;23(5):600–6.



[117] Alexander J. The heart and obesity. In: Hurst J, Schlant R, Rackley C, Sonnenblick E, Wenger N, editors. The heart. Arteries and veins, 7th ed. New York: McGraw-Hill Inc, 1990; pp. 1538–42.

[118] Okin PM, Roman MJ, Devereux RB, Kligfield P. Electrocardiographic identification of left ventricular hypertrophy: test performance in relation to definition of hypertrophy and presence of obesity. J Am Coll Cardiol 1996;27(1):124–31.

[119] Norman J. E. J, Levy D. Adjustment of ECG left ventricular hypertrophy criteria for body mass index and age improves classification accuracy. The effects of hypertension and obesity. J Electrocardiol 1996;29 Suppl:241–7.

[120] Gasowski J, Grodzicki T, Korman W, Adamkiewicz-Piejko A, Kocemba J. Inclusion of obesity indices improves sensitivity of ECG in detecting left ventricular hypertrophy. J of Hypertension 1998;16 Suppl:225.

[121] Huiskamp GJ, Van Oosterom A. Heart position and orientation in forward and inverse electrocardiography. Med Biol Eng Comput 1992;30(6):613–20.

[122] Hoekema R, Uijen GJ, Van Erning L, Van Oosterom A. Interindividual variability of multilead electrocardiographic recordings: influence of heart position. J Electrocardiol 1999;32(2):137–48.

[123] Dougherty JD. Change in the frontal QRS axis with changes in the anatomic positions of the heart. J Electrocardiol 1970;3(3):299–308.

[124] LaMonte CS, Freiman AH. The electrocardiogram after mastectomy. Circulation 1965;32(5):746–54.

[125] Horton JD, Sherber HS, Lakatta EG. Distance correction for precordial electrocardiographic voltage in estimating left ventricular mass: an echocardiographic study. Circulation 1977;55(3):509–12.

[126] McAnulty J, Metcalfe J, Ueland K. Heart disease and pregnancy. In: Hurst J, Schlant R, Rackley C, Sonnenblick E, Wenger N, editors. The heart. Arteries and veins, 7th ed. New York: McGraw-Hill Inc, 1990; pp. 1465–78.

[127] Metcalfe J, Ueland K. Maternal cardiovascular adjustments to pregnancy. Prog Cardiovasc Dis 1974;16(4):363–74.

[128] Carruth JE, Mivis SB, Brogan DR, Wenger NK. The electrocardiogram in normal pregnancy. Am Heart J 1981;102(6 Pt 1):1075–8.

[129] Wenger N, Hurst J, Strozier V. Electrocardiographic changes in pregnancy. Am J Cardiol 1964;14:774.

[130] Schwartz DB, Schamroth L. The effect of pregnancy on the frontal plane QRS axis. J Electrocardiol 1979;12(3):279–81.

[131] Hollander A, Crawford J. Roentgenologic and electrocardiographic changes in the normal heart during pregnancy. Am Heart J 1943;26:364.

[132] Benjamin D. Duration of the QT interval in normal pregnant women. Am Heart J 1949;38:119.

[133] Oram S, Holt M. Innocent depression of the S-T segment and flattening of the T-wave during pregnancy. Obstet Gynecol 1961;68:765–70.

[134] Van Doorn MB, Lotgering FK, Struijk PC, Pool J, Wallenburg HC. Maternal and fetal cardiovascular responses to strenuous bicycle exercise. Am J Obstet Gynecol 1992;166(3):854–9.

[135] Asher UA, Ben-Shlomo I, Said M, Nabil H. The effects of exercise induced tachycardia on the maternal electrocardiogram. Br J Obstet Gynaecol 1993;100(1):41–5.

[136] Veille JC, Kitzman DW, Bacevice AE. Effects of pregnancy on the electrocardiogram in healthy subjects during strenuous exercise. Am J Obstet Gynecol 1996; 175(5):1360–4.

[137] Upshaw C. B. J. A study of maternal electrocardiograms recorded during labor and delivery. Am J Obstet Gynecol 1970;107(1):17–27.

[138] Brody DA. A theoretical analysis of intracavitary blood mass influence on the heart-lead relationship. Circ Res 1956;4:731–8.

[139] Rudy Y, Plonsey R. A note on the "Brody-effect". J Electrocardiol 1978;11(1):87–90.

[140] Van Oosterom A, Plonsey R. The Brody effect revisited. J Electrocardiol 1991; 24(4):339–48.

[141] Ishikawa K, Shirato C, Yanagisawa A. Electrocardiographic changes due to sauna bathing. Influence of acute reduction in circulating blood volume on body surface potentials with special reference to the Brody effect. Br Heart J 1983;50(5):469–75.

[142] Talbot S, Kilpatrick D, Jonathan A, Raphael MJ. QRS voltage of the electrocardiogram and Frank vectorcardiogram in relation to ventricular volume. Br Heart J 1977;39(10):1109–13.

[143] Amoore JN. The Brody effect and change of volume of the heart. J Electrocardiol 1985;18(1):71–5.

[144] Einthoven W, Fahr G, De Waart A. Über die Richtung und die manifeste Grosse der Potentialschwankungen im menschlichen Herzen und über den Einfluss der Herzlage auf die Form des Elektrocardiogramms. Pflug Arch ges Phys 1913; 150:275.

[145] Dougherty JD. The relation of respiratory changes in the horizontal QRS and T-wave axes to movement of the thoracic electrodes. J Electrocardiol 1970;3(1):77–85.

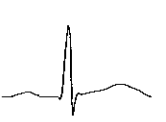
[146] Riekkinen H, Rautaharju P. Body position, electrode level, and respiration effects on the Frank lead electrocardiogram. Circulation 1976;53(1):40–5.

[147] Yamada N. Effects of respiration on the vectorcardiogram obtained with the Frank lead system. Acta Med Okayama 1985;39(4):297–313.

[148] Sutherland DJ, McPherson DD, Spencer CA, Armstrong CS, Horacek BM, Montague TJ. Effects of posture and respiration on body surface electrocardiogram. Am J Cardiol 1983;52(5):595–600.

[149] Singh MM, Rizvi SI, Bhusan S, Bagga CB, Seth P. Effect of respiration of mean frontal P. QRS and T axis in normal young adults. Indian J Physiol Pharmacol 1980;24(2):131–5.

[150] Lawrence GP, Home PD, Murray A. Repeatability of measurements and sources of variability in tests of cardiovascular autonomic function. Br Heart J 1992; 68(2):205–11.



[151] Mimbs JW, deMello V, Roberts R. The effect of respiration on normal and abnormal Q waves. An electrocardiographic and vectorcardiographic analysis. Am Heart J 1977;94(5):579–84.

[152] Pallas-Areny R, Colominas-Balague J, Rosell FJ. The effect of respiration-induced heart movements on the ECG. IEEE Trans Biomed Eng 1989;36(6):585–90.

[153] Ikeda K, Kubota I, Yasui S. Effects of lung volume on body surface electrocardiogram. Isointegral analysis of body surface maps in patients with chronic pulmonary emphysema. Jpn Circ J 1985;49(3):284–91.

[154] Almasi JJ, Schmitt OH. Basic technology of voluntary cardiorespiratory synchronization in electrocardiology. IEEE Trans Biomed Eng 1974;21(4):264–73.

[155] Hirsch JA, Bishop B. Respiratory sinus arrhythmia in humans: how breathing pattern modulates heart rate. Am J Physiol 1981;241(4):H620–9.

[156] Thayer JF, Peasley C, Muth ER. Estimation of respiratory frequency from autoregressive spectral analysis of heart period. Biomed Sci Instrum 1996;32:93–9.

[157] Noshiro M, Inaoka H, Yoshida Y, Ryumae S, Shimizu M, Hoshino H, et al. Indirect measurement of tidal volume by multi-lead electrocardiograms: effect of filtering, deep breath, lead and posture. Front Med Biol Eng 1996;7(3):207–19.

[158] Hirsch JA, Bishop B, York JL. Recovery of respiratory sinus arrhythmia in detoxified alcoholic subjects. J Appl Physiol 1993;74(4):1816–23.

[159] Fischmann E, Cosma J, Pipberger HV. Beat to beat and observer variation of the electrocardiogram. Am Heart J 1968;75:465–73.

[160] Oakley D. Cardiac hypertrophy in athletes. Br Heart J 1984;52(2):121–3.

[161] Oakley DG, Oakley CM. Significance of abnormal electrocardiograms in highly trained athletes. Am J Cardiol 1982;50(5):985–9.

[162] Bjørnstad H, Storstein L, Meen HD, Hals O. Electrocardiographic findings in athletic students and sedentary controls. Cardiology 1991;79(4):290–305.

[163] Stolt A, Kujala UM, Karjalainen J, Viitasalo M. Electrocardiographic findings in female endurance athletes. Clin J Sport Med 1997;7(2):85–9.

[164] Storstein L, Bjornstad H, Hals O, Meen HD. Electrocardiographic findings according to sex in athletes and controls. Cardiology 1991;79(3):227–36.

[165] Bjørnstad H, Smith G, Storstein L, Meen HD, Hals O. Electrocardiographic and echocardiographic findings in top athletes, athletic students and sedentary controls. Cardiology 1993;82(1):66–74.

[166] Bjørnstad H, Storstein L, Meen HD, Hals O. Electrocardiographic findings of left, right and septal hypertrophy in athletic students and sedentary controls. Cardiology 1993;82(1):56–65.

[167] Bjørnstad H, Storstein L, Dyre Meen H, Hals O. Electrocardiographic findings according to level of fitness and sport activity. Cardiology 1993;83(4):268–79.

[168] Bjørnstad H, Storstein L, Meen HD, Hals O. Electrocardiographic findings of repolarization in athletic students and control subjects. Cardiology 1994;84(1):51–60.

[169] Bjørnstad H, Storstein L, Meen HD, Hals O. Ambulatory electrocardiographic findings in top athletes, athletic students and control subjects. Cardiology 1994; 84(1):42–50.

[170] Simonson E, Alexander H, Henschel A, Keys A. The effect of meals on the electrocardiogram in normal subjects. Am Heart J 1946;32:202–14.

[171] Simonson E, McKinlay CA, Henschel A. The effect of meals on the electrocardiogram of cardiac patients. Proc Soc Exper Biol Med 1946;63:542.

[172] Colles P, Juneau M, Gregoire J, Larivee L, Desideri A, Waters D. Effect of a standardized meal on the threshold of exercise-induced myocardial ischemia in patients with stable angina. J Am Coll Cardiol 1993;21(5):1052–7.

[173] Shapiro W, Berson AS, Pipberger HV. Differences between supine and sitting Frank-lead electrocardiograms. J Electrocardiol 1976;9(4):303–8.

[174] Bergman KS, Stevenson WG, Tillisch JH, Stevenson LW. Effect of body position on the diagnostic accuracy of the electrocardiogram. Am Heart J 1989;117(1):204–6.

[175] Nørgaard BL, Rasmussen BM, Dellborg M, Thygesen K. Temporal and positional variability of the ST segment during continuous vectorcardiography monitoring in healthy subjects. J Electrocardiol 1999;32(2):149–58.

[176] Adams MG, Drew BJ. Body position effects on the ECG: implication for ischemia monitoring. J Electrocardiol 1997;30(4):285–91.

[177] Jernberg T, Lindahl B, Högberg, Wallentin L. Effects on QRS-waveforms and ST-T-segment by changes in body position during continuous 12-lead ECG, a preliminary report. Comput Cardiol 1997;24:461.

[178] Schiffer F, Hartley LH, Schulman CL, Abelmann WH. The quiz electrocardiogram: a new diagnostic and research technique for evaluating the relation between emotional stress and ischemic heart disease. Am J Cardiol 1976;37(1):41–7.

[179] Hijzen TH, Slangen JL. The electrocardiogram during emotional and physical stress. Int J Psychophysiol 1985;2(4):273–9.

[180] Ruttkay-Nedecky I. Effect of emotional stress on cardiac repolarization vectors. Adv Cardiol 1978;21:284–5.

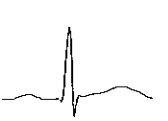
[181] Doi Y, Yoshida T, Hiroki T, Arakawa K. Effect of mental stress on R wave amplitude of the electrocardiogram in young healthy male subjects. Jpn Heart J 1983; 24(2):189–98.

[182] Ax A. The physiological differentiation between fear and anger in humans. Psychosom Med 1953;15:433–42.

[183] Huang MH, Ebey J, Wolf S. Responses of the QT interval of the electrocardiogram during emotional stress. Psychosom Med 1989;51(4):419–27.

[184] Rautaharju PM, Wolf H, Piironen P, Aikas E, Karvonen MJ. Thermal stress and the electrocardiogram: a technical study. J Appl Physiol 1966;21(6):1875–9.

[185] Kuaity J, Wexler H, Simonson E. The electrocardiographic ice water test. Am Heart J 1969;77(4):569–71.



[186] Kirchhoff KT, Holm K, Foreman MD, Rebenson-Piano M. Electrocardiographic response to ice water ingestion. Heart Lung 1990;19(1):41–8.

[187] Aigner A, Berghold F, Muss N. [Investigations on the cardiovascular system at altitudes up to a height of 7,800 meters (author's transl)]. Z Kardiol 1980;69(9):604–10.

[188] Das BK, Tewari SC, Parashar SK, Akhtar M, Grover DN, Ohri VC, et al. Electrocardiographic changes at high altitude. Indian Heart J 1983;35(1):30–3.

[189] Karliner JS, Sarnquist FF, Graber DJ, Peters R. M. J, West JB. The electrocardiogram at extreme altitude: experience on Mt. Everest. Am Heart J 1985;109(3 Pt 1):505–13.

[190] Malconian M, Rock P, Hultgren H, Donner H, Cymerman A, Groves B, et al. The electrocardiogram at rest and exercise during a simulated ascent of Mt. Everest (Operation Everest II). Am J Cardiol 1990;65(22):1475–80.

[191] Suarez J, Alexander JK, Houston CS. Enhanced left ventricular systolic performance at high altitude during Operation Everest II. Am J Cardiol 1987;60(1):137–42.

[192] Halperin BD, Sun S, Zhuang J, Droma T, Moore LG. ECG observations in Tibetan and Han residents of Lhasa. J Electrocardiol 1998;31(3):237–43.

[193] al Nashash H, Qassem W, Zabin A, Othman M. ECG response of the human body subjected to vibrations. J Med Eng Technol 1996;20(1):2–10.

[194] Qassem W, al Nashash H, Othman MO. Effect of vibrations with two different amplitudes on the ECG. J Med Eng Technol 1997;21(3-4):135–40.

[195] De Bruyne MC, Kors JA, Visentin S, Van Herpen G, Hoes AW, Grobbee DE, et al. Reproducibility of computerized ECG measurements and coding in a nonhospitalized elderly population. J Electrocardiol 1998;31(3):189–95.

[196] Van Den Hoogen JP, Mol WH, Kowsoleea A, Van Ree JW, Thien T, Van Weel C. Reproducibility of electrocardiographic criteria for left ventricular hypertrophy in hypertensive patients in general practice. Eur Heart J 1992;13(12):1606–10.

[197] Farb A, Devereux RB, Kligfield P. Day-to-day variability of voltage measurements used in electrocardiographic criteria for left ventricular hypertrophy. J Am Coll Cardiol 1990;15(3):618–23.

[198] Larkin H, Hunyor SN. Precordial voltage variation in the normal electrocardiogram. J Electrocardiol 1980;13(4):347–51.

[199] McManus CD, Doyle JT, Pipberger HV. Year-to-year variation of the orthogonal electrocardiogram and vectorcardiogram among 243 normal white males. J Electrocardiol 1984;17(2):107–14.

[200] Ishikawa K, Batchlor C, Pipberger HV. Reduction of electrocardiographic beat-to-beat variation through computer wave recognition. Am Heart J 1971;81(2):236–41.

[201] Willems JL, Zywietz C, Arnaud P, Van Bemmel JH, Degani R, Macfarlane PW. Influence of noise on wave boundary recognition by ECG measurement programs. Recommendations for preprocessing. Comput Biomed Res 1987;20(6):543–62.

[202] McLaughlin SC, Aitchison TC, Macfarlane PW. Methods for improving the repeatability of automated ECG analysis. Methods of Information in Medicine 1995; 34.

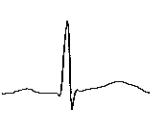
[203] Kors JA, Van Herpen G, Van Bemmel JH. Variability in ECG computer interpretation. Analysis of individual complexes vs analysis of a representative complex. J Electrocardiol 1992;25(4):263–71.

[204] Fayn J, Rubel P. CAVIAR: a serial ECG processing system for the comparative analysis of VCGs and their interpretation with auto-reference to the patient. J Electrocardiol 1988;21 Suppl:173–6.

[205] Sörnmo L. Vectorcardiographic loop alignment and morphologic beat-to-beat variability. IEEE Trans Biomed Eng 1998;45(12):1401–13.

[206] Appel ML, Berger RD, Saul JP, Smith JM, Cohen RJ. Beat to beat variability in cardiovascular variables: noise or music? J Am Coll Cardiol 1989;14(5):1139–48.

[207] Sörnmo L, Wohlfart B, Berg J, Pahlm O. Beat-to-beat QRS variability in the 12-lead ECG and the detection of coronary artery disease. J Electrocardiol 1998;31(4):336–44.



3 Interpolation of Body Surface Potential Maps

R.J.A. Schijvenaars*, J.A. Kors*, G. van Herpen*,
F. Kornreich†, J.H. van Bemmel*

* Dept. of Medical Informatics, Erasmus University Rotterdam, The Netherlands
† Unit of Cardiovascular Research and Engineering, Free University of Brussels, Belgium

Journal of Electrocardiology 1995;28:104–109

Abstract

The performance of four methods for interpolation of body surface potential maps (BSPMs) for different electrode grid densities was assessed. This study is part of a research project on the influence of the variability of 12-lead electrocardiograms on computer interpretation due to small electrode position changes. Interpolated BSPMs can be used to simulate this variability. The set of BSPMs studied, derived from a 117-electrode grid with relatively many electrodes on the left precordial part of the thorax, consisted of 232 cases without abnormalities, 277 with infarction, and 237 with left ventricular hypertrophy. The interpolation methods used were fast Fourier transforms, Chebyshev polynomials, linear functions, and cubic splines (*CS*). In the horizontal plane, a reference signal was first interpolated and, thereafter, resampled using 11 different sets of electrodes, the number of electrodes ranging from 18 down to 8. In the vertical direction, five grids with electrodes only on the front of the thorax and 9 grids with electrodes on the front and the back were examined. As a performance measure for interpolation, mean absolute error (*MAE*) was used: the absolute differences between the reference signal and the interpolated signal, averaged over the QRS on all maps. All methods showed deteriorating performance for decreasing grid density. In the horizontal direction, *CS* proved to be slightly superior to other methods for the left precordial electrodes for all but the densest grid (e.g., MAE = 22.8 μV vs. $MAE > 24.8$ μV for a 12-electrode grid). For electrodes not in that area, *CS* performed the best as well (MAE = 16.1 μV for the same grid), with differences with the other methods being small ($MAE > 16.4$ μV). In the vertical direction *CS* showed best results on the front, both for the dense non-periodic (MAE = 19.1 μV vs. $MAE > 26.6$ μV for a 6-electrode grid) and periodic grids (MAE = 25.1 μV vs. $MAE > 26.6$ μVfor a 12-electrode grid). Linear functions performed best for sparse non-periodic grids and sparse periodic grids for electrodes on the back, the difference with *CS* for the last case being small. The method *CS* performed best overall, and is recommended for interpolating BSPMs.

Keywords: body surface potential maps, electrode grids, fast Fourier transforms, Chebyshev polynomials, linear functions, cubic splines.

3.1 Introduction

The benefit of body surface potential maps (BSPMs) as a technique for obtaining more detailed electrocardiographic information than is possible with standard ECG recordings has been shown in various studies in the past (see De Ambroggi et al. [1] for an overview). The presentation of information is mostly in the form of isopotential surface maps, in which points of equal potential are joined to make isopotential contours. Recently, we have used BSPMs to study the effect of electrode position changes on ECG interpretation [2]. In that study, electrode position changes were simulated by generating ECGs from BSPMs.

To obtain detailed isopotential surface maps as well as to simulate electrode position changes, the number of measured locations needed is much larger than the number of electrodes that can be used in practice. Therefore, interpolation is necessary to achieve a detailed enough map. Research on which interpolation method to use has been done in the past by Monro and Attwood [3–5]; however, the maps to be interpolated in these studies were relatively sparse (eight electrodes in both directions) and consisted of electrodes placed on an equidistant grid. Because a large fraction of the information is located near the middle of the thorax, where the heart is closest, a nonuniform electrode grid may therefore yield more information. Attwood and Monro analyzed several interpolation methods, both for horizontal and vertical interpolation, but quantitative results on the performance of these methods were given for only 10 BSPMs of mainly normal subjects and for one grid density [4]. However, not all methods perform equally well for all grid densities, so the performance of the interpolation methods should be related to the grid density.

For some interpolation methods, the signals to be interpolated are assumed to be periodic. In the horizontal direction, the signals are space-periodic since traversal of a horizontal row ends where one started. In the vertical direction this is not the case, and Monro introduced spatial periodicity by using two additional electrodes, one near the umbilicus, and one on the neck [3]. These two extra electrodes were used to introduce periodicity for all vertical rows. Because such electrode compositions can introduce undesired aliasing frequencies in the spatial signal, using an interpolation method that does not require periodic signals may be a better alternative for vertical interpolation. This has not been investigated in previous studies.

The purpose of our study was to assess the performance of several methods for interpolation of BSPMs for different grid densities. In this study, some of the limitations of previous studies were avoided. First, the grid we used was not equidistant in terms of electrode distance, and the BSPMs we used comprised 117 torso electrodes and extended to the umbilical level. Many electrodes were placed in the left precordial area, as can be seen in Fig. 3.1 on the following page

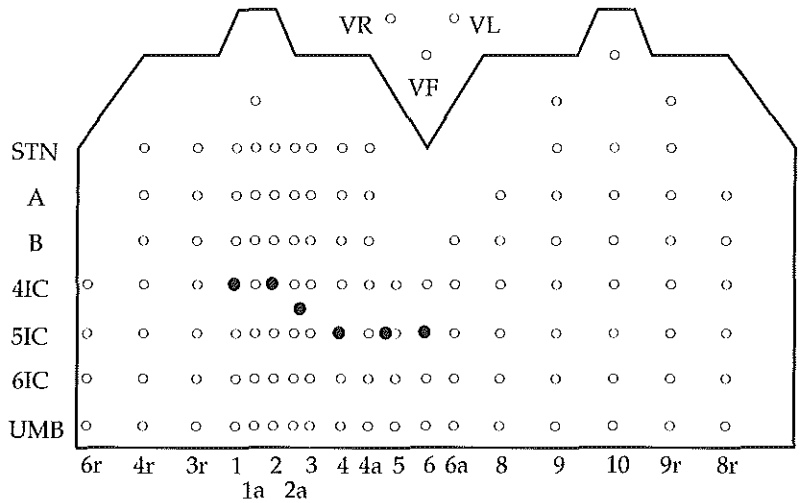


Figure 3.1. An unfolded thorax showing the body surface potential mapping electrode positions. The left part of the image denotes the front and the right part the back of the thorax. The electrode rows were placed 5 cm apart, with the fourth row from the bottom located on the fourth intercostal space (4IC). The top row is at the level of the sternal notch (STN) and the bottom row is near the umbilical level (UMB). The horizontal placement was done using anatomic landmarks [6]. The filled circles indicate the standard lead V_1–V_6 electrode positions. The three upper electrodes labeled VR, VL and VF denote the electrodes on the extremities.

that shows a schematic unfolded thorax with the BSPM lead positions. Second, we used a set of 746 BSPMs from healthy subjects and patients with infarction or left ventricular hypertrophy. Third, the performances of the methods could be related to grid density, since the methods were used to interpolate a number of different grids.

3.2 Materials and Methods

3.2.1 Body Surface Potential Maps

A set of 746 BSPMs was used for the experiments. The procedure for recording the BSPMs has been described previously [6,7]. A BSPM consisted of 117 torso and 3 limb electrode sites. Figure 3.1 shows a schematic unfolded thorax with lead positions. The set of BSPMs consisted of 232 (31%) cases without

abnormalities (normal subjects), 277 (37%) with infarction, and 237 (32%) with left ventricular hypertrophy. These diagnoses were assessed on the basis of ECG-independent evidence [7]. For each electrode position, a representative complex was obtained by coherent averaging.

3.2.2 Interpolation Methods

Four interpolation methods have been studied: interpolation using fast Fourier transforms, Chebyshev series interpolation, linear interpolation, and cubic spline interpolation. In the following, it is assumed that a signal x of length N (samples x_0 to x_{N-1}) is interpolated to obtain a signal y of length pN (samples y_0 to y_{pN-1}), with $p \in \mathbb{N}$.

Fourier Interpolation

For Fourier interpolation, a signal is transformed using the discrete Fourier transform: zeros are added in the middle of the transformed signal and this signal is then transformed back again to yield an interpolated signal (see Rabiner and Gold [8] for a more elaborate discussion of the method). Let X and Y be the discrete fast Fourier transforms of x and y respectively. For N even, Y can be computed from X using:

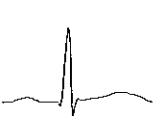
$$Y_r = \begin{cases} X_r & \text{for } r = 0, \ldots, N/2-1, \\ X_r/2 & \text{for } r = N/2, \\ 0 & \text{for } r = N/2+1, \ldots, pN/2-1. \end{cases} \tag{3.1}$$

The second half of Y_r (for $r = pN/2, \ldots, pN-1$) is defined as the Hermitian symmetric of the first half. For N odd, the second right-hand side expression is not used. To obtain the interpolated y (y_0 to y_{pN-1}), the inverse discrete fast Fourier transform is applied to Y.

Fourier interpolation is based on three assumptions. First, the sampling frequency should be sufficiently high to rule out aliasing. Second, the signal has to be periodic. Third, sampling must be equidistant. If one of these conditions is not met, the interpolation method will generate false harmonics.

Chebyshev Interpolation

Interpolation using Chebyshev series is closely related to Fourier interpolation. The method uses a least-squares polynomial fit. One of the definitions of the Chebyshev cosine series is (see, Monro [3] for more details):



$$C_n = \frac{2}{N} \sum_{j=0}^{N-1} x_j \cos\left\{\frac{(2j+1)n\pi}{2N}\right\}, \qquad n = 0, 1, \ldots, N-1, \tag{3.2}$$

where C is an array of N real Chebyshev coefficients. The interpolated sample y_k is:

$$y_k = \frac{C_0}{2} \sum_{n=1}^{N-1} C_n \cos\left\{n \arccos\left(\frac{k}{p}\right)\right\}. \tag{3.3}$$

Another way of describing the method is by using Fourier series [3]. Of the methods studied, only the Chebyshev and the Fourier interpolations use the complete signal for interpolating between samples. Like Fourier interpolation, Chebyshev polynomials require uniform sampling. Periodicity is not required for Chebyshev polynomials. Interpolation using Chebyshev polynomials has some useful properties: it results in an interpolated signal having a small maximum error while oscillations are evenly distributed across the signal.

Linear Interpolation

Linear interpolation uses only 'local' information to interpolate:

$$y_k = Ax_j + Bx_{j+1}, \qquad \text{for} \quad pj \leq k < p(j+1) \tag{3.4}$$

where

$$A = \frac{p(j+1) - k}{p}, \qquad B = 1 - A \tag{3.5}$$

Equations (3.4) and (3.5) show that the interpolation result between two samples x_j and x_{j+1} is independent of the values of other, nonadjacent samples.

Cubic Spline Interpolation

Spline interpolation is a method that divides a signal into sections that are separately interpolated using polynomials [9]. In fact, linear interpolation is a form of spline interpolation using polynomials of degree 1. In the case of cubic spline interpolation, the polynomials are cubic. The goal of cubic spline interpolation is to get an interpolated signal that is smooth in the first derivative and continuous in the second derivative, both within the interval and at its boundaries. Imposing these conditions on the coefficients of the interpolation polynomial between samples x_j and x_{j+1} results in:

$$y_k = Ax_j + Bx_{j+1} + Cx''_j + Dx''_{j+1} \qquad \text{for} \quad pj \leq k \leq p(j+1) \tag{3.6}$$

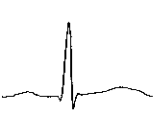
with A and B as in equation (3.5) on the facing page, and x'' is the second derivative. Further,

$$\begin{aligned} C &= \frac{p^2}{6}\left(A^3 - A\right), \\ D &= \frac{p^2}{6}\left(B^3 - B\right). \end{aligned} \tag{3.7}$$

Requiring continuity of the second derivative across interval boundaries gives $N-2$ equations in the N unknown x''_j, for $j = 0, \ldots, N-1$. To solve for the N unknowns, two more equations are required. The two are often specified as preset values for x'_0 and x'_{N-1}, or zero values for x''_0 and x''_{N-1}. When the signal is assumed to be periodic, N equations result because x''_0 and x''_{N-1} can then be computed in the same way as the other second derivatives using $x_j = x_{j \bmod N}$.

3.2.3 Experiments

The performance of the different interpolation methods was assessed separately for the vertical and horizontal directions. For the horizontal direction, the row at the level of the fifth intercostal space at V_4 was used (row 5IC in Fig. 3.1 on page 46). For the vertical direction, one column on the front of the thorax and one on the back (columns 4 and 8r in Fig. 3.1) were used. Figure 3.2 on the following page shows a schematic of the experimental setup. Based on the original samples in the row or column (indicated by the top array in Fig. 3.2), a set of densely spaced samples were derived, to be used as the underlying reference signal r (array 2). We assumed that the original samples were an



adequate representation of the underlying signal, so that the Fourier approach could be used to produce the underlying signal by exact interpolation. Different grid densities were then simulated by taking equidistant samples from this reference signal (array 3). The density of the densest grid thus constructed was equal to the density of the original grid. This signal was interpolated again by the different methods studied. The resulting interpolated signal s (array 4) was then compared to the reference signal r.

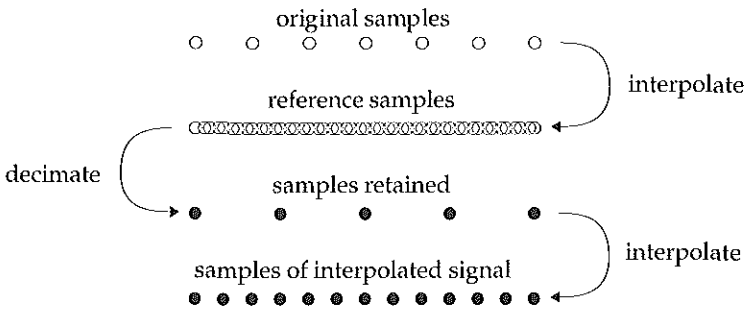


Figure 3.2. Experimental setup. The top array indicates the original samples in the BSPM. Interpolation of this original signal yields underlying reference samples (array 2). This reference signal is resampled by retaining a small subset of samples. The signal thus produced (array 3) is submitted to the four interpolation methods. The samples of the interpolated signal (array 4) are then compared to the corresponding samples in the reference signal (array 2).

The performance of each interpolation method was assessed by computing the mean absolute error (*MAE*) between the reference signal r and its interpolated counterpart s over the time-normalized QRS complex and over all BSPMs:

$$MAE = \frac{1}{MN} \sum_{n=1}^{N} \sum_{m=1}^{M} |s_{n,m} - r_{n,m}| \tag{3.8}$$

where M is the number of BSPMs used ($M = 746$), and N the number of time instances in the time-normalized QRS complex ($N = 35$). The *MAE*, like the root-mean-square error, is a global indicator for the similarity between two signals and has been chosen for the sake of comparability; Attwood and Monro used the same measurement in their experiments [4].

For each grid density, the reference signal was resampled several times with different phase, that is, with a slightly different offset. The resulting *MAE* values were averaged, thus increasing the accuracy of the *MAE* values.

Horizontal Experiments

In the horizontal experiments, the reference signal was based on the samples from the 18 electrodes in row 5IC (Fig. 3.1 on page 46). The position of these electrodes relative to each other is indicated in a cross-section of the thorax in Fig. 3.3. The reference signal was obtained by interpolating this signal using fast Fourier transform, since the signal was space-periodic.

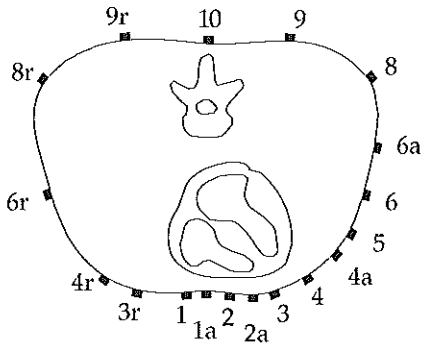


Figure 3.3. Horizontal cross-section of the thorax, showing the approximate locations of the electrode columns with respect to the heart center and the thorax center. The column names correspond with those in Fig. 3.1 on page 46.

The sizes of the grids studied were 18 down to 8 samples. These grids will be referred to as H18, H17, etc. The *MAE* values were computed for two subsets of the reference locations. The first subset consisted of the reference locations from column 1 to 6, that is, encompassing the left-precordial area. The other subset consisted of all other reference locations.

Vertical Experiments

In the vertical experiments, three reference signals were used: one periodic and two nonperiodic. The periodic reference signal was generated using the 16 electrode samples. In addition to the electrodes in columns 4 and 8r, the top electrodes in columns 9r, 1a and 10 were used. The interpolation method used to generate the reference signal was fast Fourier transform.

The sizes of the periodic grids studied were 16 down to 8 samples (referred to as VP16, VP15, etc.). The *MAE* values were averaged over two subsets of the reference locations. One subset (encompassing the front thorax) comprised the locations from the sternal notch to the umbilical level, the other one consisted of all reference locations in column 8r (A to the umbilical level).

The two nonperiodic reference signals were obtained by interpolation using the cubic spline method instead of the fast Fourier transform since the latter method assumes a periodic signal. The first nonperiodic reference signal was based on the samples in column 4 and the top electrode in column 1a. The other nonperiodic reference signal was based on the samples on the back of the thorax: all samples in column 8r and the top electrodes of columns 9r and 10.

The nonperiodic grids had sizes from eight down to four samples (VN8, VN7, etc.). The *MAE* values were averaged over the same subsets as in the periodic vertical experiment.

3.3 Results

Horizontal Experiments

Figures 3.4a, b on page 53 show that the performance of all interpolation methods decreased with grid density. The linear functions method performed worst on the left-precordial area for all grid densities. The fast Fourier transform and Chebyshev polynomials methods (the results of which were almost identical) performed worst on the back for most grids. Even though the reference signal was obtained by interpolation using fast Fourier transform, the cubic spline method performed best for all grids but H17 and H18.

Vertical Experiments

The results for the periodic vertical experiments are shown in Figs. 3.4c, d on page 53. The linear functions method performed worst for the electrodes on the front of the thorax but performed well on the back, where the fast Fourier transform and Chebyshev polynomials performed worst. Cubic spline interpolation performed relatively well for both areas.

Figures 3.4e, f on page 53 show the performances of the interpolation methods for the nonperiodic vertical grids. Note that only in these experiments was the reference signal obtained by interpolation using the cubic spline method. The performance of the Chebyshev polynomials method was relatively good on the front but worst on the back. Fast Fourier transform probably suffered from the fact that its condition of a periodic signal was not met. Cubic spline interpolation performed best for most grids. These results are difficult to compare with those of the periodic vertical experiments, since different reference signals were used.

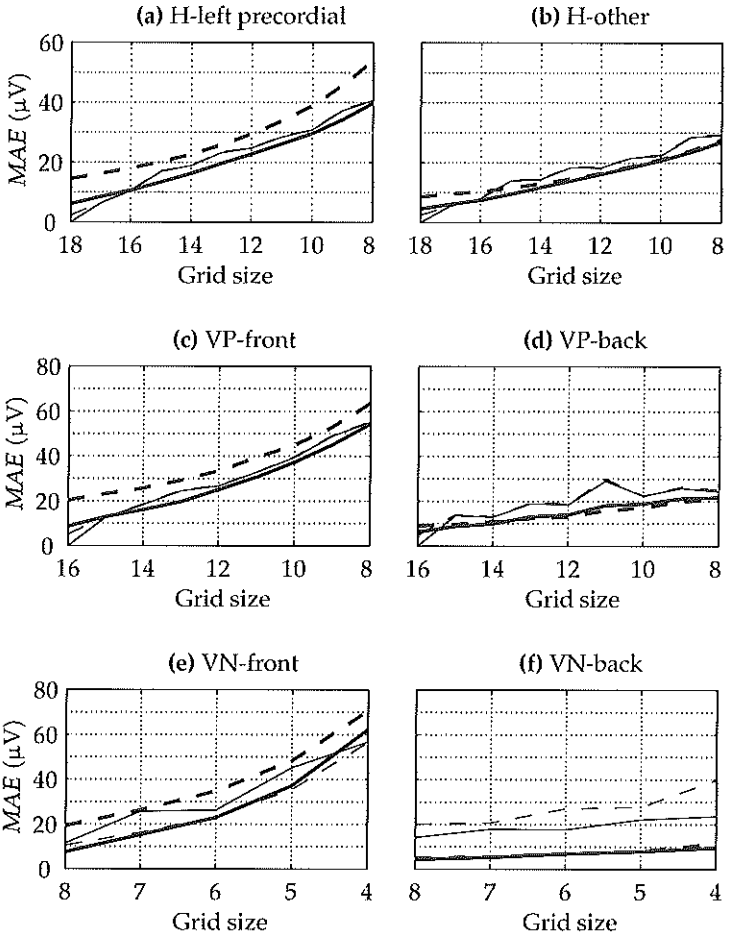


Figure 3.4. Mean absolute error (*MAE*) values of the different experiments as a function of grid size. (a) Horizontal interpolation, results for the electrodes in the left precordial region. (b) Horizontal interpolation, results for the other electrodes (not in the left precordial region). (c) Vertical interpolation of periodic signals, results for electrodes on the front of the thorax. (d) Vertical interpolation of periodic signals, results for electrodes on the back of the thorax. (e) Vertical interpolation of the nonperiodic reference signal, results for electrodes on the front of the thorax. (f) Vertical interpolation of nonperiodic signals, results for electrodes on the back of the thorax. Thin lines denote the results of fast Fourier transform (solid line) and Chebyshev polynomials (dashed line), thick lines denote the results of cubic spline interpolation (solid) and linear functions (dashed). If only one thick line is shown, fast Fourier transform and Chebyshev polynomials coincide.

The results of the experiments show a superior performance for fast Fourier transform and Chebyshev polynomials for the densest horizontal grids (H18 and H17) only. In most other grids, cubic spline interpolation performed better for the three experiments.

3.4 Discussion

The results suggested that the best overall interpolation method for electrode grids of different densities was cubic spline interpolation. Only for the densest periodic grids was its performance surpassed by that of the fast Fourier transform and Chebyshev methods. For fast Fourier transform, the *MAE* increase when going from grid VP15 to VP14 was much less than the *MAE* increase going from VP16 to VP15. This suggests a relatively high contribution to the signal power of the highest frequency sampled in the reference signal. That may have been due to disturbances generated by large differences between the electrodes at the umbilical level of columns 4 and 8r. This is supported by the fact that the difference between electrodes 6IC and UMB on column 1a (cf. Fig. 3.1 on page 46) averaged over all maps is lower than the averaged difference between electrodes UMB on columns 1a and 10: a 'jump' in the signal occurs. If so, then connecting columns in order to introduce periodicity will not improve accuracy.

For horizontal interpolation, Attwood and Monro reported *MAE* values averaged over 10 patients of 75.5, 48.6 and 54.1 μV for the fast Fourier transform, Chebyshev polynomials, and cubic spline methods, respectively for a grid of eight electrodes [4]. Corresponding values in this study can be obtained by averaging the *MAE* values of grid H8 over all reference electrodes. The resulting values are 34.3, 34.5 and 32.7 μV respectively. There is a striking difference between these two sets of values, which may be caused by a difference in the pattern of electrode locations. Attwood and Monro's horizontal grid was equidistant, meaning there were as many electrodes on the front of the thorax as on the back. The horizontal grids used in this study were not equidistant, since relatively many electrodes were located on the front of the thorax (cf. Figs. 3.1 and 3.3).

For the vertical periodic grid, Attwood reported *MAE* values of 44.0, 42.6 and 35.3 μV for the fast Fourier transform, Chebyshev polynomials, and cubic spline methods, respectively. Corresponding errors resulting from averaging the *MAE* values of grid VP8 over all reference electrodes were 47.2, 47.6, and 40.9 μV. In this case our results are slightly higher than those of Attwood.

It is difficult to assess which interpolation method is best for interpolation of even denser sampled BSPMs. In the horizontal direction Fourier interpolation

is theoretically preferred if the sampling frequency is high enough and the spatial signal not aliased. Antialias filtering of the spatial signal is not possible, however. Cubic spline interpolation performed best in most other grids and does not impose any requirements, so it might be an alternative. The differences between the performance of the methods for such dense grids will be small, however.

The question of how many electrodes to use for interpolating a BSPM depends on the accuracy desired. Typical root-mean-square noise values in the precordial leads are less than 15 μV [10]. This corresponds with an *MAE* value of about 20 μV for noise with a Gaussian distribution. For interpolation accuracy higher than this value the horizontal grid should consist of more than 13 electrodes. In vertical direction, 14 electrodes should suffice (7 on the front of the thorax and 7 on the back). Cubic spline interpolation is recommended. It should be noted that the performance of the interpolation methods was assessed using a quantitative measure of signal differences; however, the relationship between signal differences and changes in diagnostic interpretation is not straightforward. Thus, the accuracy of the methods studied may not faithfully represent their ability in preserving diagnostic information.

3.5 References

[1] De Ambroggi L, Musso E, Taccardi B. Body-surface mapping. In: Macfarlane PW, Lawrie TDV, editors. Comprehensive Electrocardiology. Oxford: Pergamon Press Inc, 1989; pp. 1015–49.

[2] Schijvenaars RJA, Kors JA, Van Herpen G, Kornreich F, Van Bemmel JH. The effect of electrode position changes on ECG interpretation by computer. J Electrocardiol 1997;30(3):247–56.

[3] Monro DM. Interpolation methods for surface mapping. Comput Program Biomed 1980;11:145–57.

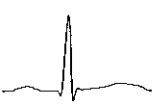
[4] Attwood RD, Monro DM. Reconstruction methods for body surface maps from 26 leads. In: Progress in Electrocardiology. 1979; .

[5] Monro DM, Guardo RAL, Bourdillon PJ, Tinker J. A Fourier technique for simultaneous electrocardiographic surface mapping. Cardiovascular Res 1974;8:688–700.

[6] Montague TJ, Smith ER, Cameron DA, Rautaharju PM, Klassen GA, Felmington CS, et al. Isointegral analysis of body surface maps: surface distribution and temporal variability in normal subjects. Circulation 1981;63(5):1166–72.

[7] Kornreich F, Montague TJ, Rautaharju PM. Body surface potential mapping of ST segment changes in acute myocardial infarction. Circulation 1993;87:773–82.

[8] Rabiner LR, Gold B. Theory and Application of Digital Signal Processing. Englewood Cliffs, New Jersey: Prentice-Hall, 1975.



[9] Press WH, Teukolsky SA, Vetterling WT, Flannery BP. Numerical Recipes in C. The art of scientific computing. Cambridge university press, 1988.

[10] Kors JA, Van Herpen G, Van Bemmel JH. Variability in ECG computer interpretations: analysis of individual complexes vs analysis of a representative complex. J Electrocardiol 1992;25:263–71.

Effect of Electrode Positioning on ECG Interpretation by Computer

R.J.A. Schijvenaars*, J.A. Kors*, G. van Herpen*,
F. Kornreich†, J.H. van Bemmel*

* Dept. of Medical Informatics, Erasmus University Rotterdam, The Netherlands
† Unit of Cardiovascular Research and Engineering, Free University of Brussels, Belgium

Journal of Electrocardiology 1997;30:247–256

Abstract

The aim of this study was to assess the variability in automated electrocardiogram (ECG) interpretation due to electrode positioning variations. Such variations were simulated by using a set of 746 body surface potential mappings from apparently healthy individuals and patients with myocardial infarction or left ventricular hypertrophy.

Four types of electrode position changes were simulated, and the effect on ECG measurements and diagnostic classifications was determined by a computer program.

At most 6% of the cases showed important changes in classification for longitudinal shifts. Transversal shifts caused less than 1.5% of important changes. An expert cardiologist, who analyzed a subset of 80 cases, agreed with the computer in 38 of 40 cases in which it made no change. In the 40 cases with large diagnostic changes, the cardiologist made no change in 18 cases.

The effect of electrode position changes on ECG classification by an expert cardiologist was about half of the effect determined by computerized ECG classification. The effects on classification are significant; therefore, correct placement of chest electrodes remains mandatory.

Keywords: electrocardiography, body surface potential mapping, computer-assisted ECG interpretation, precordial electrodes, variability

4.1 Introduction

Variation in the positioning of the chest electrodes may cause differences in electrocardiogram (ECG) interpretation, since different electrode positions result in different ECG signals. A recent study of precordial electrode variability showed that the average difference in electrode position is almost 3 cm [1]. Therefore, when interpreting ECGs, one should anticipate the possibility of electrode malposition. Of course, this applies to interpretations by both man and computer. Especially in serial ECG analysis (comparing ECGs of the same patient but recorded at different times), electrode displacements may generate differences in interpretation, not originating from changes in heart condition. Automatic serial ECG interpretation is becoming increasingly important, since more and more ECG interpretation systems are part of an ECG management system in which ECGs are stored and mutually compared on a routine basis. Also, longitudinal studies in epidemiology could benefit from automatic serial ECG analysis as opposed to automatic analysis of individual ECGs [2].

Some researchers have looked into the effects of changes in electrode position on measurements in the vectorcardiogram (VCG) [3–5]. These authors sequentially recorded VCGs using different electrode positions and compared measurements. These studies were limited in that only the VCG was studied, not the ECG, and that the VCGs were recorded sequentially, so that temporal variations of the heart itself (e.g., its position inside the thorax) or other factors (e.g., respiration) may have been additional sources of variation. Even beat-to-beat variability can be responsible for many changes [6–8]. In these studies, no attempts were made to assess changes in diagnostic classification.

Although the ECG has far wider acceptance in clinical practice than the VCG, only a few studies addressing ECG chest electrode displacements have been reported in the literature [9–11]. However, all ECGs were recorded sequentially in these studies. One study evaluated changes resulting from limb lead positions only [12]. To completely exclude variations caused by patient position, heart condition or beat-to-beat variation, ECGs from different electrode positions should be recorded simultaneously. Although the variability between recordings taken a few minutes one after another will be low, it is definitely zero only if all ECGs are recorded simultaneously instead of sequentially. Furthermore, because changes in diagnostic classification depend in a noncontinuous way on variations in ECG measurements, the diagnostic effects of position variations should be assessed as well. In the past, these effects were studied in a few (13) patients only [11], without relating them to the extent and direction of the lead displacements [10], or in lead reversal experiments only [9].

This study, which aims to investigate to what extent electrode position variations influence ECG classification, avoids the limitations of the earlier studies.

First, the ECG was studied and not the VCG. Second, to ensure simultaneous recording we simulated the displaced ECGs from each patient's body surface potential map (BSPM). Third, not only measurement changes were assessed, but also changes in diagnostic classification.

Both measurements and diagnostic classifications were obtained with our ECG analysis program MEANS (Modular ECG Analysis System) [13]. The program has been tested in the international study Common Standards for Quantitative Electrocardiography (CSE) [14]. In the CSE study the performances of nine computer programs and eight renowned human experts were compared. The MEANS program proved to have a higher than average performance for both waveform recognition and diagnostic classification [15,16]. In another study, the program also proved to have a good stability of measurements in the presence of noise [17].

In order to assess the importance of the variation found, a comparison was also made between variability induced by electrode displacements and reported year-to-year, day-to-day and beat-to-beat variations.

4.2 Materials and methods

For each individual, the 12-lead ECG taken from the BSPM at the standard precordial electrode positions was used as a reference. These standard precordial electrode positions are indicated with a solid circle in Fig. 4.1 on the facing page, which shows a schematic unfolded thorax with the BSPM electrode positions. As can be seen, the V_1, V_2, V_4 and V_6 positions are a regular part of the grid. For V_3 and V_5 however, interpolation is necessary from surrounding grid positions. Similarly, one can obtain an ECG from any desired set of new electrode positions by interpolating the signals for those positions that do not happen to coincide with grid electrode locations. Measurements and classifications of the "displaced" ECGs were then compared to those of their respective reference ECGs. A subset of 160 ECGs was interpreted by a human expert so that the variability in diagnostic classification of the program could be compared with that of the human expert. The specific method used to construct the subset is explained below.

4.2.1 Body Surface Potential Maps

The number of BSPMs used was 746. The recording procedure has been previously described [18,19]. The grid consisted of three limb electrodes and 117 torso sites, 81 anterior and 36 posterior (Fig. 4.1 on the next page). The set

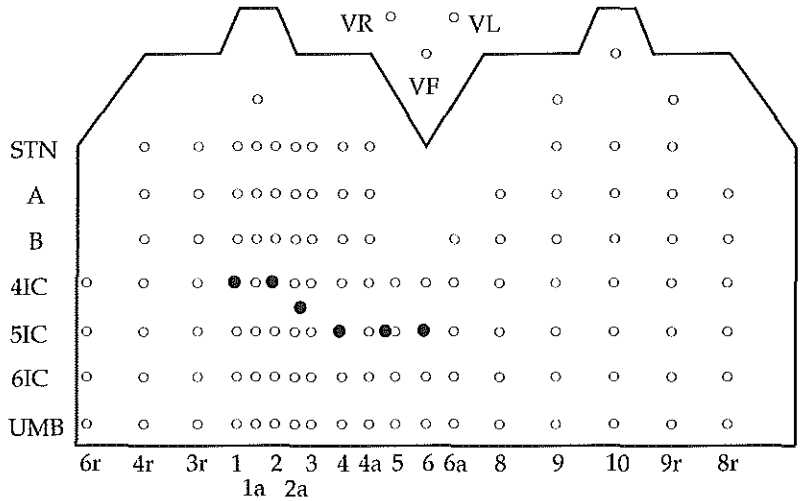


Figure 4.1. Schematic unfolded thorax, showing the BSPM electrode positions by open circles. The left part of the image represents the chest, the right part the back of the thorax. The electrode rows were placed 5 cm apart in longitudinal direction, with the fourth row from the bottom located in the fourth intercostal space 4IC. The top row is at the level of the sternal notch STN, and the bottom row is near the umbilicus UMB. The placement in the left-right direction was done by using anatomical landmarks [18]. The black dots indicate the standard V_1–V_6 electrode positions. Of these, V_3 and V_5 were not part of the recording system.

of BSPMs included 232 (31%) from apparently healthy individuals (normal subjects), 277 (37%) from patients with myocardial infarction, and 237 (32%) from patients with left ventricular hypertrophy. These diagnoses had been assessed by ECG-independent methods [20]. For the entire BSPM, one representative ECG complex was used, obtained by coherent averaging. All maps were checked for consistency and quality of the signals.

4.2.2 Body Surface Potential Map Interpolation

Interpolation of the BSPMs has the advantage that it enables the simulation of electrode positional changes without a time lag between the recording of the reference signal and the displacement ECG for any number of electrode positions. Moreover, it enables simulations of electrode position changes over distances that would be too small to be realized physically in simultaneous recordings (owing to, e.g., electrode overlap).

We examined different techniques for transverse and longitudinal interpolation. In the transverse direction the signals can be considered to be periodic in space, allowing Fourier interpolation. In the longitudinal direction, cubic spline interpolation was used. The mathematical background and validation of the transversal and longitudinal interpolations have been reported before [21]. It was shown that ECGs from any location can be accurately approximated by interpolation from a BSPM. Average absolute signal differences are in the order of a few microvolts.

4.2.3 Experiments

For simulating placement errors, the six chest electrodes were thought to be placed according to the following procedure. First, leads V_1 and V_2 are located on opposite sides close to the sternum in the fourth intercostal space. Then V_4 is placed on the left midclavicular line in the fifth intercostal space, and V_3 is placed halfway between V_2 and V_4. Next, lead V_6 is located on the left midaxillary line at the same level as V_4, and finally, V_5 is located halfway between V_4 and V_6. This method of electrode placement may lead to some general types of displacement in clinical routine, four of which were envisaged and simulated by experiments.

The first experiment (LONG) simulated an error, upward or downward, in determining the levels of the fourth and fifth intercostal space. This is also equivalent to a vertical (or longitudinal) shift of the heart position with respect to the thoracic cage, such as may occur by changing the patient's position after placing the electrodes (e.g., from sitting to supine). In this experiment, all chest electrodes, V_1–V_6, were shifted in a longitudinal direction, which is illustrated in Fig. 4.2(a) on the facing page. The variable Δ, which ranges from -5 cm (upward) to $+5$ cm (downward) in steps of 1 cm, indicates the distance over which the electrodes were moved.

Transverse displacements ware believed to follow two possible patterns. If leads V_1 and V_2 are placed at opposite sides close to the sternum, they are likely to be placed in the same transverse plane. However, the distance at which V_1 and V_2 are placed on opposite sides of the sternum may vary. Thus, in the second experiment, transverse displacements were studied for the lead group V_1–V_3, since shifting the V_2 position will result in shifting V_3 as well. This is illustrated in Fig. 4.2(b) on the next page (experiment TRANS-V123). The variable Δ ranges from -3 cm (V_1 leftward, V_2 and V_3 rightward) to $+3$ cm (V_1 rightward, V_2 and V_3 leftward) in steps of 6 mm.

The third experiment simulated an error in placing lead V_6 in position in the transverse plane. In the longitudinal direction, lead V_5 and V_6 will remain in

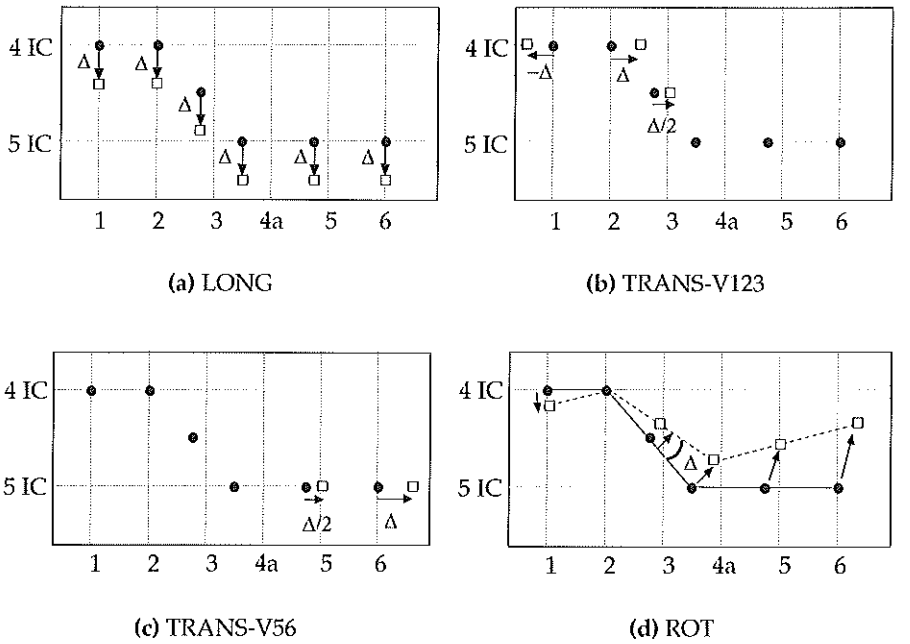


Figure 4.2. Electrode shifts in experiments (a) LONG, (b) TRANS-V123, (c) TRANS-V56, and (d) ROT. The standard positions are indicated by a black dot, the displaced positions for one specific (positive) displacement Δ by open squares.

line, but a transverse displacement of lead V_6 will influence the transversal position of V_5 as well. This experiment is illustrated in Fig. 4.2(c) (experiment TRANS-V56). The variable Δ ranges from -3 cm (rightward) to $+3$ cm (leftward) in steps of 6 mm.

In addition to these longitudinal and transverse displacements, a rotational dislocation was also performed. Rotation of the set of precordial electrodes about a center at lead V_2 is a simple simulation of rotation of the heart with respect to a fixed set of electrodes. This may occur by changes in heart orientation (e.g., a more transversally oriented heart in case of pregnancy or obesity or a more vertical heart position as in tall and slender persons). The topography of a longitudinally oriented heart is simulated by rotating the electrodes counterclockwise so that a larger angle between the apex and the electrode level is obtained. A more transversely positioned heart is simulated by a clockwise rotation of the electrode positions. Figure 4.2(d) shows the displacements studied in this

experiment. The variable Δ ranges from $+15°$ (counterclockwise rotation) to $-15°$ (clockwise rotation) in steps of 3°.

Table 4.1. Electrode displacement experiments

Experiment	V_1	V_2	V_3	V_4	V_5	V_6
LONG *	Δ	Δ	Δ	Δ	Δ	Δ
TRANS-V123 †	$-\Delta$	Δ	$\Delta/2$	—	—	—
TRANS-V56 †	—	—	—	—	$\Delta/2$	Δ
ROT ‡	rotation of V_1 and V_3–V_6 around V_2					

* In experiment LONG, the variable Δ assumes the values ± 50, ± 40, ± 30, ± 20, and ± 10 mm.

† In experiments TRANS-V123 and TRANS-V56, Δ is ± 30, ± 24, ± 18, ± 12, and ± 6 mm.

‡ In experiment ROT, the variable Δ takes on angles of $\pm 15°$, $\pm 12°$, $\pm 9°$, $\pm 6°$, and $\pm 3°$.

Table 4.1 summarizes the electrode displacements for the different experiments. The first three experiments resemble in part displacement findings in the Cornell study [1].

4.2.4 Electrocardiogram Processing

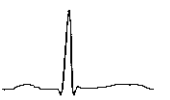
Each ECG generated from the BSPMs was processed with our ECG analysis program MEANS [13]. The program detects the onsets and ends of the P wave and of the QRS complex as well as the end of the T wave, and measures a large number of parameters. On the basis of these parameters it produces a diagnostic interpretation. The interpretative statements contain one of four qualifiers or levels of certainty: consider, possible, probable, or definite. From the set of parameters, four were selected for analysis: Q duration and Q amplitude, since these measurements are important for the diagnostic category myocardial infarction; the Sokolow index (S amplitude in V_1 + R amplitude in V_5), which is pertinent to left ventricular hypertrophy; and QRS duration, which was used to assess the stability of the program for wave durations. The classification categories studied were myocardial infarction and left ventricular hypertrophy. No distinction was made between the infarct localizations, and since the precordial leads are not used to diagnose inferior infarctions, these infarctions were not considered. The criteria used by the program to classify myocardial infarction are complicated and include, among others, Q duration and amplitude in different leads, QRS morphology, and repolarization abnormalities. The left ven-

tricular hypertrophy classification involves not only QRS amplitudes but also QRS axes, morphology, R-wave progression along the chest leads and repolarization strain features amounting to a kind of point score. For each category, the qualifiers were assigned the numbers 1 (for consider) to 4 (for definite). The number zero was used when the corresponding category was not mentioned. For each particular electrode displacement, the new diagnostic interpretation by the MEANS program was compared with that of the reference ECG. Changes in certainty levels were computed by subtracting qualifier points. For example, if an electrode displacement brought about a change from definite infarction to possible infarction and probable hypertrophy, this was coded as a change of -2 (from 4 to 2) in the myocardial infarction category, and a change of $+3$ (from 0 to 3) in the left ventricular hypertrophy category. Only such differences were considered, regardless of the levels of certainty from which they originated.

4.2.5 Comparison with Findings of Human Expert

Large changes in diagnostic classification by computer may be due to diagnostically significant changes in the ECG signals or to susceptibility of the classification algorithms to diagnostically insignificant signal changes. In the second case, a human expert will probably stay with the original diagnosis. To study whether a human expert would reproduce the classification changes of the program, the classification results of a human expert were compared with those of the program.

For both classification categories, myocardial infarction and left ventricular hypertrophy, 40 ECG pairs (one ECG originating from displaced electrodes and its reference counterpart) were randomly selected from those ECGs for which MEANS produced a large negative change (-3 or -4) in qualifiers (10 ECGs), no change (20 ECGs), and a large positive change ($+3$ or $+4$) (10 ECGs). No ECGs originating from extreme electrode shifts ($|\Delta| \geq 2.4$ cm for experiment TRANS-V123 and $\Delta = 3$ cm for TRANS-V56) were included, since such shifts are less likely to occur in practice (e.g., V_1 and V_2 at the same location). The 160 ECGs ($2 \times 40 \times 2$) were read by a cardiologist (GvH), in a blind and random order. His classifications were then scored in the same way as was done for the computer.



4.2.6 Variability due to electrode changes versus temporal variability

Our findings are difficult to compare with those of earlier studies because these either addressed specific VCG measurements (spatial amplitudes and axis orientations) [3,4,22] which cannot be compared with ours, or they applied to ECG measurements that were not included in this study [11]. Instead, a comparison was made with reported long-term (year-to-year), short-term (day-to-day), and beat-to-beat variation to assess the importance of the variations found.

4.3 Results

All measurements and classifications obtained from one electrode displacement were compared with those of the reference electrode locations shown in Fig. 4.1 on page 61. For each particular value of Δ (the displacement) in an experiment, a set of measurements and diagnostic classifications resulted for each subject. Differences were computed by subtracting the values originating from the reference ECG. The distribution of these differences over all subjects was computed for each displacement separately.

4.3.1 Measurements

When comparing measurements of Q, R, and S waves between different complexes, the problem of wave relabeling is encountered [17]. For example, if one complex is identified as an RSR′ and if in another complex the first R is not detected or has disappeared, the S and R′ waves are relabeled to Q and R waves. This complicates comparison of Q or S measurements between the complexes. Therefore, for Q durations and Q amplitudes only the complexes that contained a Q wave in the displaced ECG and/or the reference ECG were used, provided no wave relabeling occurred. The maximum percentage of cases with relabeling was 6.4%, which was reached for extreme displacements in experiment LONG. Note that the amplitude of a Q wave is taken to be negative. Thus, a negative change indicates a deeper Q wave in the displaced ECG.

For the measurements, Fig. 4.3 on the facing page shows the 2, 25, 50 (median), 75, and 98 percentiles of the distribution of changes. QRS duration is a measurement that should not change as a result of electrode position displacements. Figure 4.3a shows that most changes were indeed small; the 25 and 75 percentiles were at most -2 and $+2$ ms, respectively. Changes greater than 10 ms were due to measurement errors of the computer program.

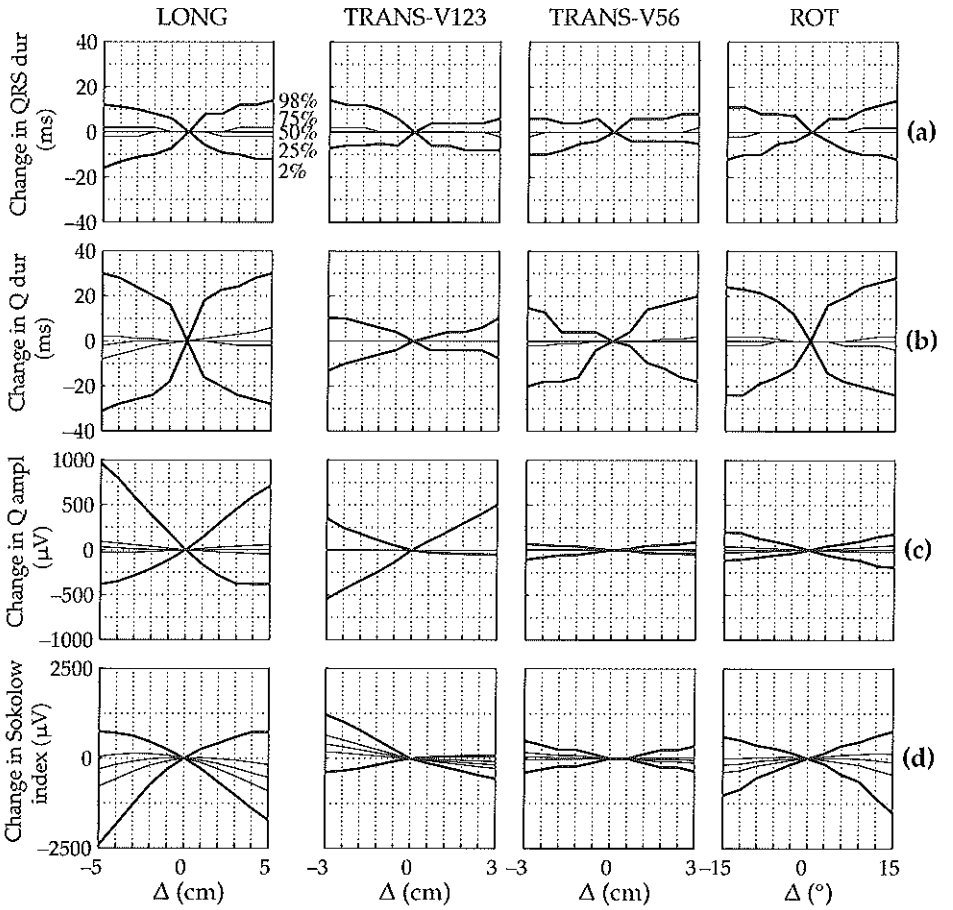


Figure 4.3. Percentiles (2, 25, 50, 75, and 98) of changes in (a) QRS duration, (b) Q duration, (c) Q amplitude and (d) Sokolow index as a function of the displacement Δ. Displacement is measured in centimeters (experiments LONG, TRANS-V123, and TRANS-V56) or in degrees (experiment ROT). The 2 and 98 percentiles are indicated by thick lines. A positive Δ means a downward shift in LONG, a leftward shift for V_2 and V_3 in TRANS-V123, a leftward shift in TRANS-V56, and a counterclockwise rotation in ROT.

The effect on Q duration (Fig. 4.3b), was in some cases quite large considering the large 2 and 98 percentiles. However, the 25 and 75 percentiles were low as compared with the 2 and 98 percentiles. The median was almost zero everywhere, indicating no change for the population as a whole. As was to be

expected, the effect on Q duration was larger than the effect on QRS duration. The distribution of changes in Q amplitudes (Fig. 4.3c) showed high 2 and 98 percentiles for experiments LONG and TRANS-V123. In all experiments the 25 and 75 percentiles were low.

In experiments LONG and ROT, large changes occurred in the Sokolow index (Fig. 4.3d). The median change was negative for all Δ. This means that at the standard electrode positions, the Sokolow values were maximal and decreased for both upward and downward displacements. In TRANS-V123 and TRANS-V56 the index increased for $\Delta < 0$ (V_1 to the left or V_5 to the right) and decreased for $\Delta > 0$. The decrease may have been due to an increased attenuation of the left ventricular forces caused by an increased amount of tissue between the electrode and the signal source. It should be noted, however, that an equal Δ in both experiments means a shift of $\Delta/2$ for V_5 (Fig. 4.2 on page 63). For all experiments, the distribution of changes was relatively smooth, since the 25 and 75 percentile lines were not very close to the median line.

4.3.2 Diagnostic Classifications

Figure 4.4 on the next page shows the percentage of cases that gave qualifier changes due to electrode displacements. The upper part of each plot indicates the percentages for positive changes in qualifier points (i.e., higher certainty for the displaced ECG), the lower part for negative changes (lower certainty). The thin lines denote the percentage for a change of $+1$ or $+2$ qualifier points (for the upper part of the plots) or a change of -1 or -2 (for the lower part). The thick lines show the same for changes of ± 3 or ± 4; they represent the more significant changes.

For the MI category (Fig. 4.4a), extreme electrode position shifts in experiment LONG resulted in large changes in classification in at most 6% of the cases. In fewer than 11% of the cases there was a small change. The experiments TRANS-V123 and TRANS-V56 produced large changes in less than 1.5% of the cases, while small changes occurred in at most 5.5%. Experiment ROT showed large changes in at most 3.5%, and small ones in less than 8.5% of the cases.

Changes in LVH (Fig. 4.4b) were in agreement with those of the Sokolow index (MEANS uses a version of the Sokolow index as one of its criteria). Large classification changes were made in at most 3% of the cases in experiment LONG. In the other experiments the percentage of cases with large changes was less than 1%, while small changes occurred in at most 7% of the cases. Similarly to the Sokolow plots, lower qualifier points (compared with the reference ECG) were produced for most ECGs in experiments LONG and ROT. A clear effect can be seen in the plot of experiment TRANS-V123. Small positive changes resulted

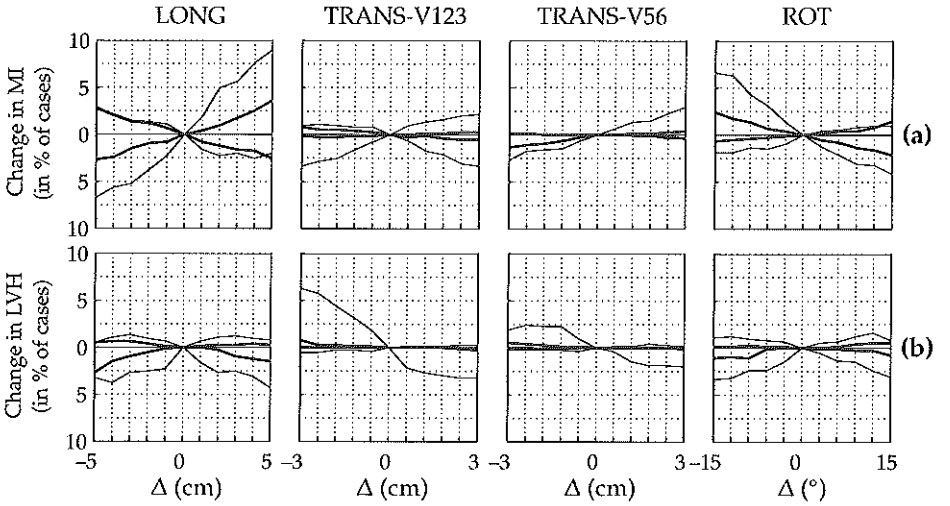


Figure 4.4. Percentages of the changes in classification for (a) myocardial infarction (MI) and (b) left ventricular hypertrophy (LVH) as functions of displacement Δ. Thick lines denote large classification changes (+3 or +4 in the upper half of the plots, −3 or −4 in the lower half); thin lines denote small changes (+1 or +2 in the upper half, −1 or −2 in the lower half). A positive Δ means an upward shift in LONG, a leftward shift for V_2 and V_3 in TRANS-V123, a leftward shift in TRANS-V56, and a counterclockwise rotation in ROT.

for $\Delta < 0$ (a leftward shift of V_1), and negative ones for opposite Δ. The same effect, but smaller, was present in experiment TRANS-V56.

4.3.3 Comparison with Human Expert

The classifications of the computer and those of the human expert were compared and are shown in Table 4.2 on the following page . For the myocardial infarction category, the cardiologist did not change his classification in 19 out of 20 cases where the program had made no change, while he made a small change of +1 in the remaining case. In 13 out of 20 cases in which the computer made a large change, the cardiologist did not, while in the other 7 cases he agreed with the direction of the change. For the left ventricular hypertrophy category, the expert agreed with the computer in most cases in which the computer made no change (19 out of 20). In the remaining case he made a small change of −1. Of the 20 cases in which the computer made a large change, he made no change in 5 cases, a small one in 9, and a large one in 6. The direction of these changes was the same as the changes made by the computer with one

exception, where he made a small negative change of −1 (from probable LVH to possible LVH) while the computer made a positive change of +3 (normal to probable LVH). Overall, the cardiologist appeared less inclined to changes than the computer.

Table 4.2. Changes in MI and LVH classifications made by the human expert compared with those made by the computer program on two different sets of 40 ECG pairs.

Computer		Human Expert					
		−4/−3	−2/−1	0	+1/+2	+3/+4	Total
MI	−4/−3	2	1	7	0	0	10
	0	0	0	19	1	0	20
	+4/+3	0	0	6	2	2	10
LVH	−4/−3	4	4	2	0	0	10
	0	0	1	19	0	0	20
	+4/+3	0	1	3	4	2	10

MI, myocardial infarction; LVH, left ventricular hypertrophy.

4.3.4 Comparison with Temporal Variability

The findings of the present study were compared with reported year-to-year, day-to-day and beat-to-beat variations [2,6–8,23–26]. Table 4.3 on the next page shows the variability values over different time intervals for those studies that addressed the same measurements as in our study and used similar variability measures. Note that in some studies measurements were obtained by hand, which introduces observer variability. In others, unmarked electrode positions were used. When different values for each lead were reported, the maximum over all leads was taken. Similar values were computed for the experiments in our study for two values of Δ; they are shown in Table 4.4 on page 72. The Δ values that were chosen represent a small positive and a large positive shift. Similar results are obtained for negative shifts, as the distribution of absolute measurement changes tended to be symmetric. The distribution of absolute changes as a function of Δ was approximately uniform for all experiments and all measurements. Values for Δs not listed can thus easily be approximated. The findings of different studies should be compared with caution, as Q measurements are presented for all leads together and changes in VCG measurements may have been compared to those in ECG measurements.

The electrode position change corresponding with the temporal QRS duration

Table 4.3. The 96% range of absolute changes from the mean value of QRS duration and Q duration and amplitude and the 96% range of the absolute paired differences of the Sokolow index reported in different variability studies.

Study	Time Scale*	Marked¶	Type	Measurement (96% range)			
				QRS Dur (ms)	Q Dur (ms)	Q Amp (μV)	Sokolow Index (μV)
Michaelis et al. [2]	Y	No	ECG	6	—	—	—
McManus et al. [23]	Y	No	VCG	5	8	140	—
Willems et al. [24]	D	Yes	VCG	5	5.5	78	—
Farb et al. [26]	D/B	No/NA	ECG	—	—	—	827/109
Fischmann et al. [6]	B	NA	VCG	19†	—	50‡	—
Borovsky et al. [7]	B	NA	VCG	11†§	—	—	—

* Y, year-to-year variability; D, day-to-day variability; B, beat-to-beat variability.

† Manual measurement.

‡ Mean + 2 SD.

§ Root-mean-square difference instead of absolute difference.

¶ Marked electrode positions. NA, not applicable.

Table 4.4. The 96% range of the absolute changes from the mean value of QRS duration and Q duration and amplitude and the 96% range of the absolute paired differences of the Sokolow index resulting from experiments for two displacements.

Experiment		Measurement (96% range)			
name	Δ_1/Δ_2	QRS dur (ms)	Q dur (ms)	Q amp (μV)	Sokolow (μV)
LONG	10 /40 mm	3 / 6	8 / 13	74 /260	269 /1256
TRANS-V123	6 / 24 mm	2 / 4	2 / 3	37 /141	107 / 404
TRANS-V56	6 / 24 mm	3 / 3	2 / 8	5 / 26	1 / 241
ROT	3 / 12°	3 / 5	8 / 12	22 / 78	145 / 902

The first value in cols. 2–6 corresponds to Δ_1, the second to Δ_2.

variability (which is assumed to be 5 ms) indicates that the influence of electrode position changes on QRS duration is small compared with the temporal variability. When temporal Q duration variability is taken to be 6 ms, it follows from Table 4.4 that even vertical electrode position changes as small as 10 mm cause larger variability. Horizontal shifts have a smaller effect. For Q amplitude, a temporal variability of about 60 μV may be assumed. Smaller changes are produced by transverse displacements of V_5 and V_6. Comparable changes are produced by transverse displacements of V_1 to V_3 and rotations, and larger changes are produced by longitudinal displacements. Table 4.3 on the preceding page shows that for the Sokolow index, the day-to-day variability is much larger than the beat-to-beat variability and is only surpassed by large vertical electrode shifts (a Δ larger than about 27 mm in experiment LONG). The relatively small beat-to-beat variability is exceeded for a vertical shift of 5 mm. Only small horizontal shifts cause lower variability.

4.4 Discussion

4.4.1 Measurements

For the QRS duration, the Q duration and the Q amplitude, there is no systematic effect on the population as a whole, since the medians of the distributions of the differences remained almost zero for all experiments. However, in individual cases, large changes can occur. These extreme changes in Q measurements in individual cases pertained to QS complexes. Even though the 2 and 98 percentiles were large, the overall effect on these measurements is small, since both

the 25 and the 75 percentiles were close to zero. As for the Sokolow index, the longitudinal position of leads V_1 and V_5 appears optimal, in the sense that the median decreases when the electrodes are moved either downward or upward. The 2 and 98 percentiles are very large, however. On top of the day-to-day variability of this index in left ventricular hypertrophy classification [26], this indicates the dependence of the index on chest electrode position.

4.4.2 Diagnostic Classifications

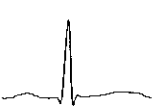
Diagnostic classifications appear sensitive to electrode displacement. For extreme electrode shifts a few percent of the cases show large changes in classification. Both diagnostic categories appear most sensitive to longitudinal electrode displacements. In practice, a displacement of this type is likely to be of the magnitude of a full intercostal space, because of an error in selecting the correct intercostal space. Moreover, displacements are often a combination of translational and longitudinal shifts, introducing even larger displacements [1].

With respect to the changes in diagnostic classification, the myocardial infarction classifications are more sensitive to displacements than the left ventricular hypertrophy classifications. The median of the difference distribution for myocardial infarction was zero for all experiments, however, which indicated no systematic change when electrodes were shifted. The left ventricular hypertrophy classifications assumed lower probability for longitudinal shifts in either direction. In conformity with the decrease in amplitude measurements, the standard locations indeed appear optimal in the sense that they show highest certainty for left ventricular hypertrophy.

Since in population studies automatic classification is often used, these effects on classification may have a considerable effect on the estimation of prevalence and incidence of pathologies in such studies.

4.4.3 Comparison with Findings of Human Expert

The cardiologist let himself be less influenced by the effects of electrode shifts than the computer. That is, human observation was more resistant to insignificant signal changes than the computer. In the cases in which the program did not generate changes in classification, the cardiologist either did not change his classification or made a small change only. In the cases in which both the computer and the cardiologist changed their classifications, the changes were made in the same direction for all but one case. The percentage of cases in which the computer changed its diagnostic classification was small and it was smaller still



for the cardiologist. Thus, diagnostic classification by a human expert was only slightly influenced by electrode displacements.

4.4.4 Comparison with Temporal Variability

The effect of electrode position changes on QRS duration was small as compared with temporal variability, even for large shifts. Conversely, for Q duration and amplitude, small vertical shifts exceeded the effects of reported temporal variability. Horizontal shifts of leads V_1 to V_3 had some effect on Q amplitude, probably because of the high occurrence of QS complexes in these leads. The day-to-day variability of the Sokolow index was small compared to the variability caused by most electrode position changes. It should be noted also that in these studies probably no variability due to different centers or different technicians was present. In practice, this may result in a systematically higher variability than the values reported in these studies.

Electrode placement variations appear to have some effect on automatic ECG classification, but classification by a human expert is less affected by electrode displacements. Variations in electrode positioning, especially in the vertical direction, are significant in a few percent of the cases. Therefore, selecting the correct electrode positions remains mandatory.

4.5 References

[1] Wenger W, Kligfield P. Variability of precordial electrodes placement during routine electrocardiography. J Electrocardiol 1996;29(4):179–84.

[2] Michaelis J, Lippold R, Gluck E, Schindler W, Scheidt E. Automated serial ECG analysis within an epidemiological study. J Electrocardiol 1987;20 Suppl:34–6.

[3] Gau GT, Smith RE. The effect of electrode position on a modified Frank electrovectorcardiographic lead system. Mayo Clin Proc 1971;46(8):536–43.

[4] Berson AS, Haisty WK, Pipberger HV. Electrode position effects on Frank lead electrocardiograms. Am Heart J 1978;95:463–73.

[5] Burger HC, Van Brummelen AGW, Van Herpen G. Correlation between subjective and objective measures of correspondence between different systems of vectorcardiography. Am Heart J 1964;67:512–5.

[6] Fischmann E, Cosma J, Pipberger HV. Beat to beat and observer variation of the electrocardiogram. Am Heart J 1968;75:465–73.

[7] Borovsky D, Zywietz C. Accuracy and beat-to-beat variation in ECG computer measurements. Adv Cardiol 1976;16:176–81.

[8] Kors JA, Van Herpen G, Van Bemmel JH. Variability in ECG computer interpretations: analysis of individual complexes vs analysis of a representative complex. J Electrocardiol 1992;25:263–71.

[9] Peberdy MA, Ornato JP. Recognition of electrocardiographic lead misplacements. Am J Emerg Med 1993;11(4):403–5.

[10] Herman MV, Ingram DA, Levy JA, Cook JR, Athans RJ. Variability of electrocardiographic precordial lead placement: a method to improve accuracy and reliability. Clin Cardiol 1991;14:469–76.

[11] Hill NE, Goodman JS. Importance of accurate placement of precordial leads in the 12-lead electrocardiogram. Heart Lung 1987;16:561–6.

[12] Pahlm O, Haisty WK, Edenbrandt L, Wagner NB, Sevilla DC, Selvester RH, et al. Evaluation of changes in standard electrocardiographic QRS waveforms recorded from activity-compatible proximal limb lead positions. Am J Cardiol 1992;69:253–7.

[13] Van Bemmel JH, Kors JA, Van Herpen G. Methodology of the modular ECG analysis system MEANS. Methods Inf Med 1990;29(4):346–53.

[14] Willems JL, Arnaud P, Van Bemmel JH, Degani R, Macfarlane PW, Zywietz C. Common standards for quantitative electrocardiography: goals and main results. Methods Inf Med 1990;29:263–71.

[15] Willems JL, Arnaud P, Van Bemmel JH, Bourdillon PJ, Brohet C, Dalla Volta S, et al. Assessment of the performance of electrocardiographic computer programs with the use of a reference data base. Circulation 1985;71:523–34.

[16] Willems JL, Abreu-Lima C, Arnaud P, Van Bemmel JH, Brohet C, Degani R, et al. The diagnostic performance of computer programs for the interpretation of electrocardiograms. N Engl J Med 1991;325:1767–73.

[17] Zywietz C, Willems JL, Arnaud P, Van Bemmel JH, Degani R, Macfarlane PW. Stability of computer ECG amplitude measurements in the presence of noise. Comput Biomed Res 1990;23:10–31.

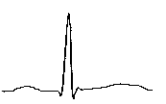
[18] Montague TJ, Smith ER, Cameron DA, Rautaharju PM, Klassen GA, Felmington CS, et al. Isointegral analysis of body surface maps: surface distribution and temporal variability in normal subjects. Circulation 1981;63(5):1166–72.

[19] Kornreich F, Montague TJ, Rautaharju PM. Body surface potential mapping of ST segment changes in acute myocardial infarction. Circulation 1993;87:773–82.

[20] Kornreich F, Rautaharju PM, Warren JW, Montague TJ, Horacek BM. Identification of best electrocardiographic leads for diagnosing myocardial infarction by statistical analysis of body surface potential maps. Am J Cardiol 1985;56:852–6.

[21] Schijvenaars RJA, Kors JA, Van Herpen G, Kornreich F, Van Bemmel JH. Interpolation of body surface potential maps. J Electrocardiol 1995;28 Suppl:104–9.

[22] Berson AS, Gamboa R, Naval A, Pipberger HV. Differences caused by changes in electrode positions in two corrected orthogonal lead systems. J Electrocardiol 1970; 3:29–42.



[23] McManus C, Doyle J, Pipberger HV. Year-to-year variation of the normal orthogonal electrocardiogram and vectorcardiogram in 243 white males. J Electrocardiol 1984;17:107–14.

[24] Willems JL, Poblete PF, Pipberger HV. Day-to-day variation of the normal orthogonal electrocardiogram and vectorcardiogram. Circulation 1972;45:1057–64.

[25] Michaels L, Cadoret RJ. Day-to-day variability in the normal electrocardiogram. British Heart Journal 1967;29:913–9.

[26] Farb A, Devereux RB, Kligfield P. Day-to-day variability of voltage measurements used in electrocardiographic criteria for left ventricular hypertrophy. J Am Coll Cardiol 1990;15(3):618–23.

6 Use of the Standard 12-lead Electrocardiogram to Simulate Electrode Displacements

R.J.A. Schijvenaars, J.A. Kors, G. van Herpen, J.H. van Bemmel

Dept. of Medical Informatics, Erasmus University Rotterdam, The Netherlands

Journal of Electrocardiology 1996;29 Suppl:5–9

Abstract

Placement of the precordial electrodes for recording a 12-lead electrocardiogram (ECG) is subject to variation. Previous research has shown that displacements, especially in the longitudinal direction, can lead to changes in diagnosis. In practice, both the displacement and the effects of displacement on an individual ECG are unknown. To assess this effect for a given ECG, the authors developed a method to simulate ECGs at different displacements using only the recorded ECG.

The material consisted of 746 body surface potential maps (BSPMs) containing 232 cases without abnormalities, 277 with myocardial infarction (MI), and 237 with left ventricular hypertrophy. By interpolating BSPMs, ECGs from closely spaced electrode positions could be derived. Taking electrode positioning errors that may be encountered in practice, 40 ECGs at different electrode displacements (displaced ECGs) for each BSPM were derived. Using half of the BSPMs, for each displacement, a transformation matrix that transforms the ECG at the standard 12-lead electrode positions (standard ECG) to the displaced ECG was determined.

Using the other half of the BSPMs, each displaced ECG was compared with the ECG yielded by the corresponding transformation matrix (transformed ECG). For each comparison, the differences were assessed between the two sets of ECG signals and between the diagnostic computer classifications of the two sets. Signal differences were expressed as mean absolute differences over the QRS. Computer interpretation of MI and left ventricular hypertrophy was graded in five levels of certainty (no, consider, possible, probable, definite). For instance, for the largest longitudinal displacement studied of about one intercostal space, the 96th percentile mean absolute difference over the test set was 204 μV. The percentage of cases showing a change in MI classification of more than two certainty levels was 2.7% for this displacement. When comparing the standard ECG with the displaced ECG, these figures were 434 μV and 8.3%, respectively.

It is concluded that ECGs from displaced electrodes can be well simulated by transforming the standard ECG, both for ECG signal and for diagnostic classifications.

Keywords: ECG, body surface potential map, displacement.

5.1 Introduction

One of the causes of variability in the electrocardiogram (ECG) is variation in the position of the precordial electrodes relative to the heart. Previous research showed that displacements, especially in the longitudinal direction, can lead to changes in the ECG signal [1–4] and in diagnosis [4], depending on the ECG at hand. The effect of positional variation can be assessed by comparing ECGs recorded simultaneously at different electrode positions. In practice, however, only one ECG is available and is assumed to have been derived from the standard electrode positions. To assess positioning effects for a single individual ECG, we developed a method to simulate ECGs at different electrode displacements using only the ECG at the standard electrode positions. When this effect can be assessed for a particular ECG, it might be used as a measure of the interpretation stability; thus, it could be taken into account when comparing the interpretations of ECGs of the same patient taken at different times.

5.2 Methods

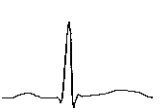
5.2.1 Material

A set of 746 Body Surface Potential Maps (BSPMs) was used. The recording procedure has been previously described [5,6]. The grid consisted of three limb electrodes and 117 torso sites, 81 anterior and 36 posterior (Fig. 5.1 on the next page). The complete set includes 232 (31%) cases without abnormalities, 277 (37%) with myocardial infarction (MI) and 237 (32%) with left ventricular hypertrophy (LVH). These diagnoses have been established by ECG-independent methods [7]. Each BSPM comprised one representative ECG complex per lead, obtained by coherent averaging. All maps were checked for consistency and quality of the signals. The set of BSPMs was divided into a learning set and a test set, each containing 373 cases. Each diagnostic category was equally divided over the two subsets.

5.2.2 Technique

Deriving ECGs from BSPMs

Two types of ECGs were derived from the BSPMs: ECGs at standard electrode locations (standard ECGs) and ECGs at nonstandard electrode locations representing positioning errors (displaced ECGs). Our goal was to investigate



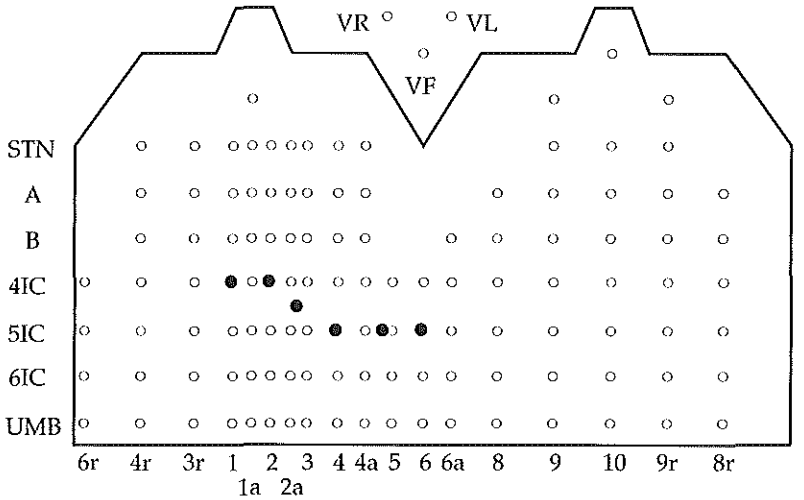


Figure 5.1. Schematic unfolded thorax showing the body surface potential map electrode positions (open circles). The left part of the image represents the chest, the right part represents the back of the thorax. The electrode rows were placed 5 cm apart in the longitudinal direction, with the fourth row from the bottom located in the fourth intercostal space (4IC). The top row is at the level of the sternal notch (STN) and the bottom row is near the umbilicus (UMB). The placement in the left-right direction was done using anatomic landmarks [6]. The black dots indicate the standard lead V_1 to V_6 electrode positions.

whether it was possible to simulate a displaced ECG by applying a linear transformation to the standard ECG, resulting in a transformed ECG.

The BSPMs were interpolated to construct the ECG signals for any desired electrode position. Previous research showed that ECGs from any location can be accurately approximated using Fourier and cubic spline interpolation [8]. The ECGs from particular electrode configurations were derived from the interpolated BSPMs by taking the signals corresponding with the electrode locations.

Simulation of any desired target chest lead V_k of a displaced ECG is obtained from the eight leads of the standard ECG as $V_k = a_{1k} \times \mathrm{I} + a_{2k} \times \mathrm{II} + a_{3k} \times \mathrm{V_1} + \ldots + a_{8k} \times \mathrm{V_6}$. Applying such a linear combination for each of the six chest leads of a standard ECG yields a transformed ECG. The coefficients a_{ik} were computed by minimizing the sum of the squared differences between the time-normalized QRS of the displaced ECG and its corresponding transformed ECG for all ECGs in the learning population. This method has been applied before for reconstruction of the vectorcardiogram (VCG) from the ECG [9], or vice

versa [10], and for reconstruction between VCGs derived from different VCG lead systems [11]. For each chest electrode configuration a different matrix must be calculated.

Electrode displacement experiments

The following procedure was regarded as the standard for chest electrode placement. First, leads V_1 and V_2 are placed on opposite sides close to the sternum in the fourth intercostal space, V_4 is placed on the left midclavicular line in the fifth intercostal space and V_6 is placed on the left midaxillary line at the same level as V_4. Lead V_3 is placed halfway between V_2 and V_4, and V_5 is placed halfway between V_4 and V_6. This method of electrode placement may lead to some general types of displacement in the clinical routine. Four of them were envisaged and simulated by experiments.

The first experiment (LONG) simulated an error, upward or downward, in determining the levels of the fourth and fifth intercostal spaces. This is equivalent to the reverse shift of the heart position with respect to the thoracic cage. The experiment is illustrated in Fig. 5.2(a) on the following page. The variable Δ, which ranges from -5 cm (upward) to $+5$ cm (downward) in steps of 1 cm, indicates the distance over which the electrodes were moved. In Fig. 5.2(a), only the downward displacement is indicated.

Transversal displacements were thought to follow two possible patterns. If leads V_1 and V_2 are placed at opposite sides close to the sternum, they are likely to be in the same transversal plane. However, the distance at which V_1 and V_2 are placed on opposite sides of the sternum may vary. Thus, in the second experiment, transversal displacements were studied for the leads V_1–V_3, since shifting the V_2 position will result in shifting V_3 as well (over half the distance). This is illustrated in Fig. 5.2(b) on the next page (experiment TRANS-V123). The variable Δ ranges from -3 cm (V_1 rightward) to $+3$ cm (V_1 leftward) in steps of 6 mm.

The third experiment simulated an error in positioning lead V_6 in the transversal plane. In the longitudinal direction, leads V_5 and V_6 will remain in line, but a transversal displacement of V_6 will change the transversal position of V_5 by half this distance. This experiment is illustrated in Fig. 5.2(c) on the following page (experiment TRANS-V56). The variable Δ ranges from -3 cm (leftward) to $+3$ cm (rightward) in steps of 6 mm.

In addition to these longitudinal and transversal displacements, a rotational dislocation was also performed. Rotation of the set of precordial electrodes about a center at lead V_2 is more or less equivalent to the reverse rotation of the

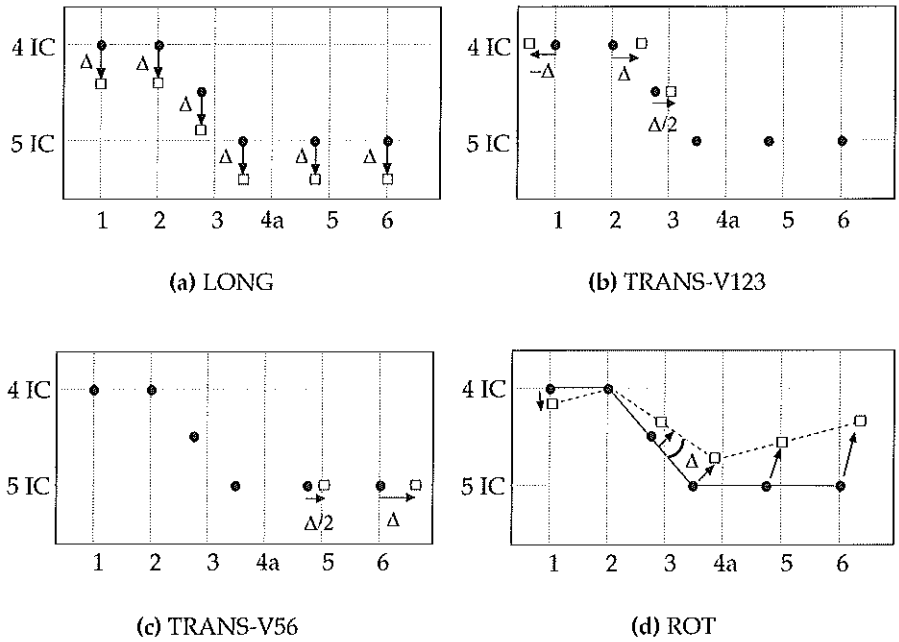


Figure 5.2. Displacements of the individual chest electrodes for the four experiments (a) LONG, (b) TRANS-V123, (c) TRANS-V56, and (d) ROT. Δ indicates the magnitude of the displacement, the black dots indicate the standard precordial electrode positions, and the open squares indicate the positions after displacement Δ. The numbers along the ordinates and abscissas correspond with those in Fig. 5.1 on page 80. IC, intercostal space.

heart with respect to a fixed set of electrodes. The topography of a longitudinally oriented heart, such as in tall and slender persons is simulated by rotating the electrodes counterclockwise so that a larger angle between the apex and the electrode level is obtained. A more upward tilting of the heart (e.g., in cases of pregnancy or obesity) is simulated by a clockwise rotation of the electrode positions. Figure 5.2(d) shows the displacements studied in this experiment. The variable Δ ranges from $+15°$ (counterclockwise rotation) to $-15°$ (clockwise rotation) in steps of 3°.

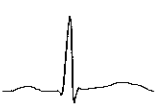
5.2.3 Performance assessment

The performance of the transformation matrices was assessed by comparing each displaced ECG (obtained by BSPM interpolation) in the test set with its corresponding transformed counterpart, obtained by transformation of the standard ECG. The comparison was done at the level of signal amplitudes and at the level of diagnostic interpretation. For signal differences, we only considered leads corresponding with displaced electrodes. As a measure of signal difference, the mean absolute amplitude difference (*MAD*) over the time-normalized QRS was computed, taken over all leads involved in the comparison. To show the relative improvement obtained by transformation, *MAD* values were also computed for the differences between the displaced and standard ECGs.

Comparison of diagnostic interpretations was done using the Modular ECG Analysis System, our ECG analysis program [12]. The interpretative statements of this program contain one of four qualifiers or levels of certainty: consider, possible, probable, or definite. The classification categories compared were MI and LVH. No distinction was made between infarct localizations and since the precordial leads are not used to diagnose inferior infarctions, these infarctions were not considered. For each category, the qualifiers were assigned values 1 (for consider) to 4 (for definite). A value of zero was assigned when the corresponding category was not mentioned. A difference in diagnostic interpretation was expressed by subtracting the qualifier points of the transformed ECG from those of the displaced ECG. For example, if a transformed ECG was interpreted as definite infarction, while the displaced ECG was interpreted as possible infarction and probable hypertrophy, the difference in diagnostic interpretation was coded as $-2(=2-4)$ in the MI category, and $+3(=3-0)$ in the LVH category. Only differences were considered, regardless of the levels of certainty from which they originated.

5.3 Results

For each of the four displacement experiments studied, 10 *MAD* values of the comparison of displaced ECG and corresponding transformed ECG were computed, one for each displacement magnitude. Figure 5.3 on the following page shows the 96th percentiles of these values over the complete test set as a function of electrode displacement in the four experiments, as well as the same values for the comparison of the displaced ECG and the standard ECG. For all displacements, in all experiments, the signal differences between the displaced ECG and the transformed ECG are about half the signal differences between the



displaced ECG and the standard ECG. The transformations therefore serve to approximate the standard ECG to the displaced ECG. Note that the *MAD* values of different experiments are difficult to compare since the average displacement of the electrodes varies between experiments: a displacement of 3 cm in experiment TRANS-V123 means a displacement of 3 cm for leads V_1 and V_2, but only 1.5 cm for lead V_3.

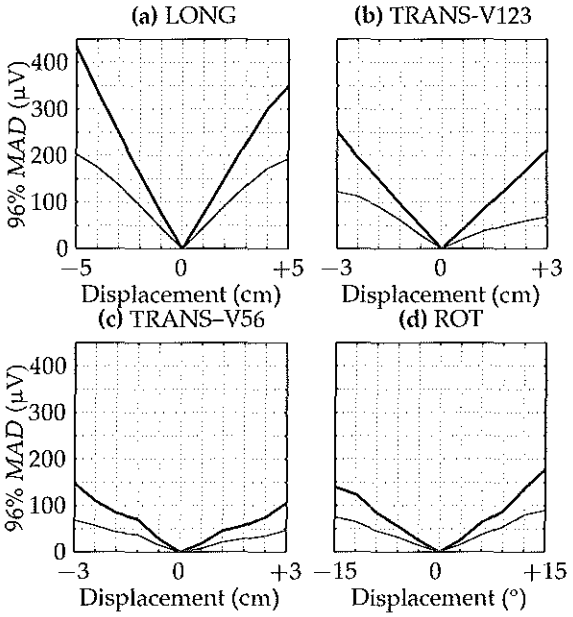


Figure 5.3. Signal differences for experiments (a) LONG, (b) TRANS-V123, (c) TRANS-V56, and (d) ROT. Shown are the 96th percentiles of the mean absolute difference (*MAD*) between the displaced and standard ECGs (thick lines) and between the displaced and transformed ECGs (thin lines), as a function of electrode displacement (compare Fig. 5.2 on page 82).

To test whether the performance of the transformation depends on the distribution of diagnostic categories in the learning population, transformation matrices were derived from learning subsets consisting of the BSPMs of one diagnostic category only. The differences in *MAD* values of transformations based on different learning sets, when tested on several homogeneous test subsets consisting of cases of one diagnostic category, were in the order of a few percent.

The comparison on the diagnostic interpretation level was done using the percentage of cases with an absolute difference in qualifier points of at least three. Figure 5.4 on the next page shows these percentages for the differences between

the transformed and displaced ECGs (thin lines) and the standard and displaced ECGs (thick lines) as a function of electrode displacement. Results are shown only for experiment LONG, which produced the largest absolute differences; the other experiments showed less pronounced differences. For the MI category (Fig. 5.4a), the percentage of large differences in diagnostic interpretation between the transformed and displaced ECGs is roughly half the difference between the standard and displaced ECGs. For LVH (Fig. 5.4b), the differences are comparable for small displacements, but for larger displacements the resemblance of the displaced ECG with the transformed ECG is much better than with the standard ECG.

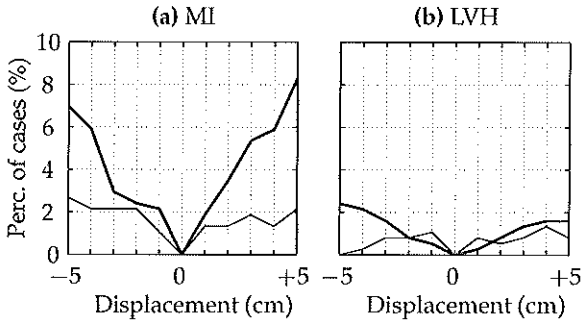


Figure 5.4. Percentage of cases with large differences in diagnostic interpretation for the (a) myocardial infarction (MI) and (b) left ventricular hypertrophy (LVH) categories. Thin lines indicate the difference between the displaced ECG and transformed ECG as a function of electrode displacement in experiment LONG. Thick lines denote the difference between the displaced ECG and the standard ECG.

5.4 Discussion and conclusion

We derived transformation matrices in order to simulate electrode positioning errors that may be encountered in practice. These transformations were evaluated by comparing the displaced and transformed ECGs derived from the standard ECG.

The results show that the transformed ECG resembled the displaced ECG substantially better than the standard ECG, both in signal amplitudes and in diagnostic classification. The results are based on an evaluation of three diagnostic categories (normal, MI, and LVH). No loss of performance was observed when transformation matrices, learned on the basis of one category, were sub-

sequently applied to others. Thus, we are confident that the transformation matrices are generally applicable and do not depend on the category.

In this study, we have demonstrated that it is possible to simulate positioning errors of the chest electrodes. Previously, it has been shown that large limb lead displacements (e.g., shoulder v.s. wrist) can have effects on the ECG as well [13]. We did not simulate such displacements because we considered them to be the result of different lead placement procedures. Our goal was to simulate electrode displacements caused by positioning errors — not to simulate other lead placement procedures.

When the transformations are applied to a given ECG, it is implicitly assumed that the ECG was recorded at the standard electrode positions. If this is not the case, the transformed ECGs cover a somewhat different set of electrode displacements. Additional research is needed to study to what extent the performance of the transformation depends on the electrode positions of the ECG to be transformed.

In practice, our method may be applied to recorded ECGs to obtain an estimate of the stability of computer interpretation. If the computer classification remains constant for all transformed ECGs, it can be regarded as more trustworthy than if the classification varies significantly between transformations. Especially in automated serial analysis, this enables a more weighted comparison between computer classification of the various serial ECGs.

We conclude that ECGs from displaced electrodes can be well simulated by transforming the standard ECG, both for the ECG signal and for diagnostic classifications.

5.5 References

[1] Peberdy MA, Ornato JP. Recognition of electrocardiographic lead misplacements. Am J Emerg Med 1993;11(4):403–5.

[2] Herman MV, Ingram DA, Levy JA, Cook JR, Athans RJ. Variability of electrocardiographic precordial lead placement: a method to improve accuracy and reliability. Clin Cardiol 1991;14:469–76.

[3] Hill NE, Goodman JS. Importance of accurate placement of precordial leads in the 12-lead electrocardiogram. Heart Lung 1987;16:561–6.

[4] Schijvenaars RJA, Kors JA, Van Herpen G, Kornreich F, Van Bemmel JH. The effect of electrode position changes on ECG interpretation by computer. J Electrocardiol 1997;30(3):247–56.

[5] Kornreich F, Montague TJ, Rautaharju PM. Body surface potential mapping of ST segment changes in acute myocardial infarction. Circulation 1993;87:773–82.

[6] Montague TJ, Smith ER, Cameron DA, Rautaharju PM, Klassen GA, Felmington CS, et al. Isointegral analysis of body surface maps: surface distribution and temporal variability in normal subjects. Circulation 1981;63(5):1166–72.

[7] Kornreich F, Rautaharju PM, Warren JW, Montague TJ, Horacek BM. Identification of best electrocardiographic leads for diagnosing myocardial infarction by statistical analysis of body surface potential maps. Am J Cardiol 1985;56:852–6.

[8] Schijvenaars RJA, Kors JA, Van Herpen G, Kornreich F, Van Bemmel JH. Interpolation of body surface potential maps. J Electrocardiol 1995;28 Suppl:104–9.

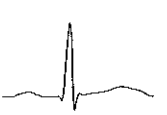
[9] Kors JA, Van Herpen G, Sittig AC, Van Bemmel JH. Reconstruction of the Frank vectorcardiogram from standard electrocardiographic leads: diagnostic comparison of different methods. Eur Heart J 1990;11:1083–92.

[10] Dower GE, Machado HB, Osborne JA. On deriving the electrocardiogram from vectorcardiographic leads. Clin Cardiol 1980;3:87–95.

[11] Burger HC, Van Brummelen AGW, Van Herpen G. Compromise in vectorcardiography displacement of electrodes as a means of adapting one lead system to another. Am Heart J 1961;62:398–400.

[12] Van Bemmel JH, Kors JA, Van Herpen G. Methodology of the modular ECG analysis system MEANS. Methods Inf Med 1990;29(4):346–53.

[13] Pahlm O, Haisty WK, Edenbrandt L, Wagner NB, Sevilla DC, Selvester RH, et al. Evaluation of changes in standard electrocardiographic QRS waveforms recorded from activity-compatible proximal limb lead positions. Am J Cardiol 1992;69:253–7.



Use of Electrocardiographic Variability to Improve the Reliability of Computer Assisted Diagnostic Interpretation

R.J.A. Schijvenaars*, J.A. Kors*, G. van Herpen*,
N.M. van Hemel†, J.F. May‡, J.H. van Bemmel*

* Dept. of Medical Informatics, Erasmus University, Rotterdam, The Netherlands
† Dept. of Cardiology, St. Antonius Hospital, Nieuwegein, The Netherlands
‡ Dept. of Cardiology, Groningen University Hospital, Groningen, The Netherlands

Submitted for publication

Abstract

The use of ECG variability for the improvement of automatic ECG classification was investigated in 470 ECGs recorded in a clinical setting and 7,437 normal ECGs from an epidemiological survey. Two sources of ECG variability were studied: technical (chest electrode position) and biological (beat-to-beat variability). For the assessment of technical variability, ECGs with displaced chest electrodes were derived from the recorded ECG by a linear combination of the recorded leads and classified automatically. Biological variability was assessed by automatically classifying in the recorded ECG all representative individual beats in addition to the averaged beat. Sensitivity and specificity for myocardial infarction (MI) were determined for an entire set of ECGs as well as for subsets having lower variability. For 272 clinical ECGs the discharge letter was studied by a cardiologist ('clinical' reference). The other 198 ECGs were reviewed by three cardiologists ('cardiologists' reference). For the 'clinical' reference, the MI specificity increased from 96% for the entire set to 100% for the set with minimal variability. The sensitivity remained constant at 73%. For the 'cardiologists' reference, MI sensitivity and specificity increased from 83% to 97% and from 89% to 100%, respectively. The specificity in the set of normal ECGs increased from 95.7% to 97%. Classification variability is useful in identifying subgroups of ECGs that can more reliably be interpreted by an ECG interpretation computer program and can be a valuable aid in selecting ECGs for overreading in, for instance, population surveys.

Keywords: electrocardiography, computers, ECG variability, electrode positions

6.1 Introduction

Various sources of variability may affect the shape of the electrocardiogram (ECG). Non-pathological, 'circumstantial' variability can be related to technical sources (e.g., electrode positioning, mains interference or patient posture) or to biological sources (non-cardiac muscle activity, respiration). In the past, much effort has been spent to assess [1–3] or to minimize [4–7] ECG variability resulting from these two sources.

Circumstantial ECG variability is always considered unwanted and is either ignored or minimized, e.g., by beat averaging. However, we hypothesized that ECG variability could also serve to improve the reliability of interpretation. We investigated whether ECG variability, expressed as a stability index, can be exploited to distinguish ECG interpretations that are more reliable, in terms of sensitivity and specificity, from less reliable interpretations. Such distinction could be instrumental in population surveys and large clinical trials to manage overreading costs or to reduce errors in automated ECG classification. Quantitative assessment of variability might also be useful for serial ECG analysis, in that serial changes between ECGs with large variation may be considered less trustworthy.

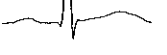
We tested our approach for both types of ECG variability. For variability resulting from technical sources (technical variability), the variability of ECGs simulated for displaced chest electrode positions [8] was used; for variability from biological sources (biologic variability), we took the beat-to-beat variability in ECG recordings.

6.2 Methods

6.2.1 Material

ECG classification performance was assessed for the detection of myocardial infarction in two databases of ECGs recorded in a clinical setting. For one set of ECGs the 'clinical truth' was taken as the reference, for another set the interpretations of a group of cardiologists was obtained. To determine the usefulness of the technique in an epidemiological setting as well, a third database of ECGs from a population survey was used to assess the accuracy for normal ECGs.

The first database contains 272 ECGs (patients' mean age 57.4 years; 34.6% female) selected from a large collection of ECGs collected during clinical routine in the cardiology department of the university hospital in Rotterdam, The Netherlands. A reference classification was determined by a cardiologist



who interpreted the discharge letter. Prevalences according to this reference classification were 69.1% for myocardial infarction (MI) and 30.9% for non-MI (normal and other abnormalities). This set is referred to as the 'clinical' database.

The second database of 198 ECGs (patients' mean age 56.0 years; 38.9% female) was collected during clinical routine in cardiology departments in five European hospitals and was classified by three cardiologists. The average kappa value of their classifications was 0.72, indicating substantial agreement [9,10]. A reference classification was computed by combining the cardiologists' classifications as will be explained lateron. This set is referred to as the 'cardiologists' database. Prevalences according to the reference classification were 27.3% for MI, and 72.7% normal or other abnormalities.

The third database used to study the performance on the 'normal' diagnostic category consists of ECGs recorded in a cohort study among apparently healthy individuals based on analysis of clinical examinations, lab tests and patient questionnaires [11]. The 7,437 ECGs were assumed to be without pathology (mean age individuals 41.2 years; 30.7% female). This database is referred to as the 'normal' database.

Two types of reference classifications can be used to assess the performance of an ECG analysis program: the 'clinical truth', determined by clinical evidence such as echocardiographic findings, cardiac catheterization or enzyme changes, or the 'cardiologist's ECG interpretation' using only the ECG. In the latter case one could view the ECG as a laboratory test that should be performed without bias from additional, ECG-independent information. Both reference types were taken into account in assessing classification performance.

6.2.2 Electrode displacement simulation

To assess the variability due to electrode positioning, ECGs were simulated to mimic recordings as would have been obtained after upward and downward displacement of all six chest electrodes in five steps of 1 cm (such as may occur when the intercostal spaces are incorrectly determined). In a previous study [12], this type of displacement proved to have the largest effect on ECG interpretation among several other possible chest electrode position variations. Each precordial displaced ECG was constructed through a linear combination of leads I, II, and the six precordial leads of the original ECG. The coefficients for these linear combinations were not patient-specific. They had been derived previously with the help of body surface potential maps (BSPMs), through a least- squares fit between the signals from the standard ECG positions in the

BSPMs and those from each of the five upward and five downward electrode locations. The simulated ECGs were shown to mimic ECGs from really displaced electrode positions in good approximation [8]. In the present experiments for each original ECG 10 simulated displaced ECGs were obtained.

6.2.3 Electrocardiographic processing

The 12 leads of each ECG were recorded simultaneously at a sampling rate of 250 samples per second for 10 seconds. Processing was done by our Modular ECG Analysis System (MEANS) [13], of which the performance has been extensively tested, e.g., in the study 'Common Standards for Quantitative Electrocardiography' (CSE) in which the performance of a number of ECG analysis programs was assessed [14].

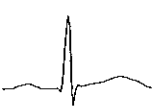
MEANS can produce one averaged representative beat for each lead for further analysis, or can analyze each representative beat separately. For the computation of the displaced ECGs, averaged beats were used. For the beat-to-beat variability experiments, the non-displaced, original ECG was used. The diagnostic categories studied were normal (NOR) in the 'normal' database and myocardial infarction (MI) in the 'clinical' and 'cardiologists' databases. For each category, MEANS could assign one of three qualifiers: 'possible', 'probable', or 'definite', coded as 1, 2, or 3, respectively. When a category was absent, a value of zero was coded. Thus, a set of qualifier codes is obtained for each diagnostic category: one for the original representative averaged beat, one for each averaged beat of the 10 simulated displaced ECGs, and one for each individual representative beat of the original ECG.

The coding procedure was also used to derive a reference classification for each of the 198 original ECGs in the 'cardiologists' database: the three cardiologist's interpretations were coded with qualifier values 0, 1, 2 and 3 for each diagnostic category and the rounded average of the 3 values was then taken as the reference classification. For example, the combination of 'definite MI', 'probable MI' and 'possible MI' would result in $(3+2+1)/3 = 2$ (probable MI). For the reference classifications of the ECGs in the 'clinical' database only codes 0 and 3 were used.

6.2.4 Stability index

A stability index I was defined as:

$$I = 1 - \frac{1}{3n} \sum_{i=1}^{n} |x_i - x_0| \tag{6.1}$$



to quantify the variability in the set of classifications. Variable x_0 denotes the qualifier of the original classification, x_i the qualifier of classification i, and $3n$ the normalization factor to make the index independent of the number of observations (e.g., the number of representative beats in a recording may vary). This index is simple to compute and to understand. Maximum stability (i.e., all qualifiers are equal to the one of the original ECG) gives a stability index one, minimum stability (i.e., the original classification is an outlier compared to all other classifications) an index zero. Two different stability indices resulted for each diagnostic category: one based on electrode position variability and one based on beat-to-beat variability. We also studied a combination index by taking the minimum of both indices, expressing variability in either beat-to-beat or electrode position changes.

6.2.5 Usage in daily routine

As pointed out by de Bruyne et al. [15], ECG interpretation by computer can be very helpful in population-based research. It can decrease the workload of research physicians or cardiologists considerably, depending on the strategy chosen for ECG interpretation. To screen for myocardial infarction, she recommends to use only a computer program or to let a cardiologist verify all ECGs the computer program classified as abnormal. In the latter case, improved performance is achieved at the expense of higher workload. We will show the advantage of using the stability index to balance workload and performance.

6.3 Results

6.3.1 Diagnostic performance

To assess the performance of the MEANS program, sensitivity and specificity were computed. Qualifier codes 2 ('probable') and 3 ('definite') were regarded as positive classifications, the other codes as negative ones. Figure 6.1 on the facing page shows the sensitivity and specificity of MEANS for ECGs in the 'clinical' database that have a stability index larger than or equal to the values indicated on the abscissa. Thus, performance estimates at stability index 0 apply to all ECGs in the database, and performance estimates at stability index 1 apply to the most stable ECGs only. MI sensitivities remained constant at about 73% for varying stability indices. However, MI specificity increased from 96.4% to 98.6% for electrode position (N=196; Fig. 6.1a on page 95) and beat-to-beat (N=186; Fig. 6.1b) and to 100% for the combined index (N=158; Fig. 6.1c).

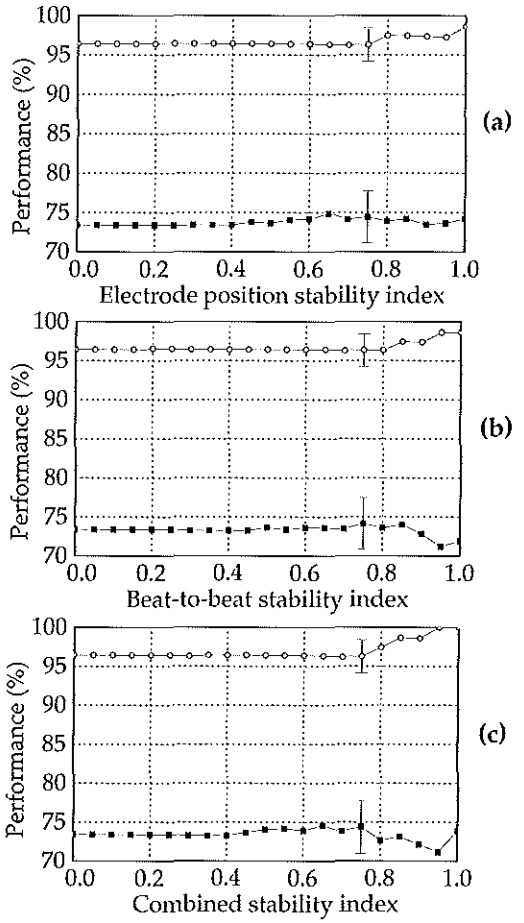


Figure 6.1. Classification performance on subgroups of ECGs from the 'clinical' database having a stability index larger than a particular value. Performance is expressed by sensitivity (solid squares) and specificity (open circles) for myocardial infarction (MI) and is plotted for the electrode position stability index (a), for the beat-to-beat stability index (b) and for the combined stability index (c). The vertical bars indicate the standard error. Note that the subgroup grows smaller when going from a stability index zero to an index one: at index zero all ECGs are considered, at index one only the most stable ones.

Figure 6.2 on the next page shows the sensitivities and specificities for myocardial infarction on the 'cardiologists' database. Both sensitivity and specificity increase with increasing stability. Sensitivity and specificity on the whole set of

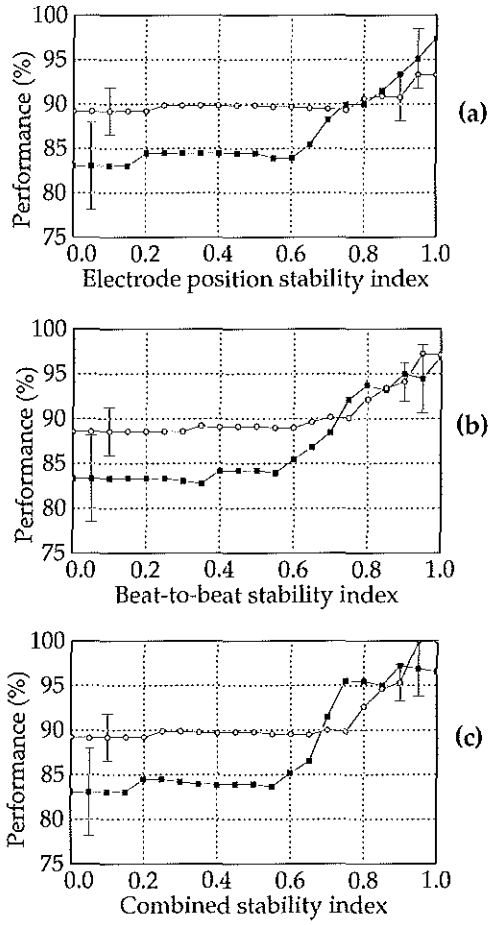


Figure 6.2. Classification performance on subgroups of ECGs from the 'cardiologists' database having a stability index larger than a particular value. Performance is expressed by sensitivity (solid squares) and specificity (open circles) for myocardial infarction (MI) and is plotted for the electrode position stability index (a), for the beat-to-beat stability index (b) and for the combined stability index (c). The vertical bars indicate the standard error.

198 ECGs were 83% and 89%, respectively. The sensitivity increased to 97% for all three indices, while the specificity increased to 93%, 97% and 100% for maximal electrode position index (N=142), beat-to-beat index (N=134), and combined index (N=114), respectively. Owing to the rather small number of ECGs in this database (N=195), standard errors of the estimates are relatively large: about 5% for sensitivity and 2% for specificity.

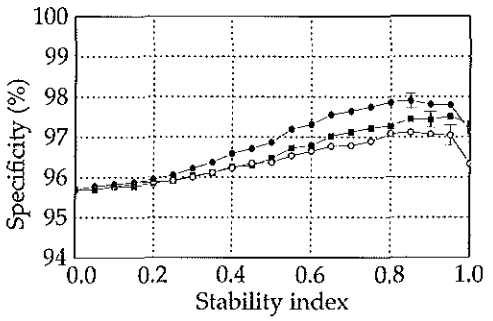


Figure 6.3. Classification performance on subgroups of ECGs from the 'normal' database having a stability index larger than a particular value. Performance is expressed by specificity. Since the 'normal' database contains no abnormal ECGs, sensitivities could not be computed. Shown are specificities for the electrode position derived stability index (solid squares), for the beat-to-beat derived stability index (open circles) and for the combined index (solid circles). The vertical bars indicate the standard error.

The performance on the 'normal' database is shown in Fig. 6.3. Since no pathology was present in this set, sensitivities could not be computed. The specificity on the entire set was 95.7% (7,116/7,436). The specificity increased for maximal electrode position stability index (solid squares) to 97.3% (5,385/5,533). For maximal beat-to-beat stability index (open circles) specificity increased to 96.3% (3,440/3,571), while for maximal combined index (solid circles) it increased to 97.1% (2,953/3,040). Maximal specificity of 97.9% was found in the subgroup with combined stability index larger than 0.85. Standard errors of all estimates were less than 0.3%. All specificities decreased somewhat at the highest stability index 1.

6.3.2 The role of the stability index

The impact of a stability index for routine ECG interpretation depends on its distribution. When only few ECGs would have an index 1, the benefit of a higher performance in this subgroup is limited. The percentages of ECGs with maximum stability index in the three databases studied, however, were large: for all databases on average about 75% for the electrode position and beat-to-beat stability indices, and 60% for the combined stability index.

Generating simulated ECGs involves considerable computing resources. For routine application it is therefore desirable to decrease the number of simulated ECGs used to compute the electrode position stability index. When we reduced the number of position shifts from 10 to 3, results changed at most

0.2%, with the exception of the 'cardiologists' database, where the MI sensitivity and specificity for the most stable ECGs became 93% and 92%, respectively, as compared to 97% and 93% for 10 displacements.

Table 6.1. Scenario demonstrating the possible use of a stability index to balance the workload of an overreading cardiologist. Shown are expected sensitivity and specificity for MI (cols. 3 and 4) and the percentage of ECGs to be overread by a cardiologist (col. 5) when a selection of ECGs to be overread is made based on a particular strategy (col. 1) and a stability index (col. 2). The estimates are based on the results presented in Fig. 6.2c on page 96. The top row shows the values in case no overreading takes place. The following three rows show the values for strategy one: the cardiologist overreads all ECGs with a stability index smaller than a specific threshold. The bottommost three rows show the values for strategy two: the cardiologist overreads all ECGs classified by the computer program as abnormal and with a stability index smaller than a specific threshold. Note that performance indices are based on a cardiologist's ECG classification as a reference. The workload is based on an MI prevalence of 5%.

Strategy	Stability index threshold	Sensitivity (%)	Specificity (%)	Workload (% of ECGs)
No overreading		83	89	0
1	< 0.8	97	94	13
1	< 1.0	98	100	40
1	All ECGs*	100	100	100
2	< 0.8	83	94	5
2	< 1.0	83	100	12
2	All abnormal Dx*	83	100	14

* No stability index threshold is used in these cases, which means that in case of strategy 1 all ECGs are overread, and in case of strategy 2 all ECGs classified as abnormal are overread.

6.3.3 The stability index and a cardiologist's workload

The 'cardiologists' database was also used to show the use of a stability index in daily routine, where computer interpretations of ECGs are reviewed by a cardiologist. Two strategies were considered: 1) the cardiologist checks all ECGs with a stability index smaller than a particular threshold; 2) the cardiologist checks all ECGs with a stability index smaller than a particular threshold that are classified as abnormal by the computer program. Table 6.1 shows expected sensitivity and specificity for MI for each strategy for a few threshold values,

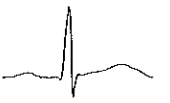
as well as the cardiologist's workload in terms of percentage of ECGs to be reviewed. The row labeled 'All ECGs' depicts the strategy where all ECGs are reviewed. The row labeled 'All abnormal Dx' depicts the strategy where all ECGs classified as abnormal are reviewed. These two strategies do not use a stability index as a selection criterion. The Table shows the results for an MI prevalence of 5%. For strategy 1 prevalence has no influence on workload. For strategy 2, a prevalence of 30% would increase the workload to 16% for a threshold value one and to about 28% for no threshold. Note that for this strategy the sensitivity remains constant since abnormal cases missed by the computer program will not be reviewed by the cardiologist.

6.4 Discussion

6.4.1 Stability index and diagnostic performance

With the use of an ECG computer program we showed that a stability index can be useful in discriminating reliable classifications from less reliable ones. For myocardial infarction, both an index based on electrode displacement simulations and one based on beat-to-beat variability help to identify subgroups that yield higher classification performance. A combination of both indices, in our study the minimum of the two, performs even better. Taking the ECG diagnosis of the cardiologists as a reference, we obtained an increase in sensitivity for myocardial infarction of up to 13% with a simultaneous increase of specificity of 11%. When a reference based on ECG-independent material was used, MI sensitivity remained constant but MI specificity increased with about 6%. On a database of normal ECGs, a specificity increase of up to 2% was found using an electrode position based stability index.

The difference in results between the 'cardiologists' database and the 'clinical' database (the latter having a lower and constant MI sensitivity) might be explained by the limited information content of the ECG itself. Possibly the ECG cannot be used to diagnose MI in all cases, so that MI sensitivity based on a clinical reference is limited to a value less than 100%. This was also illustrated in the CSE study, where the consensus ECG classification by a group of cardiologists was compared to a reference based on clinical evidence: the reported total accuracy, anterior MI sensitivity and inferior MI sensitivity were 80.3%, 85.7%, and 78.3%, respectively [16].



6.4.2 A possible explanation of the results

Apparently, classification of an averaged beat of the original ECG results in a non-optimal classification. A possible explanation is the nonlinearity of the automated ECG classification process. Classification errors made by analysis programs are often not gradual but discrete (i.e., the crossing of a threshold). Thus, small insignificant signal changes may result in large diagnostic classification changes. Combining classifications of multiple instances of the same ECG may prevent occasional errors to influence the final diagnostic interpretation. In addition, systematic measurement changes occur in the case of electrode position changes [12], and imperfect beat alignment may cause systematic wave duration increase and amplitude decrease. Providing a classification program with multiple sets of measurements enables a differentiating classification.

6.4.3 Stability index and the cardiologist's workload

The use of a stability index allows a cardiologist to balance between classification performance and workload. Instead of checking all ECGs considered abnormal by a computer program (a strategy advised by De Bruyne et al. [15]), which still implies a considerable workload, only ECGs indicated less reliable by the stability index could be checked, reducing the workload while maintaining good performance. Especially in population surveys or epidemiological studies, where large numbers of ECGs are to be processed automatically and only subsets can be overread manually, this method may help to improve overall classification performance.

In this study stability indices were computed for two diagnostic categories, normal and myocardial infarction. The method can easily be extended to other categories as well. Also different thresholds could be used for different diagnostic categories. In this way, one might optimize the performance for separate categories.

In conclusion, ECG classification variability owing to both technical and biologic sources, appears to contain information that can be used to assess the reliability of the classification. This information might also be exploited to further improve diagnostic accuracy of computerized ECG interpretation.

Acknowledgement

We are indebted to professor J. Michaelis from the Institut für Medizinische Statistik und Dokumentation der Universität Mainz who kindly provided the ECGs for the 'normal' database [11].

6.5 References

[1] Simonson E, Brozek J, Keys A. Variability of the electrocardiogram in normal young men. Am Heart J 1949;38:407–22.

[2] Farb A, Devereux RB, Kligfield P. Day-to-day variability of voltage measurements used in electrocardiographic criteria for left ventricular hypertrophy. J Am Coll Cardiol 1990;15(3):618–23.

[3] Wenger W, Kligfield P. Variability of precordial electrodes placement during routine electrocardiography. J Electrocardiol 1996;29(4):179–84.

[4] Kerwin AJ, McLean R, Tegelaar H. A method for accurate placement of chest electrodes in the taking of serial electrocardiographic tracings. Can Med Assoc J 1960; 82:258.

[5] Michaels L, Cadoret RJ. Day-to-day variability in the normal electrocardiogram. British Heart Journal 1967;29:913–9.

[6] Rautaharju PM, Wolf HK, Eifler WJ, Blackburn H. A simple procedure for positioning precordial ECG and VCG electrodes using an electrode locator. J Electrocardiol 1976;9(1):35–40.

[7] Herman MV, Ingram DA, Levy JA, Cook JR, Athans RJ. Variability of electrocardiographic precordial lead placement: a method to improve accuracy and reliability. Clin Cardiol 1991;14:469–76.

[8] Schijvenaars RJA, Kors JA, Van Herpen G, Van Bemmel JH. Use of the standard 12-lead electrocardiogram to simulate electrode displacements. J Electrocardiol 1996; 29 Suppl:5–9.

[9] Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Measurem 1960;20:37–46.

[10] Schouten HJA. Nominal scale agreement among observers. Psychometrika 1986; 51:453–66.

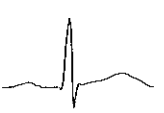
[11] Michaelis J, Lippold R, Gluck E, Schindler W, Scheidt E. Automated serial ECG analysis within an epidemiological study. J Electrocardiol 1987;20 Suppl:34–6.

[12] Schijvenaars RJA, Kors JA, Van Herpen G, Kornreich F, Van Bemmel JH. The effect of electrode position changes on ECG interpretation by computer. J Electrocardiol 1997;30(3):247–56.

[13] Van Bemmel JH, Kors JA, Van Herpen G. Methodology of the modular ECG analysis system MEANS. Methods Inf Med 1990;29(4):346–53.

[14] Willems JL, Abreu-Lima C, Arnaud P, Van Bemmel JH, Brohet C, Degani R, et al. The diagnostic performance of computer programs for the interpretation of electrocardiograms. N Engl J Med 1991;325:1767–73.

[15] De Bruyne MC, Kors JA, Hoes AW, Kruijssen DA, Deckers JW, Grosfeld M, et al. Diagnostic interpretation of electrocardiograms in population-based research: computer program research physicians, or cardiologists? J Clin Epidemiol 1997; 50(8):947–52.



[16] Willems J. Common Standards for Quantitative Electrocardiography. 10th and final progress report. Tech. Rep. CSE 90-12-31, Commission of the European Communities, 1990.

7 Improvement of diagnostic ECG Interpretation by using Intra-individual ECG Variability

R.J.A. Schijvenaars, J.A. Kors, G. van Herpen,
J.H. van Bemmel

Dept. of Medical Informatics, Erasmus University, Rotterdam, The Netherlands

Submitted for publication

Abstract

A method to increase the performance of ECG interpretation programs using ECG variability was designed. Multiple interpretations of one original ECG were produced by (1) interpreting each individual beat or (2) simulating and interpreting ECGs from shifted chest electrode positions. The interpretations for myocardial infarction were coded and combined using three methods. The hypothesis that the combined interpretations had higher performance than the original interpretation was tested on 94,010 ECGs. In about 5.5% of the cases at least one combined interpretation differed from the original one. An expert cardiologist interpreted a random set of 182 ECGs that had large interpretation differences. Taking his interpretation as the gold standard, the original computer interpretation was correct in on average 44 cases, the combined interpretations were correct in on average 104 cases. ECG interpretation variability, therefore, contains information that can be exploited to further improve the diagnostic accuracy of computerized ECG interpretation.

Keywords: electrocardiography, computer, ECG variability, electrode positions

7.1 Introduction

Computer programs for ECG interpretation perform almost as well as cardiologists [1]. However, in daily practice the ECG interpretation produced by such programs is often marred by signal disturbances and intra-individual ECG variability [2,3]. This variability is neglected in computer classification since most programs interpret one representative beat only and at one occasion only. Small differences between successive ECGs, or even between beats within one ECG, may result in different diagnostic interpretations due to the discrete nature of the interpretation algorithms. This may give an impression of unreliability. Further improvement of the performance of programs by consideration of the variability present in an ECG may increase their practicability in daily routine.

A possible method to increase the performance of an interpretation program might be to take ECG variability into account by combining multiple interpretations based on slightly different versions of the same ECG. Since the accuracy of current ECG analysis programs is high [1], the probability of misinterpretation is low. An adequate synthesis of multiple interpretations into one combined interpretation may further reduce this probability of misinterpretation. Others have reported that a combination of multiple interpretation algorithms indeed may lead to improvements in interpretation accuracy [4–6]. We tested whether a similar effect occurs when interpretations of multiple, slightly different simulated recordings of an ECG are combined.

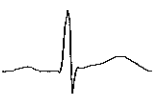
7.2 Methods

7.2.1 Material

The material consisted of 77,169 ECGs recorded in the cardiology department of the University Hospital Rotterdam, and 16,841 ECGs recorded in a cohort study among apparently healthy individuals according to clinical examinations, laboratory tests and questionnaires [7].

7.2.2 Electrocardiographic processing

The 12 leads of each ECG were recorded simultaneously at a sampling rate of 250 Hz for 10 seconds. Processing was done by our Modular ECG Analysis System (MEANS), which produces diagnostic statements using decision logic consisting of criteria that make use of measurement thresholds [8]. The



performance of MEANS has been tested extensively by ourselves [9] and by others [1]. We concentrated on one diagnostic category, myocardial infarction (MI). MEANS can assign one of three confidence qualifiers: 'possible', 'probable', or 'definite', coded as 1, 2, and 3, respectively. When no MI was reported a value of zero was coded. Multiple versions of each ECG were generated by using two types of variability that are often found in daily practice and that cause changes in signals not related to the quality of the recording. The first one is beat-to-beat variability, mostly caused by respiration. The second type of variability is caused by variations in chest electrode placement.

7.2.3 Individual beat interpretation

MEANS identifies all individual beats in the ECG and then produces one representative averaged beat for each lead for further analysis. The interpretation of this beat serves as the original interpretation. Multiple interpretations were generated by classifying each individual beat. Since the number of beats varies for each 10 second ECG, the total number of interpretations per ECG is not fixed. The interpretations of these individual beats will be referred to as 'beat-to-beat' interpretations.

7.2.4 Electrode displacement simulation

Multiple interpretations were also generated by simulating ECGs with upward and downward vertical displacement of all six chest electrodes in five steps of one cm (as may occur when the intercostal spaces are incorrectly determined). In a previous study, this type of displacement proved to have the largest effect on ECG interpretation among several other possible chest electrode position variations [3].

Each displaced precordial ECG was constructed through a linear combination of leads I, II, and the six precordial leads of the original ECG. The coefficients for these linear combinations had been derived previously from body surface potential maps (BSPMs), by applying a least-squares fit between the ECG signals obtained at the standard ECG positions and those from each of the five upward and five downward electrode locations, each 1 cm apart [10]. These simulated ECGs were shown to represent in good approximation ECGs from actually displaced electrode positions. MEANS was used to classify the averaged representative beat of each of these simulated ECGs in the same way as for the original ECG. This resulted in 10 additional interpretations with accompanying confidence qualifier: one for each of the 10 simulated ECGs. These interpretations will be referred to as 'electrode position' interpretations.

7.2.5 Combination of interpretations

Interpretation of individual beats may introduce variation in the interpretations caused by differences within the same recording. Interpretation of the averaged beat of the simulated ECGs may introduce changes in interpretation, now due to changes affecting all beats. To study the effect of these two types of variability separately, combining of interpretations was performed in two stages. In the first stage the set of 'beat-to-beat' interpretations was combined into one 'combined beat-to-beat' interpretation, and the 'electrode position' interpretations were combined into a 'combined electrode position' interpretation. Finally, in a second stage the three interpretations (original, combined beat-to-beat, and combined electrode position) were combined into one final interpretation.

Several methods to combine interpretations have been reported [1,6,11,12]. Some combination methods make use of a weighted voting scheme, using as weight a number representing the confidence of a particular interpretation: a high weight indicates a high confidence in the interpretation. Such a method can be applied in the second combination stage as well, since the variability in the two sets of individual interpretations can be used to define such a confidence measure of the combined interpretations. Previously, a stability index I has been developed, defined as:

$$I = 1 - \frac{1}{3n} \sum_{i=1}^{n} |x_i - x_0| \,, \tag{7.1}$$

with x_0 denoting the qualifier of the original interpretation, x_i the qualifier of interpretation i, and $3n$ the normalization factor for three levels of confidence to make the index independent of the number of interpretations [13]. This index is simple to compute and to understand. Maximum stability (i.e., all qualifiers are equal to the one of the original ECG) gives a stability index one, minimum stability (i.e., the original interpretation is an outlier compared to all other interpretations) an index zero. We have shown that this stability index is correlated to the performance of the original interpretation [13], which indicates its suitability as a confidence measure.

Two different combination methods were used. In the first method the median of the set of individual qualifiers was taken (method median). In the other, their rounded average was taken (method average). This latter method was also used in the study 'Common Standards for Quantitative Electrocardiography' (CSE) [1]. Averaging, however is influenced by interpretations that are exceptional (outliers), while taking a median interpretation is not.

In the second stage, these two combination methods were used as well. Method median consisted of taking the median of the three interpretation qualifiers (original, combined beat-to-beat, and combined electrode position) and did not make use of the confidence measures determined in the first combination stage. In method average the weighted average of the three interpretation qualifiers, using the stability index I as weighting factor was taken as final interpretation. The weight of the original interpretation was chosen to be one. In this way, combined interpretations resulting from a set with maximal stability (i.e., all equal qualifier codes and thus $I = 1$) had equal weights as the original interpretation. The final interpretations resulting from the two combination methods are denoted as Fin-median and Fin-average.

7.2.6 Reference interpretation

In 3,170 of the 94,010 ECGs the original interpretation showed a large difference (i.e., a qualifier code difference of 2 or 3) with at least one combined interpretation. From this set 182 ECGs were randomly selected and reviewed by an experienced cardiologist. The interpretation performances of the original interpretation and of all combined interpretations were computed taking his interpretation as the reference. Sensitivity and specificity were used as performance indices, using qualifier codes 0 and 1 ('absent' and 'possible') as non-MI interpretation, and qualifier codes 2 and 3 ('probable' and 'definite') as MI interpretations. Of the 182 cases, the cardiologist classified 143 cases as non-MI, and 39 as MI.

7.3 Results

Table 7.1 on the facing page shows the percentage of ECGs in which differences between the original and final interpretations were found for the various combination methods. Each row denotes a combination method and each column a positive or negative interpretation change of a specific size. Note that most changes are deviations of one qualifier code. The averaging method resulted in more but smaller interpretation changes, possibly because of its tendency to result in less pronounced interpretations. Also note that interpretation changes of size one (± 1) do not need to result in a change in performance as defined in our study, since a change from code 0 to 1 does not result in a change in interpretation from non-MI to MI.

In Table 7.2 on page 110 sensitivities and specificities for the 182 ECGs reviewed by the cardiologist are shown for the original interpretation and all combined

Table 7.1. The percentage (and number) of ECGs of which a combined interpretation differed from the original interpretation. Each row denotes a combination method, each column the number of ECGs with an interpretation difference of a specific size, expressed as a difference in qualifier codes (±1: small; ±2: medium; ±3: large changes).

Interpretation *	Percentage (nr) of ECGs		
	±1	±2	±3
BB-median	2.3 (2,135)	1.0 (976)	1.1 (1,064)
EP-median	1.8 (1,714)	0.7 (623)	0.7 (700)
Fin-median	0.9 (811)	0.3 (302)	0.3 (319)
BB-average	8.9 (8,404)	1.4 (1,277)	0.1 (131)
EP-average	7.9 (7,421)	1.1 (1,071)	0.1 (79)
Fin-average	4.6 (4,347)	0.1 (103)	0.0 (0)

* BB: beat-to-beat; EP: electrode position; Fin: final

interpretations. The standard errors are almost equal for all methods, and are about 8% for sensitivity, 4% for specificity, and 3.5% for MI accuracy. The performance increase is very large (about 30%) for all combination methods, although the percentage of cases involved was limited, as can be seen in Table 7.1. When the dichotomy between non-MI and MI was applied, about 0.5% of the ECGs for second stage combined interpretations and 1.5% for first stage ones showed differences.

With respect to the first combination stage, methods median and average both result in similar accuracies, except for the averaged electrode position interpretations, which has a somewhat lower accuracy. Method average causes higher sensitivities and lower specificities than method median. The second combination stage, in which the original interpretation was taken into account, showed no improvement over the first combination stage. In fact, combination method Fin-average shows lower specificity than both first stage combinations it was based on. However, it was still higher than the original interpretation. The qualifier codes produced by method average are less pronounced than those produced by method median: all qualifier codes of the Fin-average method were either 1 or 2, resulting in less interpretation changes than the other methods.

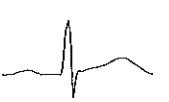


Table 7.2. The performance indices for the 182 randomly selected ECGs reviewed by a cardiologist. The first column shows the interpretation method. Columns 2 to 4 denote sensitivity, specificity and total accuracy for MI. Columns 5 to 7 show the number of cases in which the interpretation improved, remained the same, or deteriorated compared to the original interpretation.

Interpretation *	MI sens	MI spec	MI acc	better	equal	worse
Original	38.5	28.7	30.8	0	182	0
BB-median	61.5	71.3	69.2	126	0	56
EP-median	53.8	70.6	67.0	115	18	49
Fin-median	53.8	70.6	67.0	115	18	49
BB-average	69.2	69.9	69.8	119	15	48
EP-average	66.7	59.4	61.0	103	31	48
Fin-average	66.7	42.0	47.3	46	120	16

* BB: beat-to-beat; EP: electrode position; Fin: final

7.4 Discussion

We showed that combining interpretations of multiple instances of the same ECG leads to a higher ECG interpretation performance. The percentage of ECGs in which the interpretation changed was relatively low, but the number of improvements was more than twice the number of deteriorations. In addition, the sensitivity and specificity for this subgroup increased considerably.

A possible cause for this interpretation improvement may be the discrete nature of the interpretation process. The cases in which the interpretation changed probably have characteristics that are close to decision thresholds. Signal changes then may result in different interpretations. This assumption is confirmed by the distribution of interpretations of the individual beats and the simulated ECGs in the group selected for review by a human expert: a large variation of qualifiers. The original interpretation stems from one averaged beat. Due to the discrete, non-linear nature of the decision logic, this proximity to one or more decision thresholds is not necessarily reflected in the interpretation. In cases with an incorrect original interpretation, some measurements will be located at the wrong side of a decision threshold. Due to the threshold proximity, it is likely that some of the individual beats or ECGs computed by simulated chest electrode position changes will lie on the correct side of the thresholds. More strongly, the majority of these cases may lie on the correct side of the thresholds. Combining these interpretations consequently results in

a correct interpretation.

The smaller performance increase when using interpretation averaging using a weighting factor in stage two may be due to the large weighting of the original interpretation. If the set of interpretations of either beat-to-beat or electrode position simulations was not 'unanimous', the weighting factor of the combined interpretation was lower. This will cause the final interpretation to stick to the original interpretation more often. Another difference between methods median and average is that the latter tends to result in less pronounced interpretations (i.e., qualifier codes 1 or 2 instead of codes 0 or 3). This can be seen in Table 7.1 on page 109, where the Fin-average method results in a very low number of ECGs having a large interpretation change.

In conclusion, ECG interpretation variability contains information that can be exploited to further improve diagnostic accuracy of computerized ECG interpretation. Both types of variability studied can help to increase performance, but a combination of both gives no further improvement.

7.5 References

[1] Willems JL, Abreu-Lima C, Arnaud P, Van Bemmel JH, Brohet C, Degani R, et al. The diagnostic performance of computer programs for the interpretation of electrocardiograms. N Engl J Med 1991;325:1767–73.

[2] Zywietz C, Willems JL, Arnaud P, Van Bemmel JH, Degani R, Macfarlane PW. Stability of computer ECG amplitude measurements in the presence of noise. Comput Biomed Res 1990;23:10–31.

[3] Schijvenaars RJA, Kors JA, Van Herpen G, Kornreich F, Van Bemmel JH. The effect of electrode position changes on ECG interpretation by computer. J Electrocardiol 1997;30(3):247–56.

[4] Breiman L. Bagging predictors. Machine Learning 1996;24(2):123–40.

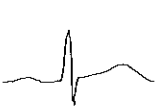
[5] Quinlan J. Bagging, boosting, and C4.5. In: Thirteenth National Conference on Artificial Intelligence. Portland, Oregon, USA: AAAI Press/MIT Press, 1996; pp. 725–30.

[6] Ali K, Pazzani M. Error reduction through learning multiple descriptions. Machine Learning 1996;24:173–202.

[7] Michaelis J, Lippold R, Gluck E, Schindler W, Scheidt E. Automated serial ECG analysis within an epidemiological study. J Electrocardiol 1987;20 Suppl:34–6.

[8] Van Bemmel JH, Kors JA, Van Herpen G. Methodology of the modular ECG analysis system MEANS. Methods Inf Med 1990;29(4):346–53.

[9] Kors JA. Expert knowledge for computerized ECG interpretation. Phd thesis, Erasmus University, 1992.



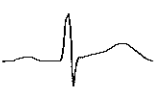
[10] Schijvenaars RJA, Kors JA, Van Herpen G, Van Bemmel JH. Use of the standard 12-lead electrocardiogram to simulate electrode displacements. J Electrocardiol 1996; 29 Suppl:5–9.

[11] Kors JA, Van Bemmel JH. Combination of interpretations from different experts. In: Lun KC, Degoulet P, Piemme TE, Rienhoff O, editors. MEDINFO 92. North-Holland Publ Comp, 1992; pp. 313–6.

[12] Bauer E, Kohavi R. An empirical comparison of voting classification algorithms: bagging, boosting, and variants. Machine Learning 1999;36:105–39.

[13] Schijvenaars RJA, Kors JA, Van Herpen G, Van Hemel NH, May JF, Van Bemmel JH. Use of electrocardiographic variability to improve the reliability of diagnostic interpretation. Tech. rep., Institute of Medical Informatics, 2000.

8 General Discussion



Over the past decades, the importance of electrocardiography as a diagnostic tool in cardiology has declined [1,2], mainly because the ECG provides only indirect information about cardiac anatomy and physiology, whereas newer techniques provide a more direct assessment. Nevertheless, cardiologists still consider the ECG to be the most important technique for detection of, e.g., acute myocardial infarction, arrhythmia and transient ischemia. In addition, the importance of computers in ECG analysis is increasing due to various developments:

- Subtle or complicated ECG characteristics have been shown to have prognostic value [3–5]. The utility of such characteristics in daily practice can be exploited only when they are determined precisely and reliably, which calls for automatic analysis.
- Very large databases of ECGs have been and are being collected in hospitals and in epidemiological studies [6]. Epidemiological studies involving interpretation of large quantities of ECGs is only practicable if computer processing is used in one or more phases of the research.
- ECG recording and interpretation is extending from the hospital to ambulances, general practitioners, etc. These latter healthcare providers have less knowledge than cardiologists on how to interpret an ECG and therefore depend more on computerized interpretation which, consequently, increases the importance of the quality of these interpretations.

The performance of computer programs for ECG interpretation has been assessed earlier [7]. It was concluded that the best computer programs perform almost as well as cardiologists in identifying seven major cardiac disorders. Whether the current programs are sufficiently reliable remains a subject of discussion [8]. There seems little doubt, however, that the stability of computerized ECG interpretations leaves room for improvement [9]. Intra-individual ECG variability, either non-pathological biologic variability or variations caused by recording circumstances, regularly result in changes in computerized diagnostic interpretation.

In the introduction, four questions were formulated to be addressed in this research. Here, we discuss the (tentative) answers to these questions.

Question 1:

What are the most prominent sources of intra-individual ECG variability?

The review of the literature presented in Chapter 2 reveals the progress made in the quality of the recording equipment. Nowadays, differences in equipment should no longer be a source of intra-individual variability. Other, trivial sources of intra-individual ECG variability, such as the lead system, are not addressed here. Intra-individual variability can be subdivided into two main types: variability *within one ECG* (beat-to-beat variability) and variability *between different ECGs* of the same individual. Variability of the first type will, of course, contribute to the variability between ECGs.

In daily routine, within one ECG the sources with the largest effect on variability are:

- Respiration. This is the major source of beat-to-beat variability. Due to changes in the volume conductor and in the position of the heart, ECGs measured on the body surface are modulated by the respiration.
- Noise. In practice, high noise levels may be less of a problem than lower levels, since signals of obviously low quality will often prompt rerecording of the ECG. Lower noise levels tend to be tolerated, but can result in significant differences in measurement between beats within one recording [10]. Although signal differences may be small, they can cause large differences in diagnostic interpretation due to, e.g., wave relabeling [11].

Normal variability between different ECGs of the same individual (excluding changes caused by cardiac pathology) is caused mainly by:

- Inaccurate chest electrode placement. This remains one of the major causes of ECG variability. Electrodes may be interchanged, or may be placed in the correct order but in an incorrect position, or both. Large changes may occur when electrodes are shifted. Human interpreters are better at handling such errors than computer programs [9].
- Age and age-related factors (e.g., weight). These effects on the ECG can be considerable and have been well documented [12–15]. Although these factors introduce changes between successive recordings, the general trend of these effects can be anticipated and should not be a major problem to a competent interpreter, either human or computer.

Other sources, such as high altitude or pregnancy, also cause smaller or larger changes in the ECG, but are relatively uncommon in daily practice. Others, such as the patientŠs position during recording (supine, sitting or standing) will influence the appearance of the ECG, but such changes can be considered

as changes in recording protocol rather than variability. If they do occur, however, the interpreter should be informed of the recording conditions. In general, little collateral information is provided with the routine ECG recording. Noise and electrode position variability result in unpredictable intra-individual ECG variability in individual cases. Thus, even if the general effects of these sources would be well known (but they are not), it is difficult to correct for them in individual ECGs.

Question 2:

What is the effect of chest electrode position changes on ECG measurement and interpretation?

Several studies have investigated the effects of chest electrode mispositioning [16–18]; however, all lacked one or more of the following desirable characteristics:

- A sufficiently large set of ECGs.
- Use of simultaneously recorded leads. When assessing the effect of chest electrode position changes, other causes of variability should preferably be excluded. This is guaranteed only if the signals from different electrode positions are recorded simultaneously.
- Assessment of the effects on interpretation, not just on ECG waveshape or measurements. Large changes in waveshape or measurements may not be diagnostically important; small changes, however, may be important.

We assessed the effect of chest electrode position changes on ECG measurements and interpretation avoiding the above limitations. From a total of 746 Body Surface Potential Maps (BSPMs) we extracted the six precordial leads from varying locations on the chest. Some of those locations were not part of the BSPMs and thus needed to be interpolated. Chapter 3 presents a study investigating which method for interpolation of BSPMs was most suitable for our purposes. Because, in the horizontal direction these maps are circular and closed, Fourier interpolation was the most suitable method. In the vertical direction, cubic splines provided the best results.

Chapter 4 presents the actual assessment of the changes in measurements and interpretation that results from changing chest electrode positions. The effect of four types of displacement was studied: a longitudinal shift of all six chest electrodes, a transversal shift of electrodes V_1 to V_3, a transversal shift of electrodes V_5 and V_6, and a rotation of all six chest electrodes simulating heart

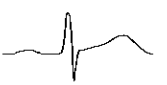
position changes due to, e.g., obesity. Longitudinal positional shifts had the largest effect. In individual cases changes in both measurement and diagnostic interpretation could be considerable.

In only a small percentage of the ECGs was the diagnostic interpretation affected by large shifts in electrode position. In screening situations, such intra-individual ECG interpretation divergence may considerably limit the accuracy of estimates of prevalence and incidence. A human expert made about half the number of changes in interpretation compared to our computer program MEANS. A conclusion from this study is, therefore, that chest electrode mispositioning can have a considerable effect on ECG measurement and interpretation. Another conclusion is that computer programs for ECG analysis may suffer more from these effects than human interpreters. Since these results were obtained with the use of only one ECG analysis computer program and only one human expert, additional research is needed to verify that these conclusions also hold in general.

Question 3:

Can intra-individual variability be assessed from the standard 12-lead ECG alone?

The susceptibility described above limits the usefulness of ECG computer programs, despite their reported good performance [9]. We investigated whether it was possible to provide a program with information about the magnitude of changes in an individual ECG that would occur if chest electrodes were shifted. If BSPMs were available, this would simply be a matter of repeating the experiments described in Chapter 4. In daily routine, however, only 12-lead ECGs are recorded rather than complete BSPMs. The study presented in Chapter 5 sought a way to circumvent this problem by simulating the ECG signals that would be recorded from translated electrode positions. Using least squares techniques, patient-independent linear transformations were determined that approximated the ECG signals recorded at translated positions using the ECG from the standard precordial electrode locations. These approximated signals were compared with their actually recorded counterparts. The differences proved to be about half the difference between the signals from the standard electrode locations and the recorded displaced ones. Differences in diagnostic interpretation were also determined. The percentages of large differences in diagnostic interpretation for extreme shifts dropped from about 8% for actually displaced vs. original ECGs, to about 2% for actually displaced vs. approximated ECGs. These general, patient- independent simulations therefore appear sufficiently accurate to indicate which ECGs are likely to show large diagnostic differences as a result of such shifts.



Question 4:

Can intra-individual variability be exploited to improve computerized ECG interpretation?

The degree to which an individual ECG interpretation is sensitive to possible electrode position misplacement or beat-to-beat variability may be used as an alert to cardiologists. However, this knowledge gains additional value if it is correlated with the performance of the interpretation: are stable classifications more accurate than instable ones? This question is investigated in Chapter 6. A stability index was designed to express the amount of variability in diagnostic interpretation for a particular ECG. Two types of variability were determined: variability due to electrode position changes by using a number of simulated, displaced ECGs, and beat-to-beat variability using individual representative beats.

ECGs with known reference were used to test whether the performance of our ECG analysis program was systematically higher for ECGs with a high stability index, i.e., a low amount of variability in diagnostic interpretation. This proved to be the case. This result allows to incorporate information about the variability of an individual ECG into its diagnostic interpretation. For example, the ECGs with a high stability index are likely to yield more reliable interpretations than ECGs with a low stability index. This may be useful in selecting ECGs for overreading in, e.g., large clinical trials or in an epidemiological setting.

In addition to estimating the reliability of the interpretation of an individual ECG, it might be possible to use information about the variability of an individual ECG to improve its interpretation. This topic is investigated in Chapter 7. In addition to the standard interpretation, two additional ones were generated, one deduced from the interpretations of the 10 simulated displaced ECGs and one extracted from the interpretations of all dominant separate beats. Differences between these three interpretations existed for a small number of ECGs. Of these, (according to a human expert) the two newly derived additional interpretations were twice as often correct as the original interpretation. This result should be interpreted with caution, since only one human expert was consulted (although his ECG interpretations proved to be in very good agreement with two other cardiologists, as shown in Chapter 6), and only one computer program (MEANS) was used. Nevertheless, this approach could offer a way to increase the performance of ECG analysis computer programs.

The research presented in Chapters 5 to 7 concentrates on the diagnostic category of myocardial infarction, the detection of which is one of the most important purposes of ECG analysis [1]. Whether our results also hold for other diagnostic categories requires additional studies.

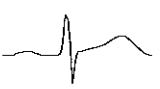
Stability improvement

Our research indicates that identification of ECGs with poor stability for beat-to-beat or chest electrode position changes is feasible. The next step is to improve this stability, but this has not been studied here. One might investigate whether the alternative interpretations presented in Chapter 7, besides being more accurate, are also more stable than the original interpretation. This method, however, does not deal with the causes of poor stability, e.g., discrete decision thresholds or loopholes in the decision logic. For that purpose other methods, involving improvement of the decision logic or even replacement of discrete decision thresholds with thresholds that incorporate imprecision, are more suitable.

In addition to intra-individual ECG variability, ECG interpretation also needs to take into account inter-individual ECG variability. The latter is generally larger than intra- individual variability (e.g., [19]), and is often quantified using normal limits, indicating between which values a particular measurement will lie in, e.g., 95% of normal ECGs. Attempts to decrease inter-individual variability using correction methods for the geometry of the thorax have not been successful so far [20].

Performance improvement

Given the large amount of work invested in designing and improving automated ECG analysis, why do human experts still perform better in ECG interpretation than computer programs? The first difficulty lies in the signal analysis. Human readers still outperform computer programs in recognizing waveforms, in particular P waves. Another cause may be the large number of measurements that must be examined. Designing algorithms that generate a reasonable diagnostic interpretation for all possible combinations of ECG measurements is a very complicated task and requires large sets of validated ECGs. As a result, current algorithms may produce unlikely interpretations for certain combinations of measurements. The same holds for the ECG waveform recognition algorithms: some ECGs will cause these algorithms to produce faulty measurements, which consequently result in erroneous interpretations. Interesting in that respect is the finding that accuracy increases by combining interpretations, either multiple interpretations by one program (as shown in this thesis), or interpretations of multiple programs [21,22]. When combining, mistakes in one interpretation are possibly overruled by correct statements in the other interpretations. The method described in Chapter 7 implements this strategy in a simplified way. Another possible approach might be that followed in pattern



recognition and expert systems design, where hybrid systems are developed that exploit different techniques of classification and combine the results.

Part of the answer may also lie in the fact that only human experts perform better, whereas physicians with only basic knowledge on ECG characteristics of cardiac pathology do not. One of the reasons that experts do better is that they possess and use insight in the underlying physical processes that result in an ECG on the body surface. If realistic models of these processes are incorporated into computer programs for ECG analysis the latter may ultimately reach a performance level comparable to that of human experts.

An alternative approach is to generate decision algorithms automatically. The amount of work to be done by human experts is one of the major bottlenecks in the further improvement of ECG analysis systems. It is difficult (and time-consuming) for human experts to formalize the knowledge they apply when interpreting ECGs, compared with providing ECG interpretations only. These interpretations can be used by automated learning techniques that generate decision rules on the basis of diagnostically documented ECGs [23]. In addition, these methods may be used to improve decision algorithms during routine use, similar to the gain of experience in human interpreters.

In conclusion, intra-individual ECG variability due to beat-to-beat variability and chest electrode positioning can be assessed for individual ECGs. This allows to distinguish between more and less reliable computerized ECG interpretations, which is valuable in the selection of ECG diagnostic interpretations to be checked manually.

8.1 References

[1] Selvester RH. A call to concerted action for change. The recording, display, and reporting of the routine electrocardiogram. J Electrocardiol 1998;31(4):367–70.

[2] Mirvis DM. What's wrong with electrocardiography? J Electrocardiol 1998; 31(4):313–6.

[3] Signal-averaged electrocardiography. J Am Coll Cardiol 1996;27(1):238–49.

[4] De Bruyne MC, Hoes AW, Kors JA, Hofman A, Van Bemmel JH, Grobbee DE. QTc dispersion predicts cardiac mortality in the elderly: the Rotterdam Study. Circulation 1998;97(5):467–72.

[5] Kors JA, De Bruyne MC, Hoes AW, Van Herpen G, Hofman A, Van Bemmel JH, et al. T axis as an indicator of risk of cardiac events in elderly people. Lancet 1998; 352(9128):601–5.

[6] Rautaharju P. Electrocardiography in epidemiology and clinical trials. In: Macfarlane P, Lawrie T, editors. Comprehensive electrocardiography. Theory and practice in health and disease, vol. 2, 1st ed. Oxford: Pergamon Press, 1989; pp. 1219–66.

[7] Willems JL, Abreu-Lima C, Arnaud P, Van Bemmel JH, Brohet C, Degani R, et al. The diagnostic performance of computer programs for the interpretation of electrocardiograms. N Engl J Med 1991;325:1767–73.

[8] RuDusky BM. Errors of computer electrocardiography. Angiology 1997; 48(12):1045–50.

[9] Hurst JW, Treasure CB, Sathavorn CS. Computer errors in electrocardiography. Clin Cardiol 1996;19(7):580–6.

[10] Zywietz C, Willems JL, Arnaud P, Van Bemmel JH, Degani R, Macfarlane PW. Stability of computer ECG amplitude measurements in the presence of noise. Comput Biomed Res 1990;23:10–31.

[11] Kors JA, Van Herpen G, Van Bemmel JH. Variability in ECG computer interpretations: analysis of individual complexes vs analysis of a representative complex. J Electrocardiol 1992;25:263–71.

[12] Rautaharju PM, Zhou SH, Calhoun HP. Ethnic differences in ECG amplitudes in North American white, black, and Hispanic men and women. Effect of obesity and age. J Electrocardiol 1994;27 Suppl:20–31.

[13] Gambill CL, Wilkins ML, Haisty W. K. J, Anderson ST, Maynard C, Wagner NB, et al. T wave amplitudes in normal populations. Variation with ECG lead, sex, and age. J Electrocardiol 1995;28(3):191–7.

[14] Simonson E. The effect of age on the electrocardiogram. Am J Cardiol 1972; 29(1):64–73.

[15] Kilty SE, Lepeschkin E. Effect of body build on the QRS voltage of the electrocardiogram in normal men. Its significance in the diagnosis of left ventricular hypertrophy. Circulation 1965;31:77–84.

[16] Herman MV, Ingram DA, Levy JA, Cook JR, Athans RJ. Variability of electrocardiographic precordial lead placement: a method to improve accuracy and reliability. Clin Cardiol 1991;14:469–76.

[17] Hill NE, Goodman JS. Importance of accurate placement of precordial leads in the 12-lead electrocardiogram. Heart Lung 1987;16:561–6.

[18] Berson AS, Haisty WK, Pipberger HV. Electrode position effects on Frank lead electrocardiograms. Am Heart J 1978;95:463–73.

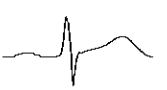
[19] Simonson E, Brozek J, Keys A. Variability of the electrocardiogram in normal young men. Am Heart J 1949;38:407–22.

[20] Hoekema R. The interindividual variability of the electrocardiogram. Phd thesis, Katholieke Universiteit Nijmegen, 1999.

[21] Willems JL, Abreu-Lima C, Arnaud P, Van Bemmel JH, Brohet C, Degani R, et al. Effect of combining electrocardiographic interpretation results on diagnostic accuracy. Eur Heart J 1988;9:1348–55.

[22] Kors JA, Van Herpen G, Willems JL, Van Bemmel JH. Improvement of automated electrocardiographic diagnosis by combination of computer interpretations of the electrocardiogram and vectorcardiogram. Am J Cardiol 1992;70:96–9.

[23] Briscoe G, Caelli T. A compendium of machine learning, vol. 1 of Ablex series in artifical intelligence. Norwood: Ablex Publishing Corp, 1996.



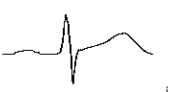
Summary

Four main research questions are addressed in this study:

1. What are the most prominent sources of intra-individual ECG variability?
2. What is the effect of chest electrode position changes on ECG measurements and interpretations?
3. Can intra-individual variability be assessed from the standard 12-lead ECG alone?
4. Can intra-individual variability be exploited to improve computerized ECG interpretation?

Chapter 2 presents the results of a literature study that was undertaken to address the first question. Two types of intra-individual ECG variability are discerned: variability within one ECG (beat-to-beat variability) and variability between successive ECGs, not due to pathological changes. The first type of variability is caused mainly by respiration and noise. The main causes of the second type of variability are variations in chest electrode position, and (as a long-term operator) age and age-related factors. An important feature of variability caused by respiration and chest electrode position is that these sources are not annotated in daily practice: neither the respiration phase nor the chest electrode positions are noted for future reference.

The effects of changes in chest electrode position on the ECG were investigated in more detail. Body surface potential maps (BSPMs) were used to ensure simultaneous recording of ECGs from many chest electrode locations. The BSPMs were interpolated in order to study those electrode positions which were not recorded. Chapter 3 presents an assessment and comparison of the performance of four interpolation methods (Fourier transforms, Chebyshev polynomials, linear functions, and cubic splines) for a number of electrode



grid densities. For the horizontal direction, Fourier transformation is the best method for the densest grid whereas cubic splines have the best performance for less dense grids. In the vertical direction, interpolation of both periodic and nonperiodic grids was studied; cubic splines again performed best for both grid types and all but the sparsest grids.

Chapter 4 addresses the effect of changes in chest electrode position on ECG measurements and interpretations. ECGs from selected chest electrode positions were extracted from the interpolated BSPMs and used to assess the effect of chest electrode position changes on the ECG. The effects on four measurements (QRS duration, Q duration, Q amplitude, and the Sokolow index) were studied, as well as the effects on two diagnostic categories: myocardial infarction (MI), regardless of location, and left ventricular hypertrophy (LVH). Four types of electrode displacements were performed: a longitudinal translation of all six chest electrodes, a transversal shift of electrodes V_1 to V_3, a transversal shift of V_5 and V_6, and a rotation of all six chest electrodes around V_2. A longitudinal translation of all chest electrodes had the largest effect on the measurements and the diagnostic interpretations. Although the effect was small for most ECG measurements, large changes of up to 15 μs in QRS duration and 1 mV in Q amplitude (in these cases QS amplitudes) occurred. The diagnostic interpretations were also most affected by a longitudinal electrode translation. In about 15% of the ECGs the diagnostic interpretation for myocardial infarction changed for electrode translations of 5 cm. The effect on LVH interpretation was smaller: 8% of the ECGs showed changes for the same 5 cm displacement. The electrode displacements had less effect on the interpretations of a human expert, although in about half of the cases in which the program found a large change, the expert also gave different interpretations. Thus, chest electrode position changes can have large effects on both the measurement and diagnostic interpretation of the ECG.

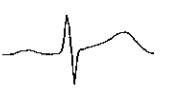
Chapter 5 discusses whether the effect of electrode position changes can be assessed for an individual standard 12-lead resting ECG. For each of the four chest electrode displacement types described earlier, linear transformations were derived (one for each magnitude of displacement) to approximate ECGs recorded at displaced electrode positions. These transformations were determined using a least squares method to fit the six standard precordial signals to the six actual ECG signals at each of the displaced positions, as recorded from the interpolated BSPMs. The mean absolute signal difference between actually displaced and approximated displaced ECG (maximally 200 μV for a 5 cm translation of all chest electrodes in longitudinal direction) was about half that of the difference between actually displaced and standard precordial ECG. The differences in diagnostic interpretation for MI and LVH decreased with about the same factor (for MI, from maximally 7% to 3% for longitudinal translation). These results led to the conclusion that ECGs from displaced electrodes can be

reasonably well simulated using a linear transformation of the standard ECG.

Chapters 6 and 7 investigate whether intra-individual variability can be exploited to improve computerized ECG interpretation. Chapter 6 explores whether the two main types of intra-individual ECG variability (beat-to-beat variability and chest electrode positioning variability) can be used to obtain an estimate of the accuracy of the diagnostic interpretation of the standard ECG. A stability index was employed to quantify the two types of variability. The performance of the diagnostic interpretation of MI (expressed in sensitivity and specificity) was assessed for a database of ECGs, the clinical reference being provided by discharge letters. In relation to the entire set of ECGs, in subsets of ECGs with lower variability specificity was higher but sensitivity was the same; using the combined ECG interpretation of three cardiologists as a reference, both sensitivity and specificity of the subset were higher. Thus, the degree of intra-individual variability of an individual ECG provides information about the accuracy of the automated interpretation of that ECG, especially when a cardiologistŠs ECG interpretation is taken as a reference. This may be useful in the selection of ECGs for manual overreading in, e.g., an epidemiological setting.

Chapter 7 investigates whether the diagnostic interpretation of the standard ECG can be improved by combining the interpretations of both the individual beats and the chest electrode displacement simulations into alternative interpretations. These combined interpretations were then compared with the original interpretation using a large database of ECGs. In only a small percentage of cases did the interpretations prove to be different: small differences occurred in (at most) 9% of the cases, large differences in (at most) 1%. From the group of cases with large differences in interpretation, a number of ECGs was randomly selected and reviewed by a cardiologist. Taking his interpretation as a reference, the combined interpretations proved to be correct in more than twice as many cases as the original interpretation.

Chapter 8 discusses the possibilities to further improve the performance and stability of ECG analysis programs in light of our investigations. We conclude that stability of a program can be assessed in the case of an individual ECG. A combination of interpretations, either from different programs or using methods similar to the ones presented in this thesis, is promising to improve performance. Improvement of stability will be difficult to achieve, probably requiring modifications in the techniques used in the interpretation process. Minimizing electrode position variability will therefore remain an important issue.



Samenvatting

Het onderzoek in dit proefschrift behandelt vier onderzoeksvragen:

1. Wat zijn de belangrijkste oorzaken van intra-individuele ECG variabiliteit?
2. Wat is het effect van borstwand elektrode positie veranderingen op ECG metingen en interpretaties?
3. Kan intra-individuele variabiliteit worden bepaald op alleen het 12-afleidingen ECG?
4. Kan intra-individuele variabiliteit worden gebruikt om automatische ECG interpretatie te verbeteren?

Hoofdstuk 2 presenteert de resultaten van een literatuurstudie gedaan om de eerste onderzoeksvraag te beantwoorden. Twee typen intra-individuele ECG variabiliteit kunnen worden worden onderscheiden: variabiliteit binnen één ECG (zogenaamde slag-op-slag variabiliteit) en variabiliteit tussen opeenvolgende ECG's, waarbij deze niet veroorzaakt worden door pathologische veranderingen. Het eerste type variabiliteit wordt voornamelijk veroorzaakt door ademhaling en ruis. De belangrijkste oorzaken van het tweede type zijn variaties in de borstwand elektrode posities en (als oorzaak op de lange termijn) leeftijd en daaraan gerelateerde veranderingen. Een belangrijk aspect van variabiliteit veroorzaakt door ademhaling en borstwand elektrode positie is dat deze in de dagelijkse praktijk niet worden genoteerd tijdens de opname: ademhalingsfase noch borstwand elektrode positie worden aangetekend voor mogelijk toekomstige naslag.

De effecten van borstwand elektrode positie veranderingen op het ECG zijn verder onderzocht. Opnamen van de potentiaalverdeling op het lichaamsoppervlak (Body Surface Potential Maps, afgekort BSPM's) zijn gebruikt om

gelijktijdige ECG opnames op een groot aantal verschillende elektrode posities te verkrijgen. Deze BSPM's werden geïnterpoleerd om ook signalen van niet opgenomen elektrode posities te kunnen bestuderen. Hoofdstuk 3 presenteert een evaluatie en vergelijking van de prestatie van vier interpolatiemethoden (Fourier transformaties, Chebyshev polynomen, lineaire functies, en zogenaamde 'cubic splines') voor een aantal elektrode roosters van verschillende dichtheid. In horizontale richting is Fourier transformatie de beste methode voor het meest dichte elektroderooster, terwijl 'cubic splines' het beste presteerde voor minder dichte roosters. In verticale richting is de interpolatie van zowel periodieke als niet-periodieke roosters onderzocht. Ook hier bleek 'cubic splines' de beste methode voor beide roostertypen en voor alle roosterdichtheden op de minst dichte na.

Hoofdstuk 4 gaat in op de tweede onderzoeksvraag over het effect van borstwand elektrode positie veranderingen op ECG metingen en interpretatie. Uit de geïnterpoleerde BSPM's zijn ECG's van bepaalde borstwand elektrode posities afgeleid en gebruikt om het effect van borstwand elektrode posities te meten. Deze effecten zijn bepaald voor vier metingen (QRS duur, Q duur, Q amplitude, en de Sokolow index) en twee diagnostische categorieën: myocard infarct (MI), ongeacht waar in het hart, en linker ventrikel hypertrofie (LVH). Vier typen elektrodeverplaatsingen zijn uitgevoerd: een longitudinale verplaatsingen van alle zes de borstwand elektroden, een transversale verplaatsing van elektroden V_1 tot V_3, een transversale verschuiving van V_5 en V_6, en een rotatie van alle borstwand elektroden rond V_2. De longitudinale verplaatsing van alle zes de borstwand elektroden bleek het grootste effect op zowel metingen als interpretatie te hebben. Hoewel het effect klein was voor de meeste ECG metingen, kwamen ook grote veranderingen tot aan 15 μs in QRS duur en 1 mV in Q amplitude (in dergelijke gevallen betrof het QS complexen) voor. Ook de diagnostische interpretatie werd het meest beïnvloed door de longitudinale elektrode verplaatsing. In ongeveer 15% van de ECG's veranderde de interpretatie voor myocard infarct als gevolg van een elektrode verplaatsing van 5 cm. Het effect op de interpretatie voor LVH was kleiner: in 8% van de ECG's traden veranderingen op als gevolg van eenzelfde verplaatsing van 5 cm. De interpretaties van een menselijke expert waren minder gevoelig voor de elektrode verplaatsingen, hoewel in ongeveer de helft van de gevallen waarin het programma een grote verandering zag, ook de expert verschillende interpretaties gaf. Het blijkt derhalve dat borstwand elektrode veranderingen grote effecten op het ECG kunnen hebben, die tot uiting komen zowel in metingen als in diagnostische interpretatie.

Hoofdstuk 5 bespreekt of het effect van elektrode positie veranderingen bepaald kan worden voor één individueel 12-afleidingen ECG. Voor elk van de eerder beschreven vier borstwand elektrode verplaatsingstypes zijn lineaire transformaties bepaald (één voor elke mate van verplaatsing) teneinde op ver-

plaatste posities opgenomen ECG's te benaderen. Deze transformaties zijn bepaald gebruik makend van een kleinste kwadraten methode die de zes standaard precordiale signalen vergeleek met zes signalen afkomstig van andere posities, geëxtraheerd uit de geïnterpoleerde BSPM's. Het gemiddelde absolute verschil tussen werkelijk verplaatst en benaderd verplaatst ECG (maximaal 200 μV voor een translatie van 5 cm van alle borstwand elektroden in longitudinale richting) was ongeveer de helft van het verschil tussen werkelijk verplaatst en standaard precordiale ECG. Het verschil in diagnostische interpretatie voor MI en LVH verminderde met ongeveer dezelfde factor (voor MI van maximaal 7% naar 3% voor een longitudinale verplaatsing). Uit deze resultaten concluderen wij dat ECG's afkomstig van verplaatste borstwand elektroden vrij goed kunnen worden gesimuleerd met behulp van een lineaire transformatie van het standaard ECG.

Hoofdstukken 6 en 7 onderzoeken of intra-individuele variabiliteit gebruikt kan worden om automatische ECG interpretatie te verbeteren. Hoofdstuk 6 gaat na of de twee typen intra-individuele ECG variabiliteit (slag-op-slag en borstwand elektrode positie variatie) gebruikt kunnen worden om een schatting van de nauwkeurigheid van de diagnostische interpretatie te verkrijgen. Een stabiliteits-index werd toegepast om de twee typen variabiliteit te kwantificeren. De prestatie van de diagnostische interpretatie van MI (uitgedrukt in sensitiviteit en specificiteit) werd bepaald op een database van ECG's, waarbij de klinische referentie gebaseerd was op de ontslagbrief. In vergelijking met de gehele verzameling ECG's was de specificiteit hoger voor deelverzamelingen ECG's met kleinere variabiliteit, maar bleef de sensitiviteit gelijk voor die groep; als de gecombineerde ECG interpretatie van drie cardiologen als referentie werd gebruikt, waren zowel specificiteit als sensitiviteit hoger. De mate van intra-individuele variabiliteit van een individueel ECG verschaft daarom informatie over de nauwkeurigheid van de automatische interpretatie van dat ECG, vooral wanneer de ECG-interpretatie van een cardioloog als referentie dient. Deze bevinding kan nuttig zijn voor de selectie van ECG's voor handmatige controle in, bijvoorbeeld, epidemiologische studies.

Hoofdstuk 7 bestudeert of de diagnostische interpretatie van het standaard ECG verbeterd kan worden door de interpretaties van zowel de afzonderlijke slagen als van de borstwand elektrode positie simulaties te combineren tot een alternatieve interpretatie. Deze gecombineerde interpretaties zijn vergeleken met de originele interpretatie op een groot bestand van ECG's. In slechts een klein percentage gevallen bleken de interpretaties te verschillen: kleine verschillen kwamen in op zijn hoogst 9% van de gevallen voor, grote verschillen in hoogstens 1%. Uit de groep waarin grote verschillen in interpretatie voorkwamen zijn vervolgens een aantal ECG's aselect gekozen en door een cardioloog geïnterpreteerd. Wanneer zijn interpretatie als referentie werd beschouwd, bleek de gecombineerde interpretatie meer dan twee keer zo vaak correct als de

originele interpretatie.

Hoofdstuk 8 bespreekt de mogelijkheden voor verdere verbetering van de nauwkeurigheid en stabiliteit van ECG analyse programma's gezien ons onderzoek. We concluderen dat de stabiliteit van een programma bepaald kan worden voor een individueel ECG. De combinatie van interpretaties, ofwel afkomstig van verschillende programma's ofwel van methodes zoals die in dit proefschrift, is veelbelovend om de prestaties te verbeteren. Verbetering van de stabiliteit zal moeilijk zijn te verwezenlijken en zal waarschijnlijk aanpassingen vergen in de in het interpretatieproces gebruikte technieken. Het minimaliseren van elektrode positie variabiliteit blijft dan ook voorlopig belangrijk.

Curriculum Vitae

Robertus Johannes Adrianus Schijvenaars was born on January 4, 1966 in Breda. He received his undergraduate education at the Gymnasium Juvenaat Heilig Hart in Bergen op Zoom from 1978 until 1984. He started studying Astronomy at Leiden University in 1984, where he was involved in the automatic determination of the proper motion (angular rate of change in position) of stars in the Scorpio system by comparing photographic plates taken more than 60 years apart.

He switched focus to Computer Science in 1989, in which he received a degree in 1991. His graduation project dealt with color image processing and methods to reduce the number of different colors in an image while preserving human color perception of the image.

In 1991 he joined the Department of Medical Informatics at Erasmus University Rotterdam, where he helped to further develop the Modular ECG Analysis System (MEANS) and to adapt it for use in commercial electrocardiographs. In 1993 he started working on the research described in this thesis.

Bob Schijvenaars continues his research as a scientific staff member of the Department of Medical Informatics at Erasmus University Rotterdam.

Dankwoord

Naast alle wetenschappelijke verhandelingen is er gelukkig ook plaats voor een persoonlijke noot. Wetenschappelijk onderzoek is vaak een kwestie van samenwerking. Zo is ook het onderzoek beschreven in dit proefschrift tot stand gekomen door intensieve samenwerking van een aantal mensen.

Prof. Jan van Bemmel is een ware kapitein én stuurman. Jan, steeds opnieuw heb je me (zonder het misschien te weten) enthousiast weten te krijgen door, niet gehinderd door kennis over de details van een experiment, vrijwel altijd de vinger op de zere plek te leggen en vol enthousiasme nieuwe experimenten te bedenken. Na werkbesprekingen met jou voelde ik me vaak, ondanks de bergen werk die het weer eens had opgeleverd, meer gemotiveerd dan daarvoor.

Jan Kors, het is een plezier met je samen te werken. Je hebt altijd ruimte voor persoonlijke aandacht, naast het werk. Je scherpzinnigheid en kritische blik verbazen me na 9 jaar samenwerking nog steeds. Hoewel je me vrij liet in mijn onderzoek heb je, gelukkig, toch een heel groot stempel gedrukt op dit proefschrift. Met grote zorgvuldigheid heb je me geholpen met het interpreteren van resultaten, het uitdenken van nieuwe experimenten en het schrijven van publicaties. Zelfs een vijfde draft lees je onveranderd kritisch door. Ik hoop nog veel van je te kunnen leren.

Ge van Herpen, onze 'menselijke expert', er zijn denk ik weinig mensen zo veelzijdig als jij. Naast een enorme kundigheid op het gebied van de electrocardiografie heb je ontelbare foute of kromme zinnen in onze publicaties verbeterd. Naast je waardevolle cardiologische inbreng waardeer ik onze gesprekken over andere zaken ook altijd zeer.

Geen onderzoek zonder data, in mijn geval ECG's. Het overgrote deel daarvan heb ik van het Thoraxcentrum Dijkzigt gekregen door de hulp van Ruud Vinke en Ron van Domburg, die met hun no-nonsense mentaliteit alles snel voor elkaar hadden. Hardstikke bedankt voor jullie hulp.

Alle andere mensen op de vakgroep vergroten absoluut het werkplezier en hoewel ze vrijwel niets met het onderzoek te maken hebben gehad hebben ze toch

op een positieve manier bijgedragen. Op dezelfde manier wisten veel mensen in mijn omgeving door hun belangstelling en de zich steeds herhalende vraag "wanneer gebeurt het nou?" mij te motiveren.

Papa en Mama, jullie hebben mij altijd aangemoedigd door te leren. Ook toen het niet zo lekker ging stonden jullie voor me klaar. Het heeft geholpen zoals je ziet. Ik ben ontzettend blij dat ik jullie trots op me kan laten zijn. Ik heb dit proefschrift dan ook mede aan jullie opgedragen.

Marjan en Sofie, woorden zijn eigenlijk niet nodig. Jullie zijn de belangrijkste mensen in mijn leven en hebben me geleerd te genieten van het leven buiten het werk. Ook aan jullie draag ik dit proefschrift op.