GE Healthcare

Marquette™ 12SL™
ECG Analysis Program

Statement of Validation and Accuracy

416791-003 Revision B

NOTE: The information in this manual only applies to the Marquette™ 12SL™ ECG Analysis Program. Due to continuing product innovation, specifications in this manual are subject to change without notice.

Contents

For your notes

Introduction

Document Purpose

This paper reviews the validation and accuracy of GE's Marquette 12SL ECG analysis program. Accuracy levels, as well as the overall clinical impact of the 12SL analysis program, are supplied from independent assessments reported in the scientific literature. For a more complete description of how the 12SL analysis program processes the electrocardiogram (ECG), please see the 12SL physician's guide (PN 416791-004).

Revision History

Each page of the document has the document part number followed by a revision letter at the bottom of the page. This letter identifies the document's update level. The revision history of this document is summarized in the table below.

Table 1. Revision History, PN 416791-003

Revision	Date	Description
A	23 October 2000	Initial release of document.
B	1 February 2007	Document updated to comply with IEC standards.

The Marquette 12SL ECG Analysis Program: A Brief History

The Marquette 12SL analysis program was first developed in 1980. It was the first commercially available ECG program to analyze all 12 leads, simultaneously recorded for 10 seconds. In 1982, the 12SL analysis program was embedded into a computerized electrocardiograph, known as the MAC-II. It was the first of its kind, generating a 12-lead interpretation at the bedside in less than 10 seconds.[1]

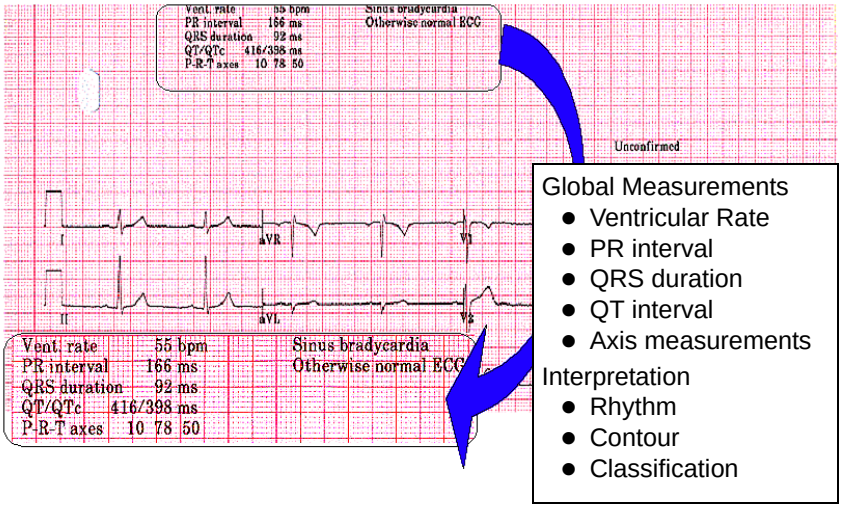
Since its inception, GE Healthcare has continued to evolve the Marquette 12SL analysis program. Furthermore, the Marquette 12SL analysis program has been validated on a variety of platforms beyond the diagnostic electrocardiograph, including bedside monitors, stress-testing systems, pre-hospital defibrillators, Holter recorders, and PC-based systems.

The timeline in Table 2 on page 2 provides is a summary of the significant advancements related to GE's Marquette 12SL analysis program.

Table 2. ECG Analysis / 12SL Timeline	
Year	Advancement
1980	12SL ECG analysis program introduced on the MUSE system[1]
1982	12SL incorporated into a computerized electrocardiograph: MAC-II[1]
1984	12SL Serial Comparison program is introduced on MUSE[2]
1986	Automated testing of 12SL using non-ECG, gold-standard databases[3]
1987	Pediatric analysis, based on Davignon tables, incorporated into 12SL[4]
1988	Analysis of extra leads, generating vector loops at an electrocardiograph
1989	Recognition of ST-elevated acute myocardial infarction (MI) in pre-hospital setting[5]
1991	12SL in a pre-hospital defibrillator equipped with 12-lead ECG[6]
1992	500 samples per second analysis, compression, and storage[7]
1993	12SL in a bedside monitor, equipped with 12-lead ECG[8]
1995	ACI-TIPI integrated into 12SL for prediction of acute cardiac ischemia[9]
1997	Automated QT dispersion and T-wave principal component analysis[10]
1998	ECG Research Workstations for systematic assessment of ECG measurements [11-13]
1999	MAC-RHYTHM: 12SL incorporates asynchronous P wave detector based on QRS subtraction[14]
2000	Gender specific acute MI criteria[15]; Improved pacemaker detection based on 4KHz sampling[16]
2002	12SL in a Holter recorder, equipped with 12 lead ECG[17-19]
2003	New 12SL QT algorithm[20]
2004	Hookup Advisor in 12SL[21]
2005	12SL cleared for measurement and trending of 12-lead ambulatory recordings[22]
2006	Recognition of acute right ventricular infarction via analysis of V4R[23]

Intended Use of GE's Marquette 12SL Analysis Program

GE's Marquette 12SL analysis program assists the physician in interpreting and measuring the resting 12-lead ECG. All computer generated measurements and interpretations should be overread by a physician. All ECG interpretations are identified as being "unconfirmed" until they have been edited by a physician.



12SL Outputs

GE's Marquette 12SL analysis program is intended for use in the general population, ranging from healthy subjects to patients with cardiac and/or non-cardiac abnormalities. The program can select different ECG criteria based on age and gender. The program has ECG criteria intended for all patient ages, including neonatal, pediatric, and adult.

The 12SL analysis program is intended for all clinical care environments that require a resting 12 lead ECG, as prescribed by a physician. This includes all departments within small or large hospitals as well as out-of-hospital environments, such as outpatient clinics, physician offices, ambulances, nursing care facilities, and home-based care.

Population-based research groups also use the 12SL analysis program for generating measurements, since it can improve their effectiveness and consistency.[24-26]

Note that the Food and Drug Administration (FDA) and the International Electrotechnical Commission (IEC) require manufacturers to provide an "intended use" for medical devices.[27] This disclosure is filed with the respective regulatory agency and used for certification and clearance of the medical device. Note that the aforementioned "intended use" for the 12SL analysis program also applies to all devices that use the 12SL analysis program and provide the capabilities of an analyzing electrocardiograph. (See IEC 60601-2-51 clause 50.102.2).

Overall Impact of Computerized ECG: Assisting the Physician

When computerized electrocardiography is used in conjunction with a physician, it can improve both the speed and accuracy of reviewing ECGs, as determined via the following clinical studies:

- "Combined cardiologist and program results demonstrated the highest accuracy, i.e., respectively 78.7% and 76.1%, higher than the result of any individual reader or program. These findings demonstrate that the combination of expert knowledge of computer programs can, similar to panel review and group analysis in clinical practice, enhance diagnostic accuracy."[28]
- "The quality of computer-assisted ECG interpretation was comparable to that of review provided by a cardiology service. Furthermore, computerized interpretation may be clinically more useful because it is immediately available."[29]
- "Computer ECG systems provide a valuable function for ECG analysis, storage, retrieval, and serial comparison. The current systems can provide quality control of technician performance, acquisition equipment, and physician over reading. Its overall acceptability and clinical usefulness is documented in a clinical practice setting with a 90.4% computer-physician agreement in more than 20,000 ECGs. Computerized ECG systems have demonstrated their clinical usefulness in patient care."[30]
- "The impact of computer assisted interpretation on cardiologists' readings of ECGs is demonstrably beneficial: the main empirical conclusion of this study is that, compared with conventional interpretation, the use of computer assisted interpretation of ECGs cuts physician time by an average of 28% and significantly improves the concordance of the physician's interpretation with the expert benchmark, without increasing the false-positive rate."[31]
- "In summary, this study has confirmed that junior doctors have a high error rate in reporting ECGs. Computer generated reports did not significantly improve this, even though the machine achieved a low major error rate compared with the junior doctors. Computer generated reports may have a role in prompting junior doctors to query their own ECG interpretation but should not replace experienced medical support."[32]

Despite the documented benefits of a computerized ECG system, it should be made clear that a computerized analysis is not a substitute for human interpretation. There are two reasons for this:

- First, statements of accuracy need to be viewed from a statistical perspective. Although accuracy levels may be high, outliers can and will exist. The computer will make mistakes, especially in the presence of artifact. As one author cautioned: "Computer decision support systems can generally improve the interpretive accuracy of internal medicine residents in reading EKGs. However, subjects were influenced significantly by incorrect advice, which tempers the overall usefulness of computer-generated advice."[33]

- Second, a computer does not have the ability to include the entire clinical picture of the patient. The ECG tracing is significant only when interpreted in conjunction with the other clinical findings associated with the patient. As quoted in the literature: “Given that computers alone cannot perform the task of cardiovascular diagnosis, and that cardiologists' ECG interpretations are greatly enhanced by ubiquitous computer assisted interpretation, it appears that the best approach is one that combines person and machine.”[31]

Development and Validation Process of the Program

GE's Marquette 12SL analysis program was introduced in 1980. All improvements to the program have been accomplished via a systematic, logical, controlled methodology. A major aspect of this methodology benefits from the use of stored ECGs.

Reanalysis of Stored ECGs

All historical ECGs analyzed by the 12SL analysis program and stored on the MUSE cardiology information system, can be reanalyzed for the purposes of validating or improving the program.[3] This is because the median QRS complex generated by the program has always been compressed and stored via a lossless Huffman encoding method.[1, 34] The first implementation of this methodology has been described in the literature,[35] was later enhanced by GE Healthcare for ECGs stored at 500 samples per second (sps),[7] and ultimately served as the basis of a new international standard.[36] This standard includes data fidelity requirements for compressed ECGs; these requirements are surpassed by the data compression/decompression methods currently employed by GE. For those who desire additional fidelity in the decompressed ECG, GE Healthcare provides another option (known as Digital View Storage DVS), which uses lossless compression on all of the raw ECG data.

Initiating a Change in the Program

Any change to the program requires a great deal of research. This effort can be instigated by a variety of sources:

- The constant pursuit of clinically correlated databases can yield statistics that indicate whether a change should be considered.
- New criteria published in the scientific literature can be evaluated and sometimes incorporated into the program.
- Consultations with cardiologists also stimulate investigations. This is especially true when they have stored ECGs that reveal a particular error.
- GE Healthcare also documents customer complaints. Although complaints are typically documented from customer interactions with GE Service, Sales, or the Call Center personnel, any GE employee who is aware of a complaint must document it. The Engineering department then tracks these complaints. Some times the complaint can be resolved by providing the customer further documentation or clarification as to how the program functions.

Measuring the Impact: Evaluation via a Library of Databases

Before a change can be instituted, it must always be evaluated in relation to the current program performance. Stored ECGs are reanalyzed and any difference due to the enhancement is scored and tracked. After this is done, the validation system automatically culls out any ECGs that scored differently between the two versions of the program. This results in an efficient method to automatically determine how a change might affect program performance.[3, 11]

An Appropriate Gold-Standard Database for Type A, B, or C Statements

In the 12SL physician's guide, each 12SL interpretive statement has been identified as either Type A, B, or C, a classification methodology approved at the Tenth Bethesda Conference on Optimal Electrocardiography.[37]

Type A statements refer to the diagnosis of anatomic lesion or patho physiologic state, such as myocardial infarction or hypertrophy. The accuracy of these statements can be determined by non-ECG evidence such as cardiac catheterization (CATH), echocardiography (ECHO), cardiac enzymes, clinical outcome, etc. These statements are evaluated with databases that have been clinically correlated with non-ECG data. The non-ECG data acts as the “gold standard”.

Type B statements cover statements referring to the diagnosis of electrophysiological changes and are therefore detected primarily by the ECG itself. This includes arrhythmias and conduction disturbances. Although intracardiac recording can be used to validate the diagnostic conclusions determined via the surface ECG, this is often not practical. As a result, a database of ECGs with the physician’s interpretation is used as the reference.

Type C statements refer to purely descriptive ECG features that usually cannot be documented by any other means. Examples of such statements include “non-specific ST-T abnormality” and “left axis deviation”. Again, a database of ECGs with the physician’s interpretation is used as the reference.

Type A Statements: Reliance on Non-ECG Correlates is Not Enough

Databases that have been correlated with non-ECG data are critical for the development and validation of Type A statements. But these databases have their limitations. Reasons include the following:

- The use of a particular “gold standard”, non-ECG correlate may force the database to contain a population that is not representative of the disease in the actual clinical setting. For example, an autopsy-proven myocardial infarction (MI) database may not be indicative of what a typical MI looks like, since many patients survive an MI. Another example would be “CATH proven normals”. In this case, the patient often receives the CATH because they were symptomatic or the ECG was “abnormal”. As a result, the ECGs from such a database may actually not be from true “normal” patients.
- Databases from most published clinical investigations have already removed the “confounding influence” of ECGs with conduction defects, etc. However, this is not the case in the real world. The algorithm must operate in the presence of ischemia, conduction defects, drug effects, etc.

- A non-ECG value may indicate the presence of an abnormality but this does not mean that the abnormality is revealed in the surface ECG. For example, an ECG can often appear “normal” even when it is clearly established that it is from a patient with an acute myocardial infarction. It is important to not force the program to identify these ECGs as positive, if the abnormality is not revealed in the signal. Otherwise, the program will overcall the abnormality in other environments.
- The database may only contain the extreme cases of normal versus abnormal. Algorithms do not operate in a black and white world.
- And finally, non-ECG data cannot be considered perfect: every test comes with its own inherent level of inaccuracy.

Thus, even when an abnormality can only be positively determined via a non-ECG correlate, a physician’s interpretation is critical as an additional check. Therefore, during development and testing, databases based on a physician’s interpretation are used in conjunction with databases that have been correlated with non-ECG data.

As an additional check, GE Healthcare uses large databases that have been gathered as part of routine care. In this case, there may be little quality control of the physician interpretation. Nevertheless, these large databases, available via a MUSE system, are useful for determining the rate at which a change in the program will generate a change in an interpretation across an entire institution. Reanalysis on over 100,000 ECGs can be done in a matter of minutes and it confronts the algorithm with multiple kinds of waveforms and varying degrees of abnormality. ECGs that changed their analysis can be further investigated with either confirmation from medical records and/or another expert opinion.

Training versus Test / Validation Sets

Different databases are used for development versus validation. This precludes us from overtraining an algorithm so that it works beautifully on the training set but cannot be applied, with the same success, to other populations. This is an important requirement for reliable pattern recognition.[38] In this document, all reported results for interpretation performance are from independent validation sets.

Porting 12SL to Multiple Platforms: Verification Process

GE’s Marquette 12SL analysis program has been implemented on a variety of platforms, including Holter recorders and pre-hospital defibrillators. In order to accomplish this, the program must be completely tested in its target environment. The use of analog ECGs to test every logic path in the target environment is not feasible. Thousands of ECGs would have to be recorded and the results manually compared. A digital solution is required. GE Healthcare invented a program for this purpose, known as EZSIM.

EZSIM Background

EZSIM is a program that generates simulated ECGs with the intent of thoroughly exercising the 12SL analysis program. After 12SL processes an ECG made by EZSIM, a checksum is computed across the complete IO12SL data structure (this data structure contains all inputs, the complete analysis output of 12SL, and many intermediate results that never get displayed on a report). Checksum mismatches indicate that 12SL produced a different output than expected on the target platform. A target implementation is only considered successful when over 70,000 ECGs have been analyzed by the target platform without any differences detected in the checksums.

EZSIM: ECGs with Variety of Shapes and Rhythms

EZSIM simulates ECGs with a vast variety of shapes and rhythms, covering all categories identified by the program. Each ECG is generated algorithmically and can run as long as samples are drawn from the simulator. ECGs are not restricted to 10 seconds or even 24 hours.

The simulator has two parts: the initialization routine and the running routine. The initialization routine uses about 109 random numbers to create a basic P wave pattern, a basic QRS pattern, a basic PVC pattern, a basic PP interval, an amount of PP variability, a basic PR interval, an amount and frequency of muscle tremor noise and an amount and frequency of baseline sway noise. The running routine uses up to 4 random numbers per sample to determine the noise, 3 random numbers per QRS or unconducted P-wave to determine when the next P-wave, QRS, or PVC will occur.

The simulator is able to overlap one QRS cycle with the next so that the P-waves at higher heart rates can creep into the T-wave of the previous cycle.

Although constructed using random numbers, these ECGs are exactly reproducible given a starting point in the random number sequence. That starting point is called the random number seed. That seed is all that is needed to reconstruct that ECG of unlimited length.

Any number can be used as the random number generator seed. All the numbers from 0 to 65535 produce different sequences of random numbers and therefore different ECGs. The simulator algorithm is the equivalent of a database but as opposed to conventional databases that retrieve stored ECGs, this database requires only about 3 kilobytes of code and no storage for the actual ECGs.

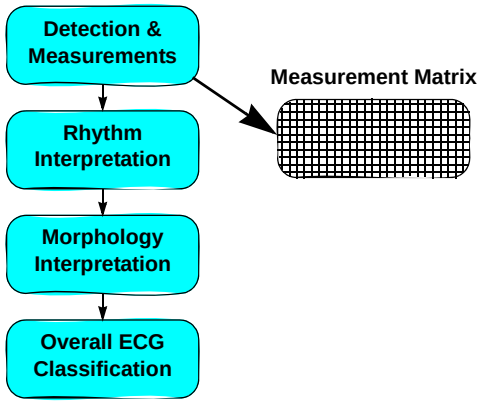
Some of the Rhythm ECG Features Supported by EZSIM

- unconducted P-waves
- modulated coupling intervals, P-P
- random occurrence of ectopy, blocked AV conduction
- dual synthesis of patterns allows overlap, P onto T, or R onto T
- atrial fibrillation - irregular with fibrillatory waves
- atrial flutter - fast, less irregularity, no fibrillatory waves
- ventricular tachycardia
- torsades, ventricular pattern is rotated gradually
- ventricular fibrillation
- muscle tremor noise, electrode motion noise, baseline sway

12SL Analysis Program Structure: Measurements Before Interpretation

Below is a simple block diagram of GE's Marquette 12SL analysis program. Note that all the interpretative statements are generated following the measurement portion of the program.

All measurements generated by the program are stored in a measurement matrix, which are then later accessed by the interpretive portions of the program. Criteria used by the program are fully described in the 12SL physician's guide. Note that these criteria never directly measure the ECG. Rather, the criteria use only the values from the measurement matrix. For any given ECG, the measurement matrix can be printed at the interpreting electrocardiograph or MUSE system.



12SL Block Diagram

Detection and Measurement

Since the interpretive portions of the program are based on measurements, it is critical that the ECG measurements be as robust and as accurate as possible.[39] The following sections address the necessary elements for generating quality measurements, with associated references to substantiate this quality.

The Digital ECG: Data Content and Fidelity

In addition to resting electrocardiographs, the 12SL analysis program operates in a variety of products, from bedside monitors to pre-hospital defibrillators. As a result, the 12SL analysis program has been designed to be configurable for different environments.

All 12 leads, simultaneously recorded for 10 seconds, is the minimum data set required by GE's Marquette 12SL analysis program (specifically leads I,II and V1-V6; leads III, aVR, aVL, and aVF are calculated via Einthoven's law). In some applications, the 12SL analysis program analyzes more than 10 seconds or more than 12 leads.

In 1979, GE introduced simultaneous recording so that the computer could use all signals from all 12 leads to properly detect and classify each QRS complex. The Common Standards for Electrocardiography independently verified the advantage of this technique:

> "Conclusion: The simultaneous recording and analysis of all 12 standard leads ... is certainly an improvement over the conventional recording of three leads at a time. Similarly ... multi-lead programs proved to be more stable than those obtained by conventional programs analyzing three leads at a time ..."[40]

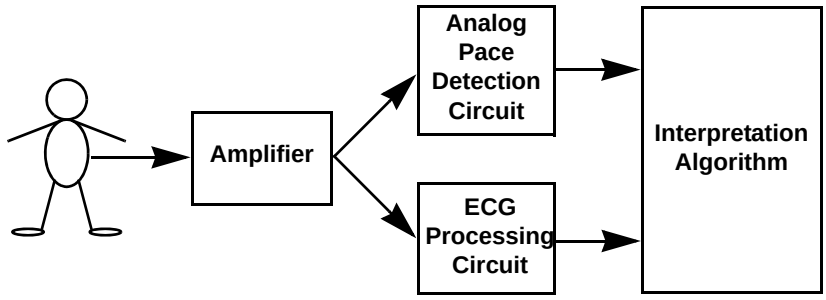
All resting electrocardiographs currently sold by GE analyze the waveform at 500 samples per second (sps). In some GE resting electrocardiographs, the ECG is sampled at a much higher rate, such as 4,000 sps. This is referred to as over-sampling and it used by the device to generate an average, cleaner signal at 500 sps. Specifications for electrocardiographs, across the industry, often cite the raw sample rate (e.g. 4K sps) without clarifying that the ECG analysis and measurement software actually executes on data with a lower sample rate. Current guidelines for resting ECG analysis cite 500 sps,[41] which is the sample rate executed by 12SL in a resting electrocardiograph.

Before the physiological data is sampled, analog filtering is applied. These filters attenuate high-frequency electrical noise that is not part of the physiological signal. If these analog filters were not present in the device, high-frequency signals could be digitized by the device and appear as low frequency noise, inter-mixed with the physiological cardiac signal. To eliminate this possible source of contamination, GE applies an analog filter, known as an anti-aliasing filter. See the 12SL physician's guide (PN 416791-004) for further discussion on anti-alias filters.

Specialized Hardware and Software Algorithm for Cardiac Pacemaker Detection

Recent advances in pacemaker technology have adversely affected the accuracy of pacemaker detection by computerized ECG analysis systems, particularly those relying solely on detecting pacemaker pulses from the digitized surface ECG. Improvements in pulse generators and lead design and the increasing use of bipolar pacing have lead to the reduction of pulse amplitudes and widths observable on the digitized surface ECG. Consequently, paced ECGs are often misinterpreted. In addition, today's pacemakers offer a wide range of operating modes and programmability options that place limitations on the inferences that can be made regarding the timing relations of pacemaker pulses.

To overcome these limitations, it is necessary to reliably detect pacemaker pulses prior to digitization of the ECG and to efficiently relay that information to the interpretation software. Therefore, in addition to sampling the physiological signal, all GE resting electrocardiographs equipped with an active, digitizing patient cable have a parallel channel specifically designed for sampling pacemaker activity. See representative schematic below.



This parallel channel measures signals with a center frequency on the order of 2KHz, as opposed to the frequencies inherent in the physiological cardiac signal that are below 250Hz.[42] In addition to the hardware detection circuit, a software algorithm contextually analyzes these high-frequency detections in relation to the physiologically recorded signal.[16] This capability reduces falsely detected pulses in high frequency noise situations. Accuracy for the detection of pacemakers is covered under the rhythm interpretation section of this document.

Signal Conditioning

It has been shown that both the physician and the 12SL analysis program are prone to make more ECG interpretation errors when presented poor-quality tracings.[43] As opposed to other applications of the ECG (like Holter or Stress), where the skin is aggressively prepared before the test, the resting ECG is typically done with little or no preparation of the skin. Therefore, it is in the best interest of the overreading physician to insist that ECGs be taken with good-quality electrodes and that the patient be kept supine, calm, and warm during the procedure in order to minimize artifacts. Nevertheless digital filters can be applied to the ECG to improve ECG quality. This process is often referred to as signal conditioning.

The 12SL analysis program automatically removes A/C interference by generating a model of the interference and then subtracting it from the raw

waveform.[44] It has been demonstrated that the 50/60Hz-learning filter in the 12SL analysis program can easily remove over 1mV of 50/60Hz noise, without distortion of the physiological signal. In some developing countries, there is no power grid. As a result, A/C interference is not locked to a specific 50 or 60Hz frequency. For the removal of AC interference in these environments, GE has developed a “Hunting Filter”.[45]

In order to remove baseline sway, the 12SL analysis program employs a high-pass filter that has a linear phase response.[46] It has been known for many years that ST segments can be faithfully reproduced with a higher filter setting if the phase response is linear.[42] This recognition led to new recommendations in the AAMI standards (EC11 1991) allowing the filter setting to be up to 0.67 Hz with an additional test to be sure the ST segment is not distorted.[41] It has been demonstrated that the 12SL program can high-pass filter the ECG at 0.32Hz without distortion of low frequency components of the ECG, such as the ST or T wave.[22]

Other sources of noise in the ECG include muscle tremor or electrode-motion artifact. Most electrocardiographs have various low-pass filter settings, including 40Hz, 100Hz, or 150Hz. The lower the filter setting, the more aggressively the filter removes high frequency signals, which include muscle tremor or electrode-motion artifact. However, these low-pass filters also operate on the entire ECG signal and attenuate all high frequency elements of the ECG signal, such as the QRS complex and pacemaker artifacts. Therefore, in order to consistently measure the resting ECG and capture the proper QRS amplitude, the 12SL program always analyzes the ECG at the AHA / AAMI recommended full bandwidth of 150Hz,[41, 42] regardless of the low-pass filter setting. As a result, these settings are sometimes referred to as “writer settings”, since they do not affect the ECG interpretation.

It should be noted that all filter settings travel with the ECG. That is, the MUSE system can be configured to either portray the ECG signal as it was acquired at the electrocardiograph or at another specified filter setting. Note that over-reliance on aggressive, low-pass filtering implies that the 12SL program is subjected to more high-frequency noise than the physician sees in a filtered ECG tracing.

Detection and Measurement of Signal Quality

ECG devices often measure the impedance across the skin-electrode interface. When this impedance exceeds 600K ohms, a GE resting electrocardiograph informs the user that a lead is off and provides no signal for that particular lead. The reason the device no longer provides a signal for a “lead-off” condition is because a dangling lead would result in extreme noise, obscuring the rest of the ECG report and making it difficult for both the analysis program and the human to interpret the ECG.

Throughout the ECG industry, impedance across the electrode-skin interface is often used as a surrogate for lead quality. However, normal skin impedance, especially without any skin preparation, can vary dramatically, from 10 to 300K ohms.[47]

Furthermore, it can be readily demonstrated that good quality resting ECGs can be obtained throughout this range, since the large impedance values exist due to the nature of the patient's skin rather than the electrode-skin interface. Stating poor signal quality below 300K ohms simply results in false-positive calls and great frustration upon the person taking the ECG. Furthermore, it has been shown that skin impedance has a poor correlation with artifacts associated with poor electrodes or a poor electrode-skin interface.[48]

As a result, GE Healthcare has adopted an alternative approach for detecting signal quality, which directly analyzes the ECG signal for muscle tremor, AC power interference, electrode motion, or baseline shifts. This software algorithm for detecting these artifacts has previously been described and is referred to as the Hookup Advisor.[21]

The Hookup Advisor assigns an ECG lead quality level of green, yellow, or red, which is also indicated on the user interface of the electrocardiograph. This was tested on a large database of over 120,000 ECGs. Lead quality distributions and rhythm interpretation discordance rates between the physician and GE's Marquette 12SL analysis program are reported in Table 3.

Table 3. Lead quality and rhythm discordance for combined test set (N = 120,698)[21]

Lead quality	N	Percent of total	Discordance rate
Green	115128	95.39%	3.9%
Yellow	5170	4.28%	7.4%
Red	400	0.33%	12.1%

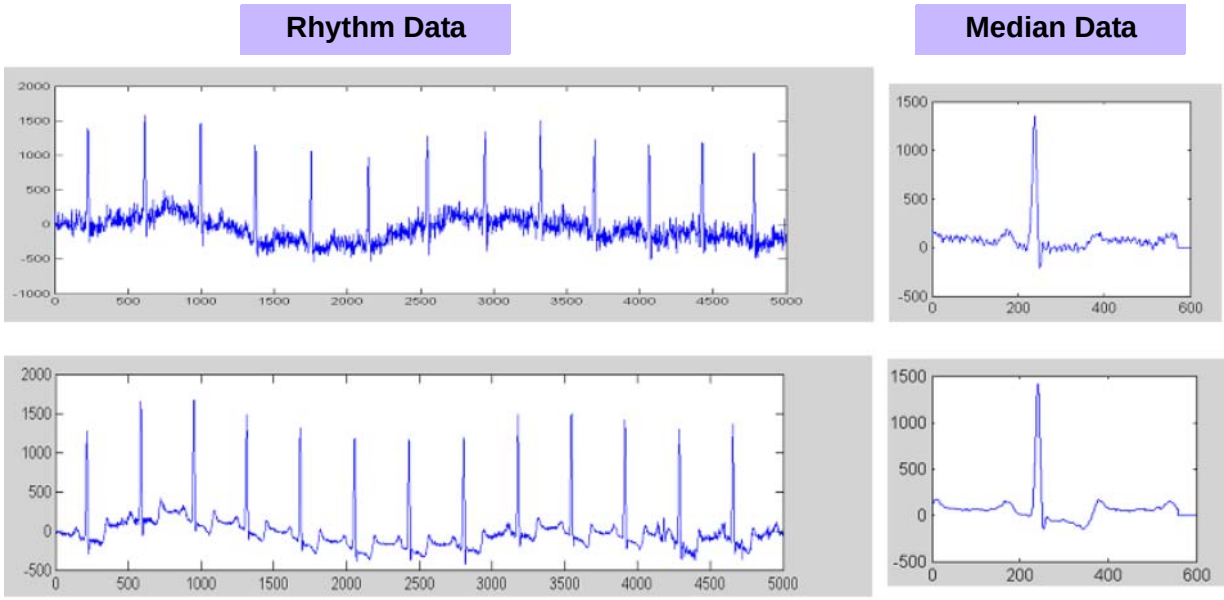
Overall, 95.4% of all ECGs were categorized as green (good) lead quality, 4.3% were assessed as yellow (marginal) lead quality, and 0.3% as red (poor) lead quality. As the primary rhythm from the 12SL reanalysis was compared to the primary rhythm in the confirmed ECG, the discordance of these two interpretations increased sharply, from 3.9% to 7.4% to 12.1% as the lead quality degraded from green to yellow to red.

Lead quality indicators can be stored on the MUSE system and used to monitor and continuously improve the quality of ECG acquisition across an institution.

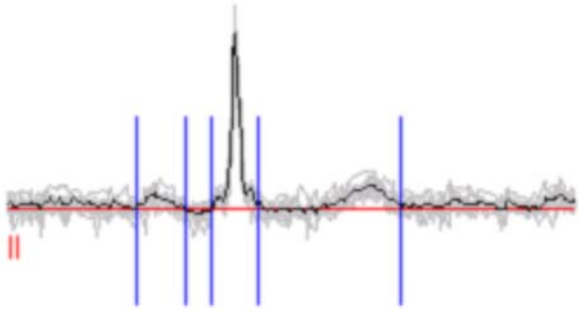
Median Beat/Signal Averaging

In addition to filtering or signal conditioning, there is another method that is employed to eliminate noise from the cardiac cycle: that is, signal averaging. Instead of analyzing a single, raw QRS complex, the GE's Marquette 12SL analysis program generates a median complex. In other words, all QRSs of the same shape are aligned in time. Next, the algorithm generates a representative QRS complex from the median voltages that are found at each successive sample time. Although more complicated than creating an average, the method results in a cleaner signal than an average.

The following figure is an example of the formation of a median from a 12-lead Holter recording.[22]



Presented below is even a closer look at the median. It shows the median complex displayed along with the raw complexes used to form the median complex. Note the noise in the raw signal versus the median complex.



Willems et. al.,[49] independently verified the value of this technique. Without the technique, onsets and offsets were shifted outward in the presence of noise. As quoted from the literature: “Increasing levels of high-frequency noise shifted the onsets and offsets of most programs outward. Programs analyzing an averaged beat showed significantly less variability than programs, which measure every complex or a selected beat. On the basis of the findings of the present study, a measurement strategy based on selective averaging is recommended for diagnostic ECG computer programs.”

Results by Zywietz[50] also showed that programs analyzing an averaged beat exhibited less variability than programs that measure every complex or a selected beat. Subsequently, Zywietz also confirmed that median beats had less noise and generated more accurate measurements than an analysis of raw beats.[51]

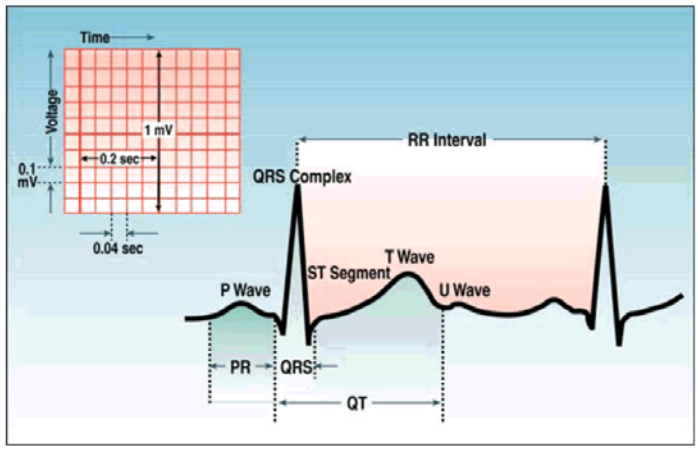
Farrell[52] also demonstrated the effectiveness of the median by testing 12SL on 90,000 “noisy” ECGs. This test used a repeatable methodology for the creation of “noisy” ECGs, which can be applied for industry-wide assessment of robustness of computerized measurements.

QRS Onset / Offset and Determination of Global Intervals

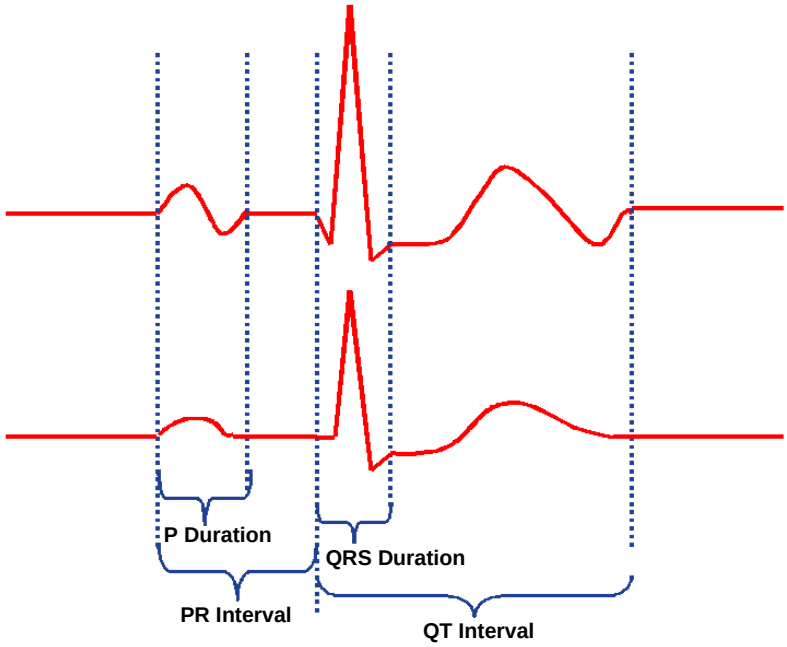
Good ECG measurements depend upon the proper identification of the fiducial points such as QRS onset and offset. Consistent with the signal-processing portion of the program as well as the physiological definitions for cardiac depolarization and repolarization, these fiducial points are determined by an analysis of the slopes in all 12 simultaneous leads. As a result, each fiducial point refers to the same sample-time in all of the time-aligned median complexes. Since these fiducial points are applied across all 12 median complexes, they are often referred to as “global” versus “lead-specific”.

P onset and P offset are also determined via the median complexes, unless the computer detects asynchronous P wave activity or an inconsistent PR coupling interval in the rhythm data. In this case, P onset and P offset remain undefined.

As opposed to the human reader, which may only inspect the QRS duration in any single lead of the ECG, the computer measures the QRS duration as a global interval. That is, it measures the QRS duration from the earliest detection of depolarization in any lead (QRS onset) to the latest detection of depolarization in any lead (QRS offset). Similarly, the QT interval is measured as global interval: that is, from the earliest detection of depolarization in any lead (QRS onset) to the latest detection of repolarization in any lead (T offset). See the following diagrams.



Basic ECG Nomenclature



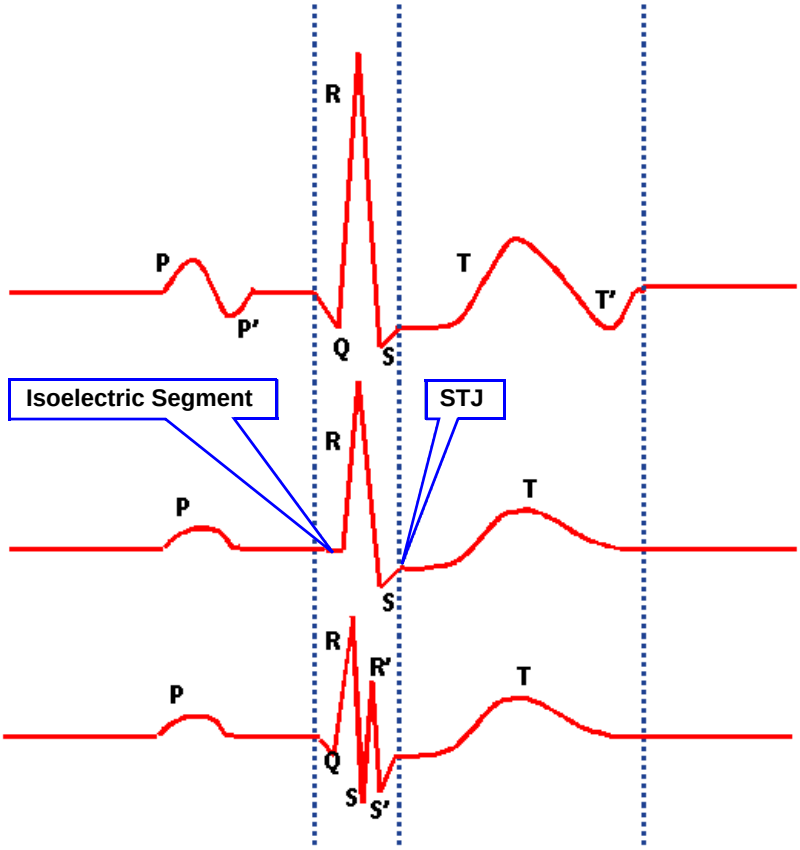
Global Fiducial Points - Across All Median Complexes

Definition and Measurement of Waves

After the global fiducial points (P onset/offset, QRS onset/offset and T offset) have been determined, the waves within each complex are measured according to published standards.[53] This is done separately for each lead. Different ECG analysis programs treat waves within the QRS complex in different ways; as a result, the IEC standard requires that this wave identification process be fully disclosed, as provided below. (See IEC 60601-2-51 clauses 50.101.2-4).[27]

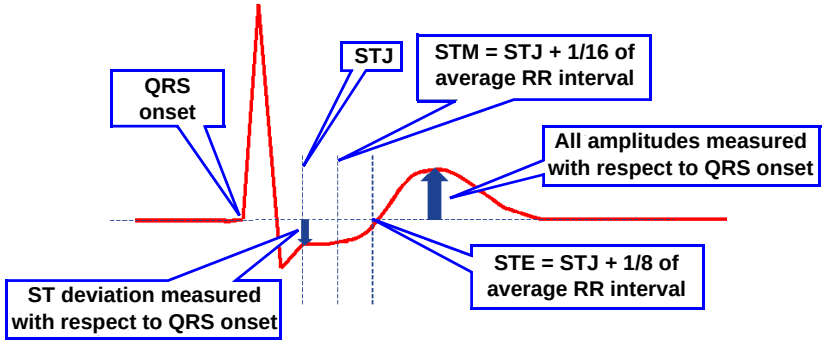
Starting at QRS onset, the program finds the points at which the ECG signal crosses the baseline within each complex. If the crossing points define a wave that has an area greater than or equal to 160 µV-ms, the wave is considered to be significant. If the area is less than this value, the program considers the wave to be insignificant, and it will not label it as a separate wave. Sections of the complex that do not exceed the minimum wave criteria of 160 µV-ms are combined with the adjacent significant wave.

Since the wave of depolarization is a spatial entity, the onset of the wave will not be evident in all leads at the same time. Isoelectric sections starting at QRS onset of the complex are treated as part of the subsequent significant wave. Likewise, isoelectric sections at the end of the QRS will be incorporated into the preceding significant wave.



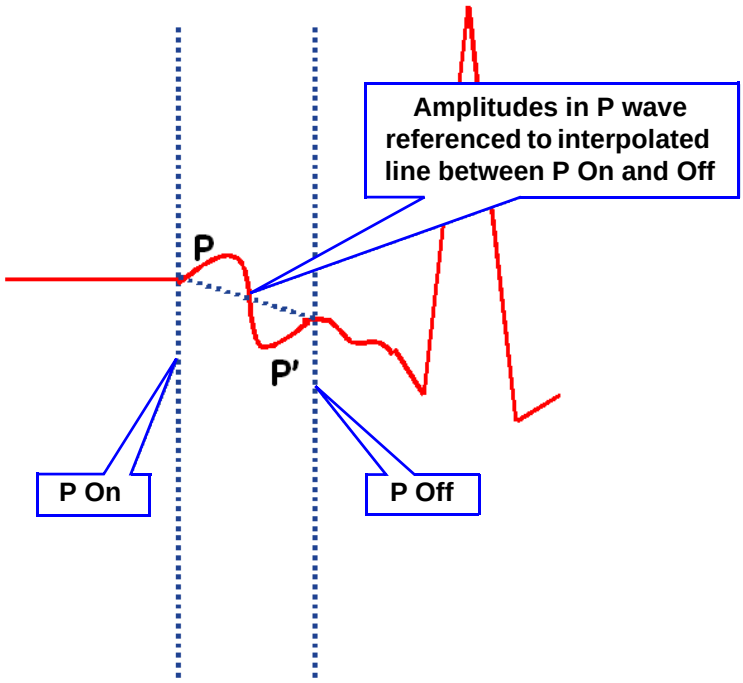
Definition of Waves Within Complex

Amplitudes of significant waves within the QRS as well as the T wave are measured with respect QRS onset. Deviation of the ST segment is also measured in relation to QRS onset. STJ is defined as QRS offset. Further definition of the ST segment is defined by STM and STE, which are two additional points along the ST segment that are 1/16 and 1/8 of the average RR-interval from STJ. See figure below.



Amplitudes of QRS and ST-T Measured in Relation to QRS Onset

Amplitudes of significant waves within the P wave are measured with respect to a baseline level that is interpolated from P onset to P offset. This accommodates the phenomena of PR segment depression. See diagram below.



These amplitudes and durations result in a measurement matrix containing more than 800 values. Measurements are then passed onto the criteria portion of the program so that it can generate an interpretation.

Measurement Accuracy: Reported Results

Common Standards for Electrocardiography (CSE) Database

In an effort to standardize and evaluate the performance of ECG computer measurement programs, a 12-lead ECG reference database was developed.[40] Typically referred to as the Common Standards for Electrocardiography (CSE) database,[54] it contains a set of 250 electrocardiograms (ECGs), with selected abnormalities, which were measured by five cardiologists. Attention was focused on the exact determination of the onsets and offsets of P, QRS, and T waves. As quoted from the literature:

> "The cardiologists performed their task on highly amplified, selected complexes from the library in a two round process. With use of a modified Delphi approach, individual outlying point estimates were eliminated in four successive rounds. In this way final referee estimates were obtained that proved to be highly reproducible and precise."[55]

All ECG waveforms in the CSE database are available to the industry. However, only one-half of these ECGs contain the measurements from the CSE referee committee. The other half does not contain these manual measurements. In other words, one-half has published measurements; the other half has unpublished referee measurements. As a result, the ECGs that contain the published referee measurements can be used by the industry for the self-assessment and reporting of measurement performance. The other 125 ECGs are unavailable for self-assessment.

Independent Evaluation Using CSE Database

The Marquette 12SL analysis program was tested using all the CSE ECGs (that is, including those without the published CSE measurements). This independent evaluation was done when the program only operated on data sampled at 250 sps. The data in the CSE database was originally acquired at 500 sps. In order to re-analyze this data at 250 sps, the ECG was down-sampled to generate data at 250 sps. The results of this independent evaluation are presented in Table 4, including the mean difference from the manual measurements and the standard deviation of the mean difference.

Table 4. Complete CSE Database Evaluation, Including Unpublished Referee Annotations[40]

Interval Measurement	N	Mean difference (ms)	Standard Deviation (ms)
P duration	218	-0.4	9.0
PR interval	218	-0.6	5.8
QRS duration	240	-0.6	5.4
QT interval	238	0.9	12.2

IEC 60601-2-51/ Reporting of Measurement Performance via CSE Database

The International Electrotechnical Commission (IEC) has issued particular requirements for recording and analyzing electrocardiographs (see 60601-2-51(c) IEC 2003)[27]. For measurement performance assessment and acceptance testing, the standard uses ECGs from the CSE database that contain the published referee measurements. As a result, this is a self-assessment, self-reporting measurement performance test.

In addition to biological ECGs, the CSE database contains analytical and calibration ECGs. These are used to evaluate the accuracy of the global interval measurements and the accuracy of amplitude and wave duration measurements within each complex of each lead. GE's Marquette 12SL analysis program has been evaluated with these analytical and calibration ECGs. With regards to amplitude measurements, no ECGs were excluded due to fiducial point errors; the program passed all of the amplitude measurement requirements as defined in IEC 60601-2-51 clause 50.101.2. With regards to global interval and wave duration measurements, one ECG was excluded from QRS duration and the S duration measurements due to a QRS offset fiducial point error. All global interval measurements were within acceptable limits. For the per-lead measurements all results are reported below. No exclusions were made. All per-lead measurements were within the acceptable limits as required in IEC 60601-2-51 clause 50.101.3.1.

Table 5. Results of Absolute Interval and Wave Duration Measurements for IEC

Measurement	Mean difference (msec)	Standard deviation (msec)	Acceptable mean difference (msec)	Acceptable standard deviation (msec)	Pass / Fail
P duration	-8.6	1.5	±10	8	Pass
PR interval	-6.0	1.6	±10	8	Pass
QRS duration	0.0	1.6	±6	5	Pass
QT interval	1.4	3.8	±12	10	Pass
Q duration	-0.8	2.8	±6	5	Pass
R duration	-0.7	2.2	±6	5	Pass
S duration	-0.9	2.7	±6	5	Pass

In addition to the calibration ECGs, the IEC requires testing on 100 biological ECGs from the 125 ECGs that contain the CSE measurements. In the performance reporting of the 100 ECGs, the IEC standard allows exclusion of up to four measurements with "obvious fiducial point errors". No obvious fiducial point errors were observed via GE's Marquette 12SL analysis program. Thus no ECGs were excluded for this reason. The standard then allows exclusion of the "four largest deviations from the mean (outliers) for each measurement". As a result, the following table contains the global interval results for 96 ECGs, analyzed at 500 sps. Included in the table are the mean difference from the CSE manual measurements, the standard deviation of the mean difference, and the IEC pass / fail criteria. The global interval measurements are well within accepted limits and pass the test. (See IEC 60601-2-51 clause 50.101.3.2).

Table 6. Global Measurement Performance for IEC standard on 96 CSE Biological ECGs					
Measurement	Mean difference (msec)	Standard deviation (msec)	Acceptable mean difference (msec)	Acceptable standard deviation (msec)	Pass / Fail
P duration	-6.7	9.0	±10	15	Pass
PR interval	-1.5	5.5	±10	10	Pass
QRS duration	-5.2	5.2	±10	10	Pass
QT interval	+1.0	8.9	±25	30	Pass

Another test includes only 10 ECGs from the CSE database that contains the published referee measurements. These 10 ECGs were analyzed by the 12SL analysis program, first without noise added and then with each of the noise types specified: 25 µV RMS high frequency muscle artifact noise, 50 µV peak-to-valley 60 Hz line frequency noise, and 1 mV peak-to-valley 0.3 Hz sinusoidal baseline noise.

For each noise type, the interval measurements were recorded and compared against the measurements of the noise-free ECGs. For each of the interval measurements of each noise type, the mean of the ten differences from the noise-free measurements was calculated. As specified by the IEC standard, two of the largest deviations from the mean were excluded from the final reported mean and standard deviation of the differences. (See IEC 60601-2-51 clause 50.101.4).

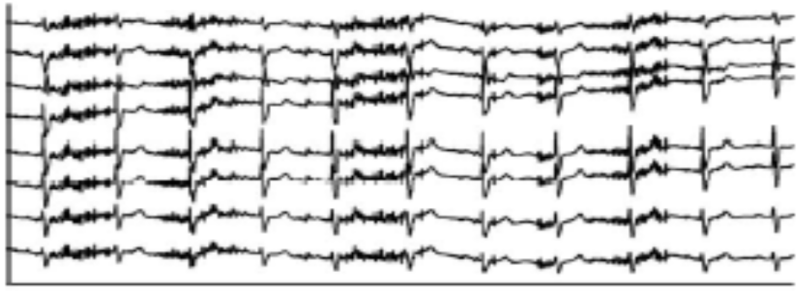
Table 7. IEC 60601-2-51, Clause 50.101.4 Mean Difference From Recordings Without Noise			
Global Measurement	Type of Added Noise	Mean Difference (ms)	Standard Deviation (ms)
P duration	high frequency	-43.5	9.9
	line frequency	-2.8	6.7
	baseline	1.5	3.7
PR interval	high frequency	-18.5	11.0
	line frequency	-1.5	2.8
	baseline	0.3	1.3
QRS duration	high frequency	-7.8	2.7
	line frequency	-1.3	4.7
	baseline	-0.3	1.7
QT interval	high frequency	-1.3	3.2
	line frequency	1.5	3.7
	baseline	-0.3	3.5

Interval Measurement Noise Immunity: Evaluation with MIT-NST & CSE Database

The 125 ECGs of the CSE (containing the published referee measurements) were merged with records from the MIT Noise Stress Test database (MIT-NST).[56] For each CSE ECG, 720 unique noise ECGs were created, for a total of 90,000 noisy ECGs. Computerized measurements from the noisy ECGs were compared to the original ECG measurements. The repeatability of the measurements was assessed as a function of a lead quality score.

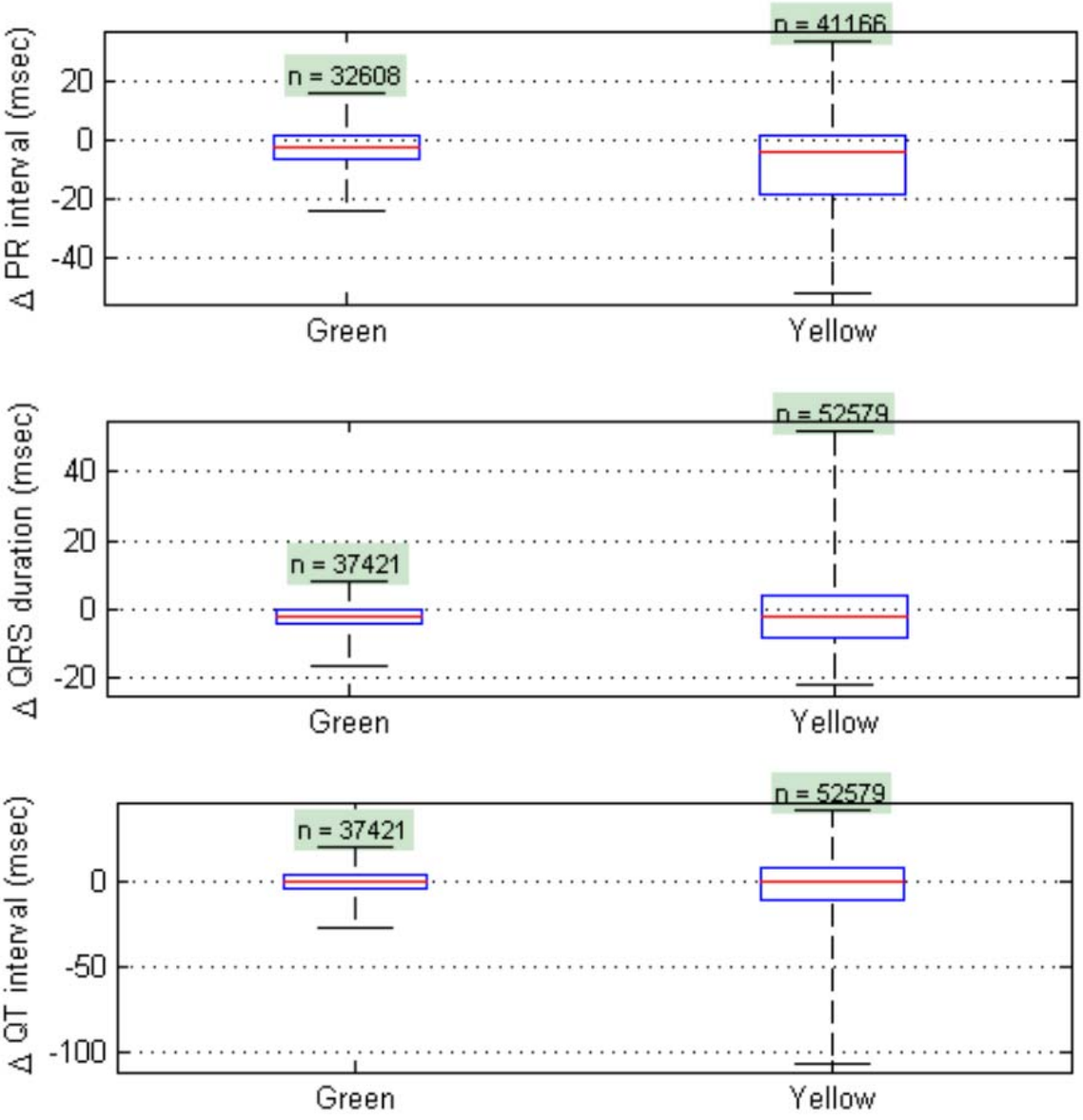
The repeatability of the measurements was found to be in excellent agreement with the original ECG measurements when the noise level was no worse than that of the original ECGs. Noise did not introduce any bias to the measurements, although not surprisingly, the variation of the errors increased as the lead quality degraded.[52]

An example of an ECG generated by the combination of the CSE and MIT-NST databases in shown below. The MIT-NST database consists of three 30-minute 2-channel noise records and is specified for the analysis of the robustness of ambulatory ECG analysis by the AAMI standard EC38.[57] The noise recordings were made using physically active volunteers and standard ECG recorders, leads, and electrodes; the electrodes were placed on the limbs in positions in which the subjects' cardiac generated signal was not visible.



Example of CSE ECG combined with MIT-NST Record

For each ECG, interval measurement differences versus the CSE annotations were obtained. These differences were grouped against the Hookup Advisor indicators[21] and the ranges of the values reported in the following figure.[52] The reported PR interval tended to shorten as the noise level increased. The mean difference of the QRS duration was relatively unaffected by noise, changing by less than 2ms. Likewise, the median difference of the QT interval was 0 ms for both lead quality levels, while the standard deviation (SD) of the QT differences went from 20.5 to 39 ms and the interquartile range went from 8 to 18ms.



PR Interval (top), QRS Duration (center), and QT Interval (bottom) Compared to CSE

Boxes in box plots denote 25th and 75th percentiles, with 50th percentile (median) inside the box. Whiskers extend to 2.5th and 97.5th percentiles, spanning 95% of the measurement differences.

Independent Assessments of 12SL Measurements

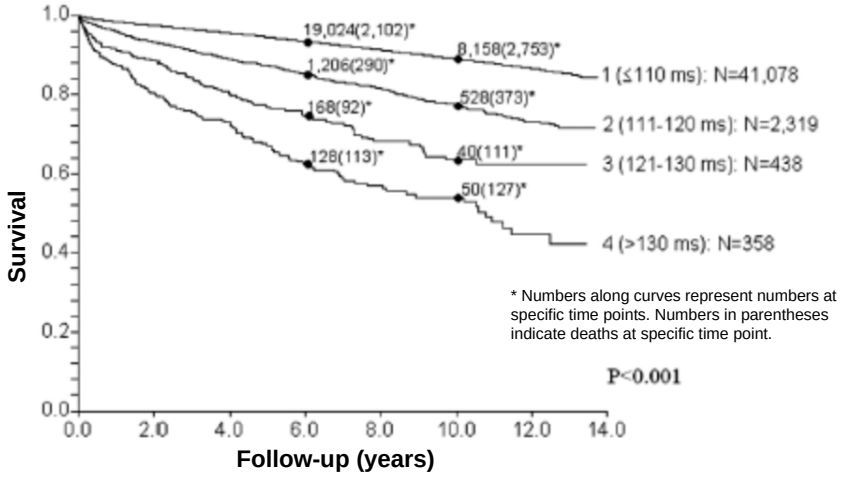
There have been several independent assessments of the measurements generated by GE's Marquette 12SL ECG analysis program, ranging from an evaluation for routine clinical use[58, 59] through to an assessment as to whether the measurements are appropriate for large clinical trials or epidemiology studies.[60]

Independent Assessment of QRS Duration

Based on the QRS duration measurement made by GE's Marquette 12SL program, several studies have explored whether QRS duration can predict death[61] or indicate the presence of congestive heart failure.[62] QRS duration has also been investigated as an indicator for patients that benefit most from cardiac resynchronization therapy.[63-65] Below are some quotes from the scientific literature with regards to GE's automated QRS duration measurement:

- "The widest QRS duration on each ECG was manually measured after magnification. ... Compared with computer measurements of QRS duration, the correlation coefficient (r) was 0.95, with a SE of 0.06, $p < 0.0001$"[66]
- "Of the 4,033 patients, 252 died during a median follow-up of 3 years. The QRS duration was univariately associated with an increased risk of death (relative risk 8.5, 95% confidence interval CI 4.4 to 16.4, $p < 0.0001$) ... A QRS duration >105 ms best identified patients at increased risk. In conclusion, QRS duration is associated with an increased risk of death, even after adjustment for clinical factors, exercise capacity, left ventricular function, and exercise-induced myocardial ischemia."[67]
- "Prolonged QRS was associated with a significant increase in mortality (49.3% vs 34.0%, P = .0001) and sudden death (24.8% vs 17.4%, P = .0004)."[68]
- "A target population of 3,471 had .... ECG data obtained from automated sources during the first year of diagnosis. .... Among the heart failure population, 20.8% of the subjects had a QRS duration > 120 ms. A total of 425 men (24.7%) and 296 women (16.9%) had a prolonged QRS duration ($p < 0.01$). There was a linear relationship between increased QRS duration and decreased ejection fraction ($p < 0.01$). A prolonged QRS duration of 120 to 149 ms demonstrated increased mortality at 60 months ($p = 0.001$), when adjusted for age, sex, and race ($p = 0.001$). Systolic dysfunction was associated with graded increases in mortality across ascending levels of QRS prolongation."[69]

- "Analyses were performed on the first electrocardiogram digitally recorded on 46,933 consecutive patients." Using computer generated QRS durations from 12SL, the following conclusion was made: "QRS duration provides a simple method to stratify patients as to their risk of cardiovascular (CV) death. In a general medical sample, without BBB or paced rhythms, those with a QRS duration greater than 130 ms experience nearly twice the risk of cardiovascular death compared with those with a QRS duration of 110 ms or less. Similarly in patients with LBBB and RBBB, QRS duration greater than 150 ms is associated with greater risk of CV death."[70]



Independent Assessment of ST Deviations

Quoting from the literature, here are some assessments of ST measurements made by the 12SL program:

- "The predictive value of nonspecific ST depression as determined by visual and computerized Minnesota Code (MC) codes 4.2 or 4.3 was compared with computer-measured ST depression > or = 50 microvolts in 2,127 American Indian participants in the first Strong Heart Study examination. .... Concludions: Computer analysis of the ECG, using computerized MC and computer-measured ST depression, provides independent and additive risk stratification for cardiovascular and all-cause mortality, and improves risk stratification compared with visual MC."[71]
- In this study, computerized ST measurements were correlated with the presence of left ventricular hypertrophy (LVH). ECGs and echocardiograms (ECHO) we done on a total of 1,595 American Indian participants without evident coronary disease.[72] "The absolute magnitude of ST segment deviation above or below isoelectric baseline was measured by computer in leads V(5) and V(6), and participants were grouped according to gender-specific quartiles of maximal STdep. Left ventricular hypertrophy was defined by indexed LV mass >49.2 g/m(2.7) in men and >46.7 g/m(2.7) in women. ... After controlling for clinical differences, increasing STdep remained strongly associated with increased prevalence of LVH ($p = 0.0001$). Conclusions: In the absence of evidence of coronary disease, increasing STdep in the lateral precordial leads is associated with increasing LV mass and increased prevalence of anatomic LVH."

- ST deviations were evaluated in 69 consecutive patients suspected of an acute coronary syndrome.[73] Bland-Altman analysis demonstrated clinically acceptable limits of agreement comparing measurements of the J point and the T wave, but clinically inadequate limits of agreement with respect to ST-segment deviation, between the electrocardiographer and the computer. But as quoted from the study: “The difference between these two methods is mainly caused by different measurement points. There is no common agreement on what time point to use to measure ST amplitude. In this study, it was measured at 80 ms after the J point by manual measurement, while the computer selected a displacement at the midpoint of the ST segment.” This measurement is known as STM, which is 1/8th of the average RR interval after the J point.

Independent Assessment of QT Measurements

The assessment of automated QT measurements has undergone a great deal of scrutiny due to the challenge of consistent measurement of small changes (<6ms) for drug-induced trials.[74] Automated measurements are desirable since the reduction of effort in performing manual measurements may result in a lower sample size and overall cost of a trial.[75]

One drug-induced QT study concluded: “Manual and automated measurements generated similar numerical results in these 3 studies in healthy volunteers, which all included a positive control. There is little evidence to suggest that manual methods have advantages over automated methods in measuring QT, and the clinical interpretations remain the same.”[76]

In another study, which evaluated normals and patients with hypertrophic cardiomyopathy, the automatic QT measurements made by GE’s Marquette 12SL analysis program were “more stable and reproducible than the manual measurements”.[25]

The stability and consistency of the 12SL analysis program was recently leveraged in for the measurement of QT in a large epidemiology study, because the QT variability of the 12SL Program “was smaller than that of the Dalhousie program.”[77] This study derived normal limits from percentile distributions for QT as well as QT and T-wave subintervals in 22,311 participants in the Women’s Health Initiative (WHI). This study advised considerable revision of the currently used limits for prolonged QT in women, with an additional race-specific adjustment in Asian women. The study also recommended that Bazett’s formula is inappropriate for testing new drugs or other applications.

Similar normative values were established in another study, which was conducted on a large drug-induced trial patient population using 12SL Program measurements and medians, available for review by a cardiologist.[78] The analysis was performed on baseline (drug-free) ECG data. The final analysis population included 13,039 baseline ECG recordings from 13,039 patients. Reference ranges from the study are stratified by important prognostic factors: age, sex, and overall ECG evaluation at baseline (normal or abnormal). From this study, proposed reference ranges may be useful for patient management and data analyses in clinical drug development, in addition, the article provides a QT correction formula to correct the QT interval for heart rate. This QT correction formula was shown to be superior to the Bazett and Fridericia corrections in a clinical trial population in the ability to minimize the correlation between QT and RR.

In 2006, a large independent study evaluated the new QT algorithm for the 12SL analysis program, which was released in 2003 and is now available in all current GE Healthcare electrocardiographs. Evaluation of computerized QT measurements from 12SL was done on over 45,000 resting ECGs obtained from two clinical trials, labeled as set "A" and "B". Set "A" (n=15,194 ECGs) exhibited substantially better signal quality than set "B" (n=29,866 ECGs). In recording set A, 95.9% of ECGs were measured automatically within 10 ms of the manual measurement. In recording set B, 83.9% of the automated measurements were within 10ms. "The study shows that (a) compared to the "old" version of the 12SL algorithm, the QT interval measurement by the "new" version implemented in the most recent GE ECG equipment is significantly better, and (b) the precision of automatic measurement by the 12SL algorithm is substantially dependent on the quality of processed ECG recordings. The improved accuracy of the "new" 12SL analysis program makes it feasible to use modern ECG equipment without any manual intervention in selected parts of drug-development program."[20]

Table 8. Percentages of ECGs with successful automatic QT measurement (n=45,060)[20]

Absolute measurement error	ECG Set A		ECG Set B	
	"New" 12SL	"Old" 12SL	"New" 12SL	"Old" 12SL
≤ 5 ms	73.7	47.8	54.4	33.5
≤ 10 ms	95.9	76.6	83.9	59.5
≤ 15 ms	99.3	91.7	94.0	77.3

Table 8 shows percentages of ECG tracings in which the error of automatic QT interval measurement was below the given threshold. For example, with a given threshold of 10ms, 95.9% of the ECGs in set A were within 10 ms of the manual measurement as opposed to only 76.6% of the ECGs with the "old" 12SL measurement algorithm.

Furthermore, GE's computerized QT / T wave measurements,[12, 13, 79] including QT dispersion and principal component analysis, have been correlated with overall mortality[26, 80-83] as well as acute ischemia.[10, 84-87]

Despite these positive statistics, it is important to note that outliers do occur in an automated interpretation. Furthermore, congenital QT abnormalities provide their own unique challenges to the program accuracy, as identified in the literature.[88, 89] Given this new established performance of the 12SL program in drug-induced trials,[20] GE Healthcare is currently re-evaluating the performance of 12SL on databases specifically developed for the management of congenital long QT syndromes.

Accuracy of Interpretive Statements: Reported Results

Purpose of Reported Results

The Statement of Validation and Accuracy is considered official product labeling and is reviewed by the Food and Drug Administration (FDA) and the International Electrotechnical Commission (IEC). This document primarily serves as a disclosure of the accuracy of the interpretive statements generated by GE's Marquette 12SL analysis program. This is in contrast to a description of how interpretive statements are generated by the program; that is the purpose of another document, known as the 12SL physician's guide.

In 1991, the FDA recommended that such a document as The Statement of Validation and Accuracy be generated for the clearance of a 1500 Series Prehospital Defibrillator[6] that incorporated GE's Marquette 12SL Program and was the first prehospital defibrillator to provide automated analysis of the prehospital 12-lead ECG.[5] Since 1991, The Statement of Validation and Accuracy has periodically been updated to keep abreast of the latest scientific findings regarding the 12SL Program. In 2003, the IEC issued a similar request for all manufacturers of ECG analysis equipment: that is, the IEC asked the manufacturers of ECG analysis programs and equipment to report the sensitivity, specificity, and positive predictive accuracy of the interpretive statements for each of the major diagnostic categories (see 60601-2-51(c) IEC 2003).[27] Like the FDA, the IEC also requested that these results be published and available to the consumer. This Statement of Validation and Accuracy fulfills this requirement.

The Marquette 12SL analysis program has continually evolved since it was first introduced in 1980. Each released version of the program contains one or more changes to it and is associated with a unique version number. This number appears on the ECG report printed by the analyzing electrocardiograph. The number is also printed on each ECG from the MUSE system. Encoded within this number are two elements: the actual 12SL version number and a product specific code, which refers to the type of product used for the analysis. The 12SL Physician's Guide (PN 416791-004) contains a table that clarifies these codes and identifies the related 12SL version numbers.

The Marquette 12SL analysis program has continually evolved since it was first introduced; however, only portions of the program are changed for any one particular software version. The rest of the executable is tested to insure that it generates the same results as the last version (see the previous description of the development and validation process for 12SL). Based on the 12SL version number, the state of revision of each portion of the program can be determined.

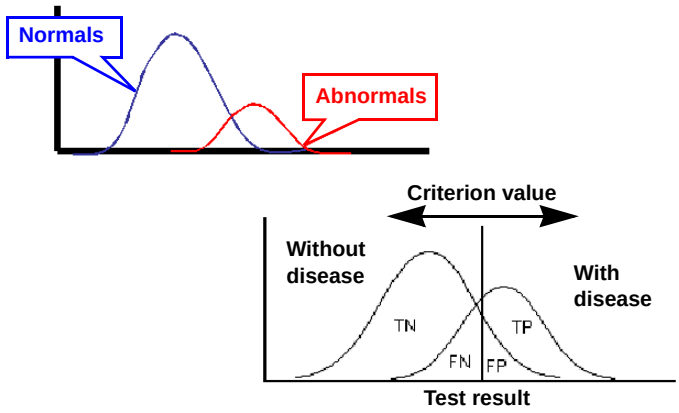
Scientific references and results presented in this document span a variety of dates. Portions of the program that have not been recently changed can rely on reported results that are older, and yet remain representative of the current state of that portion of the program. Sections of the program that have recently been enhanced require more recent publications. Depending upon which portion of the program is used for a particular diagnostic statement, different results reported in the literature can be used to characterize the performance of that particular statement as long as the results were generated subsequent to any substantial change to that portion of the program. Care has been taken to insure

that results from the literature and presented in this document are representative of the current version of the 12SL analysis program.

Although scientific references and results presented in this document reflect the current performance of the 12SL analysis program, it would be unwise to directly extrapolate these to what will occur in a particular clinical environment. Furthermore, these are statistical measures, not the performance that one should expect for a particular patient.

Definition of Sensitivity, Specificity, and Other Performance Metrics

For the purpose of this document, four key accuracy measures are explained in this section. It is assumed that the true diagnosis for a patient is known (that is, the "truth"). The ECG interpretation (classification) is called a "Test".



Performance Metrics

The following designations are applied to characterize the performance of a test.

- "Normal" correctly classified as "Normal" is called "True normal" (TN)
- "Normal" incorrectly classified as "Pathologic" is called "False pathologic" (FP)
- "Pathologic" incorrectly classified as "Normal" is called "False normal" (FN)
- "Pathologic" correctly classified as "Pathologic" is called "True pathologic" (TP)

Table 9. Tabulation of Test Results

Reference	Test	
	"Normal"	"Pathologic"
"Normal"	TN	FP
"Pathologic"	FN	TP

The following equations are calculated from a two- (or multi-) category test:

1. Sensitivity:
   probability that a “True pathologic” would be classified as “Pathologic”

   Sensitivity = TP / (TP+FN) x 100%

2. Specificity:
   probability that a “True normal” would be classified as “Normal”.

   Specificity = TN / (TN+FP) x 100%

3. Positive predictive value (PPV):
   probability that a classified “Pathologic” is a “True pathologic”.

   PPV = TP / (TP+FP) x 100%

   NOTE

   The previous explanation can be made general by substituting “Negative” for “Normal” and “Positive” for “Pathologic”.

4. Negative predictive value (NPV):
   probability that a classified “Normal” is a “True normal”.

   NPV = TN / (TN+FN) x 100%

Description of Table Format for Reporting Interpretation Metrics

In order to present the performance metrics for GE’s Marquette 12SL program, each study reported in this document uses one of the tables as presented in the following example. Note that the overall description of the study is presented in the header of the table, including the total number of ECGs for the particular study, the representative population or care environment where the ECGs were acquired for the study, and the independent scientific method used for verifying the disease or pathology. (See IEC 60601-2-51 clauses 50.102.3.1 and 50.102.3.2.)

In the following example, 110 ECGs were collected in an emergency department from patients with chest pain of unknown origin. Each patient was tested for cardiac Troponin, a very sensitive and specific indicator of an acute myocardial infarction (AMI). Such details of the study and the method used to verify the diagnosis can be pursued via the bibliography reference associated with the title of the table. (See “Bibliography” on page 66.) In this example, only 10 patients were positive for Troponin. As a result, under the column labeled “N”, the number “10” appears in the row labeled as acute myocardial infarction. Therefore, “N” has to do with the number of patients who have been verified for a particular diagnosis, “N” has nothing to do with number of ECGs that were positive or negative for the recognition of AMI. In this specific example, the program correctly identified 4 of the 10 patients as having an AMI. As a result, the sensitivity for the program is listed as 40%. Note: this does not necessarily mean that the program made an ECG interpretation error on the other 6 patients. Rather, it could mean that the ECG did not reveal any ST elevation. From the remaining 100 patients that were negative for Troponin, the program falsely recognized 1 as being an AMI. As a result, the specificity is listed as 99%. Since a total of 5 patients were called AMI by the program, but only 4 were correct, the positive predictive value is 80%.

Table 10. Example: Study "A"[Ref A]				
Representative test population . . . . . Emergency department, patients with chest pain of unknown origin Total number of test ECGs . . . . . 110 Method(s) used to verify diagnosis: . . . . . Troponin				
Verified Diagnosis	N	Sensitivity (%)	Specificity (%)	PPV (%)
Acute Myocardial Infarction	10	40	99	80

Also notice that the tables indicate that this is a "test population" and that these are "test ECGs" or a validation set. This is an important distinction for the reporting of performance of the automated recognition of disease: that is, the term test ECG / validation set means that GE's Marquette 12SL analysis program was not trained with the data that was collected for the study. Rather, the study provided results on a test set, not a training set. Typically, the performance of program will be worse on a test set than a training set.

Bayes Theorem and Intended Use: Understanding Performance Metrics

The tables in this document report sensitivity, specificity, positive predictive value (PPV) and, sometimes, negative predictive value (NPV). Depending on the distribution and prevalence of disease in a particular population, a high-level of specificity may be more important than a high level of sensitivity. In the above example, there are only 10 individuals with the disease out of a population of 110. A 10-point drop in specificity would lead to many more mistakes (10% of 100 results in 10 mistakes) as opposed a 10-point drop in sensitivity (10% of 10, results in one mistake). However, it may be important to find every sick individual if a particular therapy can be applied that cures the disease but is not detrimental to the healthy individual. In this case, a high sensitivity, which typically results in a loss in specificity, may be warranted if there is no risk for treating a false positive, healthy individual. These issues are beyond the scope of this document but are discussed in the literature.[90, 91]

Interpretation of Rhythm: Reported Results

This section provides performance metrics as reported in the literature regarding rhythm interpretations generated by GE's Marquette 12SL Program. Results are reported for the following major rhythms: sinus, ectopic atrial rhythm, atrial tachycardia, atrial fibrillation, atrial flutter, junctional rhythm, and artificially paced. In addition, results are reported for the following rhythm modifiers: 1st degree AV block, 2nd AV block, 3rd AV block, and premature atrial / ventricular beats. The IEC also requires manufacturers to disclose rhythms, without reported results, due to their low rate of prevalence. (See IEC 60601-2-51 clause 50.102.4.1.) These include idioventricular rhythm, ventricular tachycardia, ventricular fibrillation, and wandering atrial pacemaker as well as statements regarding escape or fusion beats. Also, no reported results exist for interpretations regarding the rate or character of AV conduction during atrial fibrillation or atrial flutter.

Asynchronous P-Wave Detection via QRS Subtraction

Interpretation of cardiac rhythms is highly dependent on accurate detection of atrial activity. As a result, improved P wave detection has been a major pursuit of GE Healthcare.[92-94] Since 1998, a sophisticated tool, called MAC-RHYTHM, was incorporated into GE's Marquette 12SL ECG analysis program for the detection of asynchronous P waves, hidden within the QRS or T wave.[95]

Previous versions of the program, which did not incorporate the QRS subtraction tool for P-wave detection, have been evaluated for rhythm interpretation accuracy and reported in the literature.[96, 97] The metrics in all tables presented below are from the later versions of the program, which incorporated MAC-RHYTHM.

QRS Subtraction / MAC-RHYTHM: Prospective Study on 10,761 ECGs

The value of the QRS subtraction tool was prospectively tested on 10,761 ECGs.[14] Quoting from the study:

"For three of the abnormal rhythms, namely, atrial fibrillation, junctional rhythms, and second degree atrioventricular blocks, MAC-RHYTHM gave significantly higher sensitivity in both prospective (87.5%, 92.2%, and 80.8%, respectively) and retrospective (82.0%, 81.2%, and 79.6% respectively) testing than the [old program] (65.0%, 39.6%, and 12.0% respectively). Similarly, for sinus rhythms, MAC-RHYTHM had significantly higher specificity (prospective, 91.0% and retrospective, 91.7%) than the [old program] (86.5%). The specificity for the abnormal rhythms remained very high with MAC-RHYTHM (prospective, 99.4% to 99.7% and retrospective, 99.1% to 99.7%) compared to the [old program] (99.0% to 99.9%)."

Table 11. Prospective Study Using MAC-RHYTHM[14]

Representative test population. . . . . Hospital, all departments
Total number of test ECGs 10,761
Method(s) used to verify diagnosis . . . . . Confirmed by experienced cardiologist

Verified Diagnosis	N	Sensitivity (%)	Specificity (%)	NPV (%)	PPV (%)
Sinus rhythms	9,324	98.7	91.0	91.5	98.6
Atrial fibrillation	832	87.5	99.4	99.0	92.4
Atrial flutter	106	76.4	99.7	99.8	71.7
Junctional	64	92.2	99.5	100.0	52.7, (72.8)*
2nd-degree AV blocks	26	80.8	99.6	100.0	32.8

* After excluding paced ECGs with failed pace detection.

Enhancements to QRS Subtraction, Tested on 69,957 ECGs

Since the addition of the QRS subtraction tool, several enhancements were made to the P wave detector. This included spectral analysis for the detection of atrial flutter; optimal lead selection for P wave detection; and T wave alignment to reduce subtraction artifact in the residual signals used to create a P wave detection function.[98]

As published in the literature:

> "Performance was assessed using a test set of 69,957 confirmed ECGs from four hospitals. The rhythm interpretation in the confirmed ECG was compared to the rhythm interpretations from the previous and new versions of the program. The rate of disagreements between the confirmed rhythm and the computerized interpretation decreased from 6.9% to 4.1%. Sensitivity improved for sinus, atrial fibrillation, atrial flutter, and junctional rhythms, while specificity and positive predictive value improved for all arrhythmias."[98]

Table 12. Four Hospitals, Random Selection of ECGs[98]

Representative test population . . . . . Four hospitals, all departments
Total number of test ECGs 69,957
Method(s) used to verify diagnosis . . . . . Routine confirmation by cardiologists

Verified Diagnosis	N	Sensitivity (%)	Specificity (%)	PPV (%)
Sinus	62397	98.2	85.5	98.3
Atrial fibrillation	5163	89.0	99.4	91.9
Ectopic atrial rhythm	1066	35.2	99.7	63.4
No P waves	635	63.1	99.1	38.1
Atrial flutter	576	55.0	99.6	50.7
2nd/3rd degree AVB	120	49.1	99.6	18.1

Subsequent Evaluations Yield Similar Results

Recently, Poon[99] analyzed the interpretation performance for rhythm on 3,954 non-paced ECGs analyzed by the 12SL analysis program. As quoted from the literature: "Our findings differ only modestly from the corresponding performance characteristics for sinus rhythm, atrial fibrillation, and atrial flutter recently reported by Farrell et al."

Table 13. Evaluation Done in 2005 at NY Presbyterian Hospital [99]

Representative test population . . . . . University Hospital
Total number of test ECGs 4297
Method(s) used to verify diagnosis . . . . . Confirmation by 2 cardiologists

Rhythm Category	N	Sensitivity (%)	Specificity (%)	PPV (%)
Primary Rhythms				
Sinus	3579	98.7	90.1	99.0
Atrial fibrillation	250	90.8	98.9	84.7
Atrial flutter	41	61.0	99.9	83.3
Atrial tachycardia	360	2.8	99.9	25.0
Rhythm Modifiers				
Premature atrial complexes	212	64.2	99.5	87.2
Premature ventricular complexes	162	82.7	99.1	80.2

In another study, a total of 2194 consecutive ECGs from 1856 patients were collected from a tertiary care VA Hospital from both inpatients and outpatients. The results for rhythm analysis are summarized below. Not all rhythms, for example sinus rhythms, were reported in the study.

Table 14. Evaluation of Rhythm Analysis Done in 2006 at Tertiary Care, VA Hospital[100]				
Representative test population . . . . . Tertiary care, VA Hospital Inpatients & Outpatients Total number of test ECGs . . . . . 2194 from 1856 patients Method(s) used to verify diagnosis . . . . . Confirmation by 2 cardiologists				
Rhythm Category	N	Sensitivity (%)	Specificity (%)	PPV (%)
Primary Rhythms				
Atrial fibrillation	67	76.1	99.6	85.0
Atrial flutter	41	65.9	99.9	93.1
Permanent pacemaker	56	73.2	99.9	93.2
2nd degree AV block	1	100	99.7	14.3
Rhythm Modifiers				
1st degree AV block	138	97.8	99.7	95.7
Premature ventricular complexes	150	94.0	99.5	94.0
Premature atrial complexes	94	66.0	99.5	86.1

In another study, ECGs were acquired from symptomatic patients with isolated pulmonary hypertension. The blinded and un-blinded cardiologist and computer program analysis agreed regarding the rate and rhythm in each case (n=64). Sinus rhythm was present in 96.9% of patients; one patient had an ectopic atrial rhythm and one had a junctional rhythm. The heart rate averaged 84.1 ± 15.5 b/min. Sinus bradycardia was present in 5, sinus tachycardia in 6, and first degree atrioventricular block in 7 patients; 2 patients had a complete right bundle branch block.[101]

Table 15. ECGs from Symptomatic Patients With Pulmonary Hypertension[101]				
Representative test population. . . . . University hospital, Patients with pulmonary hypertension Total number of test ECGs . . . . . 64 Method(s) used to verity diagnosis . . . . . Confirmation by 2 cardiologists				
Rhythm Category	N	Sensitivity (%)	Specificity (%)	PPV (%)
Primary Rhythms				
Sinus	62	100	100	100
Ectopic atrial rhythm	1	100	100	100
Junctional Rhythm	1	100	100	100
Rhythm Modifiers				
1st degree AV block	7	100	100	100
RBBB	2	100	100	100

Note that the aforementioned studies yield similar results, despite the different locations and environments. This increases the confidence that these results will be reproducible in other populations.

In addition to these aforementioned studies, an evaluation of the clinical consequences of misdiagnosed atrial fibrillation by a computer was performed at Henry Ford Hospital in Detroit, Michigan. A total of 2298 ECGs were identified with a computerized diagnosis of atrial fibrillation by GE's Marquette 12SL analysis program. Of these 2298 ECGs, 442 (or 19%) from 382 (35%) of the 1085 patients had been incorrectly interpreted as atrial fibrillation. The paper did not report the total number of true atrial fibrillation ECGs across the entire sampled population, only the number of "true positives" and "false positives' from the computerized interpretation. Therefore only the positive predictive value may be calculated. In 92 patients (that is, 24% of the inaccurate computerized interpretations), the physician ordering the ECG, failed to correct the inaccurate interpretation. Clinical consequences of this misdiagnosis are presented in the paper. The conclusion of this work is that greater efforts should be directed toward educating physicians about the electrocardiographic appearance of atrial dysrhythmias and the recognition of confounding artifacts.

Table 16. Evaluation of Misdiagnosis of Atrial Fibrillation by Computer[43]

Representative test population . . . . . Large, university hospital
Total number of test ECGs 2298
Method(s) used to verify diagnosis . . . . . Patient chart and follow-up

Rhythm Category	N	PPV (%)
Atrial fibrillation	2298	81.0

This value of 81% for the positive predictive accuracy for the computerized recognition of atrial fibrillation is lower but comparable to the other studies presented here. Noise in the ECG tracing is certainly a confounding factor in this study. Note that 38% of the misinterpretations by both the computer and physician were due to artifact.[43, 102] Quality control of noise is a critical factor for proper ECG interpretations by both the physician and computer.[21, 52]

Paced Rhythms

Improvements in electronic pacemaker pulse generators and lead design as well as the increasing use of bipolar pacing have led to the reduction of pulse amplitudes and widths observable on the digitized surface ECG.

In 1983, a prospective evaluation study in one hospital published by Swiryn and Jenkins reported pacemaker detection sensitivity for the GE's Marquette 12SL Program to be 87.5%.[97] By 1998, a prospective analysis of more than 10,000 ECGs analyzed by essentially the same detection algorithm in one hospital had a corresponding sensitivity of only 71.5%. Specificity in both samples was very high at 99.9%.[14] This reduction of 16% in sensitivity is most likely due to the advent of low energy pacemaker artifacts. In 2000, GE Healthcare enhanced the software that operates with the pacemaker detection circuitry in order to improve sensitivity while maintaining specificity.[16]

In 2001, this software was evaluated on 100 of 103 consecutive patients seen in a device clinic who were asked to participate in the study. Two consecutive paced ECGs were recorded from each patient with pacemaker amplitudes and pulse widths at arrival or discharge settings. In 86 patients, two additional consecutive

ECGs with underlying non-paced rhythm were recorded by lowering the pacemaker rate thresholds. The implanted devices included 44 single and 56 dual chamber devices (41 ICDs; 59 pacemakers; 92 bipolar leads). Pulse width settings ranged between 0.3 ms and 3.0 ms and voltage settings ranged between 0.9 and 6.0 V. Sensitivity for detecting paced rhythms using the new method was 87% compared to 41% using the old method (x^2=45.9, $p < 0.0005$). For both methods, specificity was 100% for this data set.

Table 17. Evaluation of Pacemaker Detection[103]

Representative test population . . . . . Pacemaker clinic, large hospital
Total number of test ECGs 372
Method(s) used to verify diagnosis . . . . . Patient History, Pacemaker Programmer

Rhythm Category	N	Sensitivity (%)	Specificity (%)	PPV (%)
Paced	200	87	100	100

Similarly, in 2002, a prospective trial was done at a different institution on 100 pacemaker clinic patients. ECGs were obtained from all patients in the clinic. At least two paced ECGs, and whenever possible, two non-paced ECGs were obtained from each patient. A total of 389 ECGs were collected and analyzed; 235 ECGs were paced and 154 were non-paced. Both the new and old algorithms had high specificity for pacemaker detection (>99.4%). The new algorithm had a sensitivity of 87% versus 30% for the old algorithm.

Table 18. Evaluation of Pacemaker Detection[104]

Representative test population . . . . . Pacemaker clinic, large hospital
Total number of test ECGs 389
Method(s) used to verify diagnosis . . . . . Patient History, Pacemaker Programmer

Rhythm Category	N	Sensitivity (%)	Specificity (%)	PPV (%)
Paced	235	87	99.4	99.5

In 2006, a large study was conducted that solely focused on pacemaker recognition and rhythm interpretation in the presence of electronic pacemakers. “Of the 7834 consecutive ECGs screened, a pacemaker (PM) was identified by the computer, the cardiologists, or both in 205 ECGs. The cardiologists detected an electronic pacemaker in 201 tracings, whereas the computer detected one in 168 tracings. In 4 ECGs that were read as having an electronic pacemaker by computer, no pacemaker was present according to both cardiologists. Therefore, in 164 of 205 ECGs (80.0%), both computer and cardiologists agreed upon the presence of an electronic pacemaker. The sensitivity of recognizing a pacemaker by computer was 82.0%, and the specificity was 99.9%. In 37 cases, the algorithm failed to recognize the presence of a pacemaker. A common error was missing the ventricular spike (16 cases). Other errors included missing both the atrial and ventricular spikes (10 cases) and, rarely, the atrial spikes alone (4 cases).”[105]

Table 19. Evaluation of Computer Analysis of Pacemaker (PM) Rhythms[105]				
Representative test population . . . . . VA Hospital - Inpatients & Outpatients Total number of test ECGs . . . . . 7834 Method(s) used to verify diagnosis . . . . . Confirmed by 2 cardiologists				
Rhythm Category	N	Sensitivity (%)	Specificity (%)	PPV (%)
Paced ECG	205	82.0	99.9	96

The article concludes that: “Automated computer ECG reading algorithms are useful tools for ECG interpretation, but they need further refinement in recognition of electronic pacemakers (PM). In 61.3% of ECGs with electronic PM, computer-drawn interpretation required revision by cardiologists. In 18.4% of cases, the ECG reading algorithm failed to recognize the presence of a PM. Misinterpretation of paced beats as intrinsic beats led to multiple secondary errors, including myocardial infarctions in varying localizations. The most common error in computer reading of ECGs with PMs is the failure to identify an underlying rhythm.”[105]

Poon reported similar results for the analysis of paced tracings. Quoting from the article: “The most common errors were related to interpretive statements involving patients with pacemakers: of 343 ECGs with pacemaker activity comprising 8.0% of the study ECGs, 75.2% (258/343) required revision, so that 45.7% of all inaccurate rhythm statements in this population occurred in patients with pacemakers. Overall, 13.2% (565/4297) of computer-based rhythm statements required revision, but excluding tracings with pacemakers, the revision rate was 7.8% (307/3954).”[99]

Pediatric Rhythm Interpretation

Recently, two studies have evaluated pediatric populations. The first was in an emergency department (ED); the other was across a large pediatric hospital.

In the first study, a total 294 cases were evaluated.[89] The patients ranged in age from 5 days to 21 years. The ED physicians interpreting the ECGs were directly involved in the patients' care and were familiar with the presenting complaint, past medical history, and physical examination. Physicians were allowed to use whatever means available to aid with ECG interpretation. The physicians were blinded to the computer interpretations. The reference standard was the ECG interpretation by a pediatric electrophysiologist.

Each electrocardiographic diagnosis, as well as the ECG as a whole, was assigned to one of the following predetermined classes: I, normal sinus rhythm; II, minimal clinical significance; III, indeterminate clinical significance; IV, those of definite clinical significance.

Both the computer and ED physician correctly interpreted all normal (class I) ECGs correctly (that is, normal sinus rhythm / normal ECG). The computer correctly diagnosed class II ECGs 82% of the time as compared to 67% by the ED physicians ($p<0.001$). The computer was also significantly more accurate than the ED physicians with regard to the class III diagnoses, correctly interpreting 73% compared to 30% by the physicians ($p<0.001$). With regard to the individual class IV ECG diagnoses, the ED physicians were more accurate than the computer (28% vs 14%), but this difference did not reach significance ($p>0.3$).

Pediatric rhythm interpretation resulted in a majority of computer errors in this study. Quoting this work: "Despite its superior ability to accurately interpret many of the simple rhythm disturbances, the computer was less accurate than the ED physicians with regards to interpreting ECGs with abnormal Supraventricular rhythms. Specifically, the computer failed to identify all 4 ECGs with junctional rhythm, 2 of 4 with supraventricular tachycardia, and 2 with intraatrial reentry tachycardia."[89]

This study did not assess specificity. "The over interpretation of ECGs by either the computer or ED physicians was not evaluated in this study."[89] As a result, the results of this study cannot be represented in the table recommended by the IEC.[27]

The second study evaluated 56,149 pediatric ECGs.[106] From this list, 2 groups of patients were selected: patients with heart disease and those without heart disease. The ECGs were systematically selected in the stratified groups to ensure balanced representation in terms of age, sex, etc. This resulted in a sample size of 1,147 ECGs. The reported results for rhythm are presented in Table 20.

Table 20. Evaluation of Pediatric Rhythm Interpretation[106]

Representative test population . . . . . Large pediatric hospital
Total number of test ECGs. . . . . 1,147 (sampled from 56,149)
Method(s) used to verify diagnosis. . . . . Confirmed by 2 pediatric cardiologists

Rhythm Category	N	Sensitivity (%)	Specificity (%)	PPV (%)
Sinus Rhythm in presence of Heart Disease	399	95.5	99	99
Sinus Rhythm in normal group	390	98.5	100	100
Sinus Arrhythmia in presence of Heart Disease	31	87	100	100
Sinus Arrhythmia in normal group	51	88	100	100
Sinus Rhythm with Ectopy in Heart Disease group	10	100	98.5	56
Sinus Rhythm with Ectopy in normal group	22	100	98	69

Interpretation of P-wave Abnormalities: Reported Results

This section provides performance metrics, as reported in the literature, for interpretation of right and left atrial abnormalities.

Table 21. Evaluation of Right and Left Atrial Abnormality at Tertiary Care, VA Hospital[100]

Representative test population . . . . . Tertiary care, VA Hospital - Inpatients & Outpatients
Total number of test ECGs . . . . . 2194 from 1856 patients
Method(s) used to verify diagnosis . . . . . Confirmation by 2 cardiologists

P Wave Abnormality	N	Sensitivity (%)	Specificity (%)	PPV (%)
Right	29	100	99.9	97
Left	97	95.5	100	100

Interpretation of QRS Abnormalities: Reported Results

This section provides performance metrics, as reported in the literature, for the computerized interpretation of QRS abnormalities. These include: right bundle branch block (RBBB), left bundle branch block (LBBB), left ventricular hypertrophy (LVH), right ventricular hypertrophy (RVH) as well as healed anterior and inferior myocardial infarction. The IEC also requires manufacturers to disclose those QRS abnormalities without reported results. (See IEC 60601-2-51 clause 50.102.3.1). These include the following statement categories:

- Wolff-Parkinson-White (WPW),
- QRS axis deviation abnormalities,
- hemi-blocks,
- low-voltage QRS, and
- pulmonary disease pattern.

In addition, isolated lateral or posterior myocardial infarctions have no reported results; instead, these statements are grouped with inferior or anterior myocardial infarctions.

Conduction

At Mount Sinai Medical Center in New York City, over 39,000 ECGs were reviewed for computer accuracy.[107]. The cardiologist was used as the reference, since interpretative statements regarding conduction are Type B statements.

A detailed inspection of the data from the Mount Sinai study showed that the cardiologist often changed the computer diagnosis to LBBB (n=97) from another conduction abnormality already stated by the program (like ILBBB or nonspecific intraventricular conduction block). If these other conduction abnormalities were included as part of the analysis, the sensitivity would increase from 78% to 88%.

Table 22. Independent Assessment of Conduction Abnormalities[107]

Representative test population . . . . . Hospital, all departments
Total number of test ECGs 39,000
Method(s) used to verify diagnosis . . . . . Confirmation by cardiologists

Verified Diagnosis	N	Sensitivity (%)	Specificity (%)	PPV (%)
RBBB	1661	90	100	100
LBBB	860	78	100	100
LBBB (grouped w/ ILBBB, IVCB)	860	88	100	100

At the Mayo clinic, the 12SL program was evaluated to determine whether it could replace an ECG program, based on XYZ Leads, with the 12SL program, which is based on the scalar 12-lead ECG.[108] In a similar fashion as the aforementioned study, over 12,000 ECGs were evaluated at the Mayo Clinic. See Table 23, “Independent Assessment of Conduction Abnormalities[109],” on page 45.

Table 23. Independent Assessment of Conduction Abnormalities[109]

Representative test population . . . . . Hospital, all departments
Total number of test ECGs 12,793
Method(s) used to verify diagnosis . . . . . Confirmation by cardiologists

Verified Diagnosis	N	Sensitivity (%)	Specificity (%)	PPV (%)
RBBB	391	91	100	100
LBBB	248	87	99.9	99.9

In another study,[100] ECGs were collected in a tertiary care facility from inpatients (36.4%), outpatients (47.6%), and emergency room patients (16.0%). There were 2194 consecutive ECGs recorded on 1856 patients.
Two cardiologists read the ECGs. Of the 2,194 tracings, 122 were excluded from analysis because of a disagreement between the cardiologists’ interpretations. Out of 2072 remaining cases, 776 (37.5%) the computer interpreted as normal and 1296 as abnormal. In 206 cases, there were discordances between the computer and cardiologists' interpretation (9.9%). There were no discordances in the ECGs interpreted as normal by the computer. Therefore, the discordances occurred in 15.9 % of all ECGs read as abnormal.

Conduction abnormalities were evaluated as part of this study. The results are reported below:

Table 24. Independent Assessment of Conduction Abnormalities by Two Cardiologists[100]

Representative test population . . . . . Hospital, all departments
Total number of test ECGs 2072
Method(s) used to verify diagnosis . . . . . Confirmation by 2 cardiologists

Verified Diagnosis	N	Sensitivity (%)	Specificity (%)	PPV (%)
RBBB	118	93.2	99.8	96.5
LBBB	33	90.9	99.9	90.9

Assessment of RBBB in a Pediatric Population

RBBB in a pediatric population is exhibited in a narrow QRS. This diagnosis was evaluated at a pediatric hospital using 56,149 ECGs stored on a MUSE system. From this list, 2 groups of patients were selected: patients with heart disease and those without heart disease. The ECGs were systematically selected in the stratified groups to ensure a balanced representation. This resulted in a sample size of 1,147 ECGs. RBBB is a Type B statement and can thus be validated by a pediatric cardiologist.

Table 25. Assessment of RBBB in Pediatric Population[106]

Representative test population . . . . . .Hospital, all departments
Total number of test ECGs 1,147
Method(s) used to verify diagnosis . . . . . .Confirmation by 2 pediatric cardiologists

Verified Diagnosis	N	Sensitivity (%)	Specificity (%)	PPV (%)
RBBB	123	79.6	99.8	99

Hypertrophy

Two independent studies have evaluated the performance of our 12SL analysis program for *left ventricular hypertrophy* (LVH) using echocardiography (ECHO).

At the Mayo Clinic, an ECHO test was performed within 30 days of the ECG. ECGs demonstrating WPW syndrome, paced rhythm, or LBBB were excluded from the study. ECHO studies were excluded for patients who were less than 21 years of age. All two dimensional and M-mode ECHO studies were technically adequate and required clear delineation of interventricular septal thickness (IVST), posterior wall thickness (PWT), and left ventricular internal dimension (LVID). Patients with IVST/PWT>1.5, segmental wall motion abnormalities, pericardial effusion, or infiltrative cardiomyopathy were excluded from the study. This resulted in a test population of 4,300 patients.

ECHO measurements were made according to the American Society of Echocardiography. ECHO studies revealed LVH in 1,029 patients.
LVH was defined as:

ECHO LV mass >265g
LV mass = 1.04 ((LVID + PWT + IVST)3 - (LVID)3) - 13.6g

The 12SL analysis program correctly identified 328 patients with LVH and 3,010 patients without LVH. The program was scored as stating LVH for the full breadth of statements that refer to the abnormality; including “minimal (and moderate) voltage criteria for LVH, may be normal.” Table 26, “LVH by ECG and Cross Correlation with ECHO[108],” on page 47 summarizes the program’s performance.

Table 26. LVH by ECG and Cross Correlation with ECHO[108]				
Representative test population . . . . . Hospital, all departments Total number of test ECGs. . . . . 4,300 Method(s) used to verify diagnosis. . . . . ECHO				
Verified Diagnosis	N	Sensitivity (%)	Specificity (%)	PPV (%)
LVH	1,029	31.9	92	57

In addition to the Mayo Clinic study, a large international study evaluated program performance for hypertrophy.[110] In this study there were a total of 1220 patients, 382 controls and 838 with cardiac disorders that were collected across five European centers. ECGs showing complete *Left Bundle Branch Block* (LBBB), *Right Bundle Branch Block* (RBBB) or other major intraventricular conduction defects were excluded; otherwise there were no other criteria for excluding ECGs. A normal individual (n=286) was defined as being free of significant cardiopulmonary disease on the basis of a health screening examination (negative history, normal physical exam, normal chest X-ray) or invasive cardiac study (n=96). Invasive studies usually entailed cardiac catheterization (CATH) for atypical chest pain or ST/T abnormalities evident at rest or during exercise. LVH was based on CATH or ECHO or both. Specific details regarding the population are contained in the article.[110]

Table 27. Performance of LVH and RVH by ECG, Validated by CATH and ECHO[110]				
Representative test population . . . . . 5 European Academic Centers, Hospitals Total number of test ECGs . . . . . 1220 Method(s) used to verify diagnosis . . . . . ECHO, CATH, Clinical History				
Verified Diagnosis	N	Sensitivity (%)	Specificity (%)	PPV (%)
Hypertrophy, all kinds	291	61.1	91.2	85
LVH	183	76.2	91.2	82
RVH	55	29.1	100	100

In another study, patients with pulmonary hypertension due to pulmonary vascular occlusive disease were evaluated in the Pulmonary Hypertension Clinic at the University of Michigan. Each underwent a thorough history, physical exam, ECG, echocardiogram, pulmonary function testing, and right heart catheterization. Symptoms (type and duration), effort tolerance, and New York Heart Association (NYHA) functional class were recorded during the initial visit. Pulmonary hypertension was defined as a mean pulmonary artery pressure >25 mmHg. Patients were excluded if they presented with evidence of chronic lung disease, left ventricular hypertrophy, mitral or aortic valve disease, congenital heart disease, coronary artery disease or cardiomyopathy.[101]

Table 28. Performance of RAE and RVH by ECG, Validated by CATH and ECHO[101]

Representative test population. . . . . Hospital, Academic Center
Total number of test ECGs 64
Method(s) used to verify diagnosis . . . . . ECHO, CATH, Pulmonary artery pressure

Verified Diagnosis	N	Sensitivity (%)	Specificity (%)	PPV (%)
RVH	64	39.1	100	100
Right Atrial Enlargement	13	46	100	100

The blinded cardiologist and computer program diagnosed RVH in 43.8 and 39.1% of patients, respectively; this is substantially lower than the 78.1%, as determined by the un-blinded reader that was provided the age and clinical parameters (i.e. symptoms associated with possible pulmonary hypertension). Right ventricular strain was present in 71.9% of patients, and was most often characterized by the blinded cardiologist and the computer program as non-specific or inferior / anterior-lateral ischemia. The most common errors by the computer and blinded cardiologist were the diagnosis of an anterior-septal infarction based on the presence of a qR in V1 (10.9%), and of an inferior-posterior myocardial infarction because of the presence of a “pathologic” Q wave in II, III and aVF associated with a prominent R in V1 (6.2%)

The study concluded that the ECG does have a high specificity for the detection of RVH in symptomatic patients with pulmonary hypertension and that correlation with the clinical parameters is essential to optimize the usefulness of the ECG. Without the clinical parameters, the computer program and blinded cardiologist often suggested myocardial infarction / ischemia.

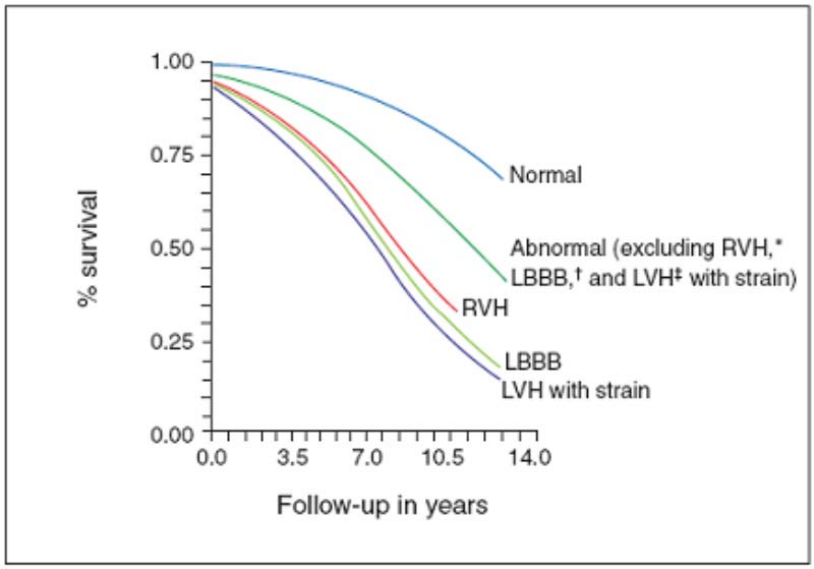
In another study, two cardiologists were considered as the gold standard. As expected, performance metrics for the program are much higher when they are based on this human standard.

Table 29. Evaluation of ventricular hypertrophy at tertiary care, VA Hospital[100]

Representative test population . . . . . Tertiary care, VA Hospital, Inpatients & Outpatients
Total number of test ECGs . . . . . 2194 from 1856 patients
Method(s) used to verify diagnosis . . . . . Confirmation by 2 cardiologists

Hypertrophy Category	N	Sensitivity (%)	Specificity (%)	PPV (%)
Right Ventricle (RVH)	15	100	99.9	66.7
Left Ventricle (LVH)	399	98.7	99.5	98

In addition to the evaluation of accuracy, GE's Marquette 12SL interpretation of LVH has been evaluated in terms of its prognostic value on 26,734 male and 3,737 female veterans.[81] The computerized interpretation was used without modification. Computer detected abnormalities associated with the lowest survival rates are presented below. Note that LVH with strain is the most predictive and that a normal ECG as defined by the 12SL program "is associated with extremely good survival".[81]



Assessment of RVH in a Pediatric Population

Criteria for RVH, in a pediatric patient, are defined by 16 different age categories.[4, 111] This diagnosis was evaluated at a pediatric hospital using 56,149 ECGs stored on a MUSE system. From this list, 2 groups of patients were selected: patients with heart disease and those without heart disease. The ECGs were systematically selected in the stratified groups to ensure balanced representation. This resulted in a sample size of 1,147 ECGs

Note that RVH is a Type A statement: that it typically requires non-ECG data for a reference gold-standard. However, in this case, the authors used the opinion of 2 pediatric cardiologists.

Table 30. Assessment of RVH in a Pediatric Population[106]

Representative test population . . . . . Hospital, all departments
Total number of test ECGs 1,147
Method(s) used to verify diagnosis . . . . . Confirmed by 2 pediatric cardiologists

Verified Diagnosis	N	Sensitivity (%)	Specificity (%)	PPV (%)
RVH	93	91.3	99.8	99

Myocardial Infarction

There are several independent studies that have evaluated the performance of GE's Marquette 12SL analysis program to recognize healed myocardial infarction (MI).[112] The term "healed myocardial infarction" implies that this section is reporting results on the ability of the program to detect QRS abnormalities (like abnormal Q-waves) associated with necrosis. Computerized interpretation of a myocardial infarction is a Type A statement, requiring independent validation from non-ECG data.

CATH as the Reference

The first series of evaluations of the 12SL program were done on ECGs from subjects that were selected from consecutive patients undergoing cardiac catheterization.[113, 114] The presence of an MI was determined via wall motion abnormalities associated with a 75% or greater obstruction of the relevant coronary artery. Patients with pulmonary disease, valvular disease, a history of previous MI, LV wall motion abnormalities suggesting multiple MIs, and patients with a history of previous cardiac surgery were excluded. Normals were defined as having normal LV motion and coronary arteries. This resulted in a study population of 734 patients with an MI and 406 patients defined as normal. The infarction group consisted of 84% males with an average age of 55 years. The average age of the 121 female patients was 57 years. ECGs selected for analysis were obtained on average 3 days before the CATH in 92% of the infarction group patients. The remaining 8% were done within 30 days following the CATH procedure. The normal group consisted of 41% males with an average age of 46 years. The average age of the 238 female patients was 52 years. ECGs were obtained, on average, within 4 days before the CATH in 99% of the normal patients.

The results for the performance of the program versus CATH are presented in Table 31. Note that the physician had a similar level of sensitivity (69%) but maintained a higher level of specificity (97%).

Table 31. Performance of MI: Group All Statements Indicating MI[113]

Representative test population . . . . . Hospital
Total number of test ECGs 1140
Method(s) used to verify diagnosis . . . . . CATH, Clinical History

Verified Diagnosis	N	Sensitivity (%)	Specificity (%)	PPV (%)
Myocardial Infarction	734	70	92	94

Influence of Modifiers “Cannot Rule Out”, “Possible”

This same study also evaluated the performance of statements that were preceded by the modifiers *cannot rule out* and/or *possible*. When these statements were not considered diagnostic for MI, the sensitivity was reduced to 54% while the specificity improved to 98%.

Table 32. Performance MI Statements Without Modifiers *Cannot Rule Out*, *Possible*[113]

Representative test population . . . . . Hospital
Total number of test ECGs 1140
Method(s) used to verify diagnosis . . . . . CATH, Clinical History

Verified Diagnosis	N	Sensitivity (%)	Specificity (%)	PPV (%)
Myocardial Infarction	734	54	98	98

Inferior Myocardial Infarction

Using the same aforementioned source of data, an evaluation of inferior MI was conducted,[114] which demonstrated that the 12SL program had a sensitivity of 76% and a specificity of 95% while the physician had a lower sensitivity (75%) but a higher specificity (97%) than the computer.

Anterior Myocardial Infarction

In a separate study conducted at a Veterans Administration hospital, 137 patients were evaluated via cardiac catheterization using similar methods for data acquisition and analysis as the aforementioned study but, in this case, the focus was anterior myocardial infarction. Patients who had significant valvular heart disease, left bundle branch block or paced rhythm were excluded. However, no attempt was made to identify and exclude patients with either left ventricular enlargement or chronic obstructive pulmonary disease, conditions that can reduce the specificity of ECG criteria for anterior myocardial infarction. All the ECGs were obtained on or near the day of each patient's catheterization. Of the 137 patients, the normal group consisted of 82 patients and the anterior MI group consisted of 55 patients. Below are the reported results for the 12SL analysis program:

Table 33. Performance of Anterior MI by ECG, validated by CATH[115]

Representative test population . . . . . Veterans Administration Hospital
Total number of test ECGs 137
Method(s) used to verify diagnosis . . . . . CATH

Verified Diagnosis	N	Sensitivity (%)	Specificity (%)	PPV (%)
Anterior MI	55	64	99	99

Evaluation of Both Inferior and Anterior Myocardial Infarction via Cath

Another large international study also used CATH as the reference but relied solely on the assessment of wall motion abnormalities, not including coronary obstruction. The results are presented in Table 34:

Table 34. Performance of Anterior and Inferior MI by ECG, validated by CATH [110]

Representative test population. . . . . 5 European Academic Centers, Hospitals
Total number of test ECGs 1220
Method(s) used to verify diagnosis . . . . . CATH, wall motion studies

Verified Diagnosis	N	Sensitivity (%)	Specificity (%)	PPV (%)
Anterior MI	170	66	98	84
Inferior MI	273	65	97	86

Evaluation of Old Myocardial Infarctions Based on Cardiologist Opinion

In another study, two cardiologists were defined as the standard. As expected, the performance metrics of the program are markedly higher using this human standard.

Table 35. Evaluation of Ventricular Hypertrophy at Tertiary Care, VA Hospital[100]

Representative test population. . . . . Tertiary care, VA Hospital – Inpatients & Outpatients
Total number of test ECGs . . . . . 2194 from 1856 patients
Method(s) used to verify diagnosis . . . . . Confirmation by 2 cardiologists

Category	N	Sensitivity (%)	Specificity (%)	PPV (%)
Old myocardial infarctions	399	98.8	99.5	97.4

MI Sizing / Electrocardiographic Damage Scores

There are several electrocardiographic damage scores, which are used to predict the size or severity of the myocardial infarction. These scores primarily rely on an analysis of QRS abnormalities, such as Q waves. Based on measurements generated by GE's Marquette 12SL analysis program, the following damage scores have been evaluated: Selvester Score,[116] Simplified Selvester Score,[117] and Cardiac Infarction Injury Score (CIIS).[118]

A fully automated version of the Selvester Score was validated versus manual measurements and the results demonstrated that it "had a high correlation with manual application ($r = 0.94$) and was superior regarding time, training, reader bias, reproducibility and precision of measurement."[119] This automated version evaluated ECGs from 1,344 normal subjects, 706 patients with a single myocardial infarction (366 with inferior infarction, 277 with anterior infarction and 63 with posterolateral infarction), and 131 patients with combined inferior and anterior infarction.[120] The presence and location were determined by CATH criteria, similar to the other aforementioned study done by Haisty.[114] A score greater than 4 yielded a sensitivity of 67% for anterior infarction, 41% for inferior infarction, 32% for posterolateral infarction and 72% for multiple infarcts. However, 7 of 32 criteria failed to achieve 95% specificity and 10 of 35 criteria in criteria sets had a sensitivity that was even lower than their false positive rate. Quoting from the literature: "the automated Selvester QRS scoring system currently has limitations that are attributable to development of the original manual system, which used manual scoring techniques and established criteria limits from middle-aged men".[120] Note that the sensitivity and specificity of the Selvester Score are all less than the reported results of the standard interpretation of myocardial infarction by the 12SL analysis program, even though the score is often used as a reference.[121, 122]

In a more recent study, ECGs from 46,933 patients were used to evaluate the prognostic value of these electrocardiographic damage scores.[123] The Simplified Selvester Score, the Cardiac Infarction Injury Score (CIIS), and a Q-wave score were calculated based on the computerized measurements generated by the 12SL program. The main outcome was cardiovascular mortality. During a mean follow-up of 6 years, the CIIS outperformed all other ECG classifications in determining prognosis.

There is renewed interest in MI-sizing via a QRS score, due to advances in cardiovascular magnetic resonance as a new reference for myocardial size[124, 125] and the possible implications that MI size could have on prophylactic ICD therapy.[126] However, more work needs to be done on this promising technology.[127]

Repolarization Abnormalities: Reported Results

Repolarization abnormality computerized interpretations are composed of Type A and C statements. Recall that Type C statements refer to purely descriptive ECG features that usually cannot be documented by any other means. Examples of such statements include *non-specific ST-T abnormality*. This document will primarily be reporting results of the Type A statements, which are verified by non-ECG data such as cardiac enzymes, patient outcomes, etc.

This document reports results for ST-elevated acute myocardial infarction (STEMI). Other statements associated with ST elevation, namely early repolarization and acute pericarditis are not directly reported. However, these other ST elevation interpretations are analyzed appropriately as part of the assessment of STEMI: that is, they would be classified as FN (false normal) or TN (true normal) with respect to the enzyme data. In addition to STEMI, ST segment depression and T wave abnormalities associated with the interpretation of ischemia are presented.

The IEC requires manufacturers to disclose those interpretative repolarization statements that have no reported results in the literature. (See IEC 60601-2-51 clause 50.102.3.1). These include interpretations regarding subendocardial injury, an abnormal QRS-T angle, non-specific ST or T-wave abnormality and digitalis effect. In addition, repolarization abnormalities interpreted as part of old infarctions or hypertrophies are not reported separately from the commensurate QRS abnormality. In this document, the interpretation of prolonged QT has been included under the assessment of the automated QT measurement.

ST Elevated Acute Myocardial Infarction

The recognition of *ST-elevated acute myocardial infarction* (STEMI) has been a major focus of GE Healthcare. This is because the ECG is so vital in selecting an appropriate treatment path for acute myocardial infarction[128] as well as reducing time-to-treatment for STEMI.[129]

Prehospital Electrocardiography

GE Healthcare was the first to introduce a prehospital diagnostic 12 lead ECG as a small, compact unit for the ambulance that could acquire and transmit the ECG digitally so that there would be no distortion of the ST/T waveform.[130] This led to several studies that demonstrated that a prehospital ECGs can be practically acquired,[131] significantly cuts total time-to-treatment,[132-134] and has “the potential to significantly increase the diagnostic accuracy in chest pain patients.”[135]

Based on data collected from the prehospital environment,[136] GE’s Marquette 12SL analysis program was modified to recognize earlier forms of STEMI, using reciprocal depression as the primary discriminating characteristic to discern STEMI versus early repolarization.[5] This approach, combined with enhancements, allowed the sensitivity to double without a loss of specificity.[137, 138] Several tests have since verified that reciprocal depression is a highly specific marker of STEMI.[139-141]

GE's Marquette 12SL analysis program (Version 14) is used in prehospital defibrillators currently offered by other vendors (Medtronic-PhysioControl, Zoll).[142, 143] GE's resting electrocardiographs use a later version that includes such features as gender and age-specific criteria for the recognition of STEMI[144] and the detection of right ventricular involvement in the presence of an acute inferior infarction.[23] As a result, the following reported results for STEMI are presented in two groups: one that applies to the results of the program in the prehospital defibrillator and one for the results of the program in GE's resting ECG equipment. Note that both versions of the program analyze data of the same fidelity and content, generating fiducial points and medians at 500 sps.[7]

STEMI - Reported Results, Prehospital ECGs

The following series of reported results are from prehospital ECGs and are representative of version 14 of the 12SL analysis program.

In Australia, a GE Healthcare portable prehospital electrocardiograph[145] was used for the automatic diagnosis of acute myocardial infarction via GE's Marquette 12SL analysis program. "This automated program diagnosed acute evolving Q wave myocardial infarction with 71% sensitivity and 98% specificity. Specificity was 100% when patients with a known previous Q wave myocardial infarction were excluded."[141, 146]

Table 36. Results from GE's Prehospital Electrocardiograph[141]

Representative test population. . . . . Prehospital ECGs
Total number of test ECGs 526
Method(s) used to verify diagnosis . Physician interpretation, serial ECG analysis, & clinical outcome

Verified Diagnosis	N	Sensitivity (%)	Specificity (%)	PPV (%)
Acute MI	Unknown	71	98	Unknown
Acute MI, no previous MI	Unknown	71	100	100

As part of the NIH sponsored Myocardial Infarction Triage and Intervention (MITI) Project,[147] the 12SL analysis program accuracy for recognizing STEMI was evaluated. This was a large prehospital study (n=1,189) that acquired ECGs from patients within 6 hours of the onset of chest pain. This study used cardiac enzymes as the "gold standard". Their conclusion: "the positive predictive value of the computer- and physician-interpreted ECG was, respectively, 94% and 86% and the negative predictive value was 81% and 85%."[148] The authors also stated: "The present algorithm is clearly adequate for first line screening of patients with chest pain by paramedics or in the emergency department. Its sensitivity is no worse than that of the emergency physician and its specificity is superior to the trained electrocardiographer. ... Although more sensitive, the electrocardiographer had an overall incidence of a 5% false positive diagnosis, including a 22% incidence of false positive diagnoses in patients with isolated ST segment elevation. In contrast, the computer was nearly perfect at excluding patients without acute myocardial infarction, but did so at the expense of diminished sensitivity." The raw numbers for algorithm performance are given in the following Table 37.

Table 37. Results from the MITI Trial Based on Cardiac Enzymes[148]				
Representative test population. . . . . Prehospital ECGs, large city Total number of test ECGs . . . . . 1,189 Method(s) used to verify diagnosis . Cardiac enzymes				
Verified Diagnosis	N	Sensitivity (%)	Specificity (%)	PPV (%)
Acute MI	391	52	98.5	94

The results of the MITI trial were also analyzed for the recognition of STEMI as opposed to solely using cardiac enzymes as the reference. That is, an analysis was done as to whether or not ST elevation was present along with the positive cardiac enzyme result. In this case, the program achieved a sensitivity of 71%. As stated in the literature: "The computer algorithm was developed to help differentiate early repolarization and nonspecific ECG changes from those of acute injury and, unlike the electrocardiographer, did not presume that ST elevation in a patient with chest pain was more likely than not to indicate acute infarction. Although more sensitive, the electrocardiographer has an overall incidence of 5% false positive diagnoses, including a 22% incidence of false positive diagnoses in patients with isolated ST segment elevation."[148]

Table 38. Results from the MITI Trial Based on Cardiac Enzymes and Presence of ST Elevation[148]				
Representative test population. . . . . Prehospital ECGs, large city Total number of test ECGs . . . . . 1,189 Method(s) used to verify diagnosis . . . . . Cardiac Enzymes and ST elevation				
Verified Diagnosis	N	Sensitivity (%)	Specificity (%)	PPV (%)
STEMI	286	71	98.5	94

In another study, clinical data and ECG findings on 264 consecutive patients admitted to a coronary care unit with suspected acute myocardial infarction were prospectively evaluated with the same portable prehospital electrocardiograph as in the aforementioned prehospital studies. Eighty-six (86) patients (32.5%) had confirmed acute infarction and of these 85% had some form of ST elevation on their initial ECG. The area under the receiver operator curve (ROC) of the interpretations made by the 12SL analysis program was 83.9%.[139]

A recent survey of 365 hospitals in the United States, found that hospitals that used the results of prehospital "electrocardiography, that were called in or transmitted by emergency medical services to activate the catheterization laboratory while the patient was still en route to the hospital, had significantly faster door-to-balloon times than did hospitals that waited for the patient to arrive before activating the catheterization laboratory (P = 0.001)."[149] Furthermore, this survey found that "false alarms were reported to be infrequent."[149] The authors also stated that the perception "about the number of false alarms are probably as important" in determining "whether non-cardiologists are permitted to activate the catheterization laboratory".[149]

STEMI - Reported Results for Resting Electrocardiographs

The following series of reported results are representative of the current version of the 12SL analysis program.

In the following study, body surface mapping (80 leads) was compared with GE's Marquette 12SL analysis program for the recognition of acute myocardial infarction on ECGs taken over a 3-month period from 103 chest pain patients in the ED.[150] Of these, 53 had an acute myocardial infarction as defined by positive enzymes. Only 24 met ECG criteria for STEMI.

The purpose of this study was to not only detect STEMI but to detect non-ST elevated acute myocardial infarction. The motivation of the study was to reveal that body surface mapping is superior because it can detect non-ST elevated acute myocardial infarction. Note that the 12SL analysis program is designed not to detect non-ST elevated acute myocardial infarction; rather it will indicate ST depression or T wave inversion. Based on the severity of these abnormalities, the current program will state, *marked ST depression, consider subendocardial injury* or *marked T wave abnormality, consider ischemia*. It remains controversial as to whether the ECG can diagnose non-ST elevated acute myocardial infarction: this diagnosis is currently the sole domain of cardiac enzyme data.[151]

See the reported results of this study below. The admitting physician correctly diagnosed 24 patients with AMI (sensitivity 45%, specificity 94%). Of the 24 patients correctly diagnosed, 20 received thrombolytic therapy. According to care guidelines, thrombolytic therapy should only be applied in the case of a STEMI.[128] The automated analysis program correctly diagnosed 17 patients with STEMI (sensitivity 32%, specificity 98%).

Table 39. Results for STEMI Based on Cardiac Enzymes[150]

Representative test population . . . . . Emergency Department
Total number of test ECGs 103
Method(s) used to verify diagnosis . . . . . Cardiac Enzymes (CK-MB, Troponin)

Verified Diagnosis	N	Sensitivity (%)	Specificity (%)	PPV (%)
Acute MI	53	32	98	98

Table 40. Results for STEMI Based on Cardiologist[150]

Representative test population . . . . . Emergency Department
Total number of test ECGs 103
Method(s) used to verify diagnosis . . . . . Positive STEMI by Cardiologists

Verified Diagnosis	N	Sensitivity (%)	Specificity (%)	PPV (%)
STEMI	24	71	98	98

In the next study, 75 electrocardiograms were interpreted. “Two criteria were compared for thrombolysis eligibility: (1) measurement of > or =1 mm ST-segment elevation in 2 contiguous leads (measured) and (2) criterion 1 plus the subjective opinion that the changes represented acute transmural injury (interpretive). The results were compared with computerized interpretations by the Marquette 12SL system.”[152]

The ECGs for this study[152] were manually selected in a CCU and were roughly evenly divided among (1) normal, (2) those showing evidence of acute transmural injury, and (3) those showing other ST-segment or T-wave abnormalities (such as early repolarization, acute pericarditis, etc.) Note: this distribution of patient abnormalities is not representative of an ED, CCU, or emergency medical service that typically has a much lower incidence of acute transmural injury (that is, on the order of 10-15%).[153]

This paper states that “strict reliance on measured electrocardiographic criteria alone would have resulted in overuse of thrombolysis among all 3 raters. Based on the consensus opinion, the absolute overuse of thrombolysis would have been approximately 15% (P <.0034).” In contrast, the computer had 100% specificity.

Table 41. ECGs from Cardiac Care Unit (CCU) Evaluated by 3 Cardiologists, Consensus Opinion[152]

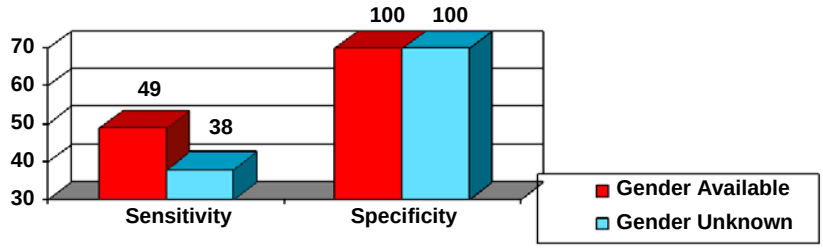
Representative test population . . . . . Emergency Department
Total number of test ECGs 75
Method(s) used to verify diagnosis . . . . . Consensus of 3 Cardiologists

Verified Diagnosis	N	Sensitivity (%)	Specificity (%)	PPV (%)
STEMI	26	61.5	100	100

STEMI - Gender Specific Criteria in GE Healthcare Resting Electrocardiographs

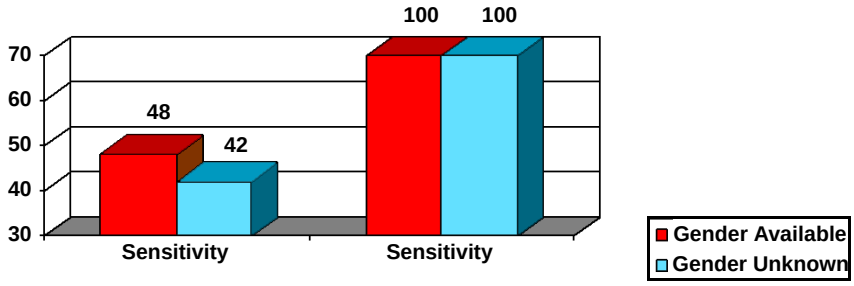
GE has done considerable research in gender specific differences in the ECG. Testing was done via data collected at the Mayo Clinic and the Medical College of Wisconsin. Results of testing, and an analysis of the ECG differences based on gender, have been broken down by location of myocardial infarction: that is, anterior versus inferior.

For acute inferior MI patients under age 60, women had lower ST elevation than men (lead II STJ average: 57µV for 99 females versus 86µV for 340 males, P value <.02). The opposite was true for patients over age 60. In the older patient population, women had larger ST elevation than men (lead AVF STJ average: 130µV for 378 females versus 84µV for 522 males, P value < .04). The following figure displays a comparison of the results, between the two program versions, for the recognition of acute inferior myocardial infarction in women less than 60 years of age.[15]



Acute Inferior MI Detection in Women, age <60

For acute anterior MI patients under age 60, women had lower ST elevations than men (lead V2 STE average, 307µV for females versus 432µV for males, P value < .007). Over age 60 years, this difference becomes less pronounced (lead V2 STE average, 336µV for females versus 421µV for males, P value < .009). The figure displays a comparison of the results between the two program versions for the recognition of acute anterior myocardial infarction in women less than 60 years of age.[154]



Acute Anterior MI Detection in Women, age <60

Test results show that the program is more sensitive for the recognition of acute myocardial infarction in women less than 60 years of age. For ages 60 and over, the program performance is the same as in previously published studies.

Table 42. ECGs from Cardiac Care Unit (CCU) Evaluated by 3 Cardiologists, Consensus Opinion[152]

Representative test population . . . . . Emergency Department
Total number of test ECGs 3,457
Method(s) used to verify diagnosis . . . . . Cardiac Enzymes, Clinical Outcomes

Verified Diagnosis	N	Sensitivity (%)	Specificity (%)	PPV (%)
Acute Inferior MI	1,339	49	100	100
Acute Anterior MI	1,305	48	100	100

STEMI - Right Ventricular Involvement in GE Healthcare Resting Electrocardiographs

AHA / ACC guidelines recommend that patients with inferior STEMI and hemodynamic compromise should be assessed with a right precordial lead V4r to detect ST segment elevation to screen for right ventricular (RV) infarction.[128] This is a class I recommendation, meaning that there is evidence and/or general agreement that a given procedure or treatment is beneficial, useful, and effective. RV involvement in acute inferior infarction may be accompanied by significant hemodynamic consequences including a lowering of cardiac output and systemic blood pressure.[155] In addition, the in-hospital mortality of an acute inferior infarct is worsened when complicated by RV involvement.[156]

The 12SL ECG analysis program uses a threshold of 100 µV in lead V4r in interpreting all cases of right ventricular involvement, except under very specific circumstances.[23] Specifically, the program reduces the threshold to 50 µV in the presence of an acute inferior STEMI with high-degree AV block and a rightward ST vector (i.e., STE in III > II).[157-159] The prevalence of high-degree AV block (i.e., 2nd or 3rd degree AV block) in the general population is extremely rare and a person with an acute inferior STEMI and concomitant high-degree AV block is more than twice as likely to have RV involvement than not.[160]

ST elevation of 100 µV in lead V4r is a highly specific indicator of right ventricular involvement in the presence of acute inferior infarction. A threshold of 100 µV has been reported to have sensitivities of 57% - 100% and specificities of 68% - 100%, depending on the gold standard used (post-mortem examination, hemodynamic measures, angiography, etc).[161] A threshold of 50 µV has been reported to have sensitivities of 76% - 100% and specificities of 40% - 86%, again depending on the gold standard.[161, 162] Morgera[163] analyzed both thresholds in the same study with the same patient population and reported a specificity increase from 86% to 100% as the threshold went from 50 to 100 µV, with a sensitivity decrease from 76% to 57%. However, one should note that the diagnostic accuracy of right ventricular involvement statements have not been assessed in patients with certain conditions such as chronic lung disease and pericardial disease.

Although the lower ST elevation threshold in lead V4r will increase sensitivity and decrease specificity, this decreased specificity is offset by the requirement of concomitant ST elevation in lead III exceeding ST elevation in lead II and high-degree AV block, both of which are associated with right ventricular involvement. Using only the criteria of ST in III > II, Saw[157] reported a sensitivity of 97% and a specificity of 56% for the detection of right ventricular involvement in the presence of an acute inferior infarction. The reported incidence of high degree AV block in patients with RV involvement is 43%, compared to only 13% in patients with acute inferior infarction without RV involvement.[160]

GE Healthcare developed a 16-lead ECG database in conjunction with several chest-pain centers. A total of 1,343 16-lead ECGs were acquired and analyzed from 712 chest-pain patients. Each ECG record contained the standard 12-lead ECG, simultaneously acquired with leads V4r, V7, V8, and V9. GE Healthcare, in conjunction with the contributing investigators, analyzed and reported on the characteristics of the additional leads in relation to acute myocardial infarction and outcome.[164-166] The interpretation of GE's Marquette 12SL analysis program was compared to patient outcomes, as registered in this 16-lead ECG

database. An acute STEMI was detected in 143 ECGs. Of these, 101 were diagnosed as being an acute inferior STEMI (including inferolateral and inferior-posterior). When V4r was withheld from the analysis, *consider RVI* was stated in 84 of the 101 IMI ECGs. When V4r was included in the analysis, the *with RVI* modifier was added in 34 of the 101 IMI ECGs. With one exception, all 12-lead ECGs that stated *consider RVI* also stated *with RVI* when V4r was added.

The sensitivity of the *consider RVI* statement for predicting positive ST elevation in V4r was 97% (33 / 34), while the positive predictive accuracy was 39% (33 / 84). The result here of 34% (34 / 101) of all acute inferior STEMIs having RVI is consistent with the percentages of 30 - 50% reported in the literature.[167].

Repolarization Abnormalities Associated with Acute Cardiac Ischemia

ACI-TIPI[168] uses the measurements of GE's Marquette 12SL program. Based on the presence of pathologic Q waves and/or the presence of repolarization abnormalities, the ACI-TIPI algorithm reports the probability of acute cardiac ischemia. The logistic regression formula used by ACI-TIPI[169] was implemented in all GE electrocardiographs and tested in the emergency department (ED)[170] as well as the prehospital environment.[9]

A large prospective trial was accomplished across 10 different emergency departments, with 30-day follow-up of clinical outcomes. A total of 10, 689 patients were evaluated: 8150 were not ischemic, 673 had stable angina, and 1866 had acute cardiac ischemia (that is, unstable angina or an acute myocardial infarction. Quoting from the literature:[171]

> "Reductions in admissions for patients without acute cardiac ischemia were greater among patients with ACI-TIPI-predicted ischemia probabilities in the lower ranges, reflecting a greater effect with stronger probabilistic advice not to admit (that is, a dose-response effect). Of note, in settings in which use of the ACI-TIPI reduced unnecessary admissions, appropriate hospital and CCU admission did not deteriorate for patients with true acute ischemia (unstable angina or acute infarction). Given these results of this 'effectiveness' trial ACI-TIPI seems to be safe and effective for general use."

ACI-TIPI had a larger impact when the attending physician was inexperienced (that is, an unsupervised resident). In this case, "use of ACI-TIPI was associated with a reduction in CCU admissions from 14% to 10%, a change of -32% (CI, -55% to 3%); a reduction in telemetry unit admissions from 39% to 31%, a change of -20% (CI, -34% to -2%) and an increase in discharges to home from 45% to 56%, a change of 25% (CI, 8% to 45%; overall P = 0.008)."

The purpose of this study was to measure the impact of care based on whether ACI-TIPI was available or not available. Within the same ED, ACI-TIPI was available on alternate months. The effect of improved triage with ACI-TIPI was reproducible, even after the physician had several months of experience with the device.

Repolarization Abnormalities Stated as Ischemia

Using two cardiologists as the reference, the following results were reported for the interpretations of ischemia by computer:

Table 43. Evaluation of ST/T abnormalities stated as ischemia at tertiary care, VA Hospital[100]

Representative test population . . . . . Tertiary care, VA Hospital – Inpatients & Outpatients
Total number of test ECGs . . . . . 2194 from 1856 patients
Method(s) used to verify diagnosis . . . . . Confirmation by 2 cardiologists

ST/T Abnormality	N	Sensitivity (%)	Specificity (%)	PPV(%)
Ischemia	199	100	99.8	98

Overall Classification: Reported Results

Several studies have addressed the issue of whether or not the computer can reliably classify the ECG as either normal or abnormal. The following studies reported the following:

- "the program is reliable in diagnosing normality: even the disagreements are arguable."[59]
- "From a practical point of view, the eventual consensus opinion of the cardiologists was that only one tracing reported as normal by the system definitely should have been reported as abnormal to a family doctor, resulting in a negative predictive value of 98.4%. In view of the cardiologists inter-observer variation with regard to what is normal, this may well be higher than an individual cardiologist's negative predictive value and suggests that the system examined may safely be used to exclude major abnormalities which would affect clinical management".[59]
- "A total of 39, 238 electrocardiograms were reviewed ... The program placed the ECG into the following diagnostic classifications: normal 22%, otherwise normal 6%, borderline 5%, abnormal 66%. The reviewing physician agreed with this classification in 96.3% of all cases ... The most striking information shows the agreement of the physicians with the computer diagnosis of an abnormal electrocardiogram in 97.7% of the 25,295 tracings. In only 204 records out of 25,987 tracings (.8%), the physicians edited a computer-called abnormal electrocardiogram and changed it to normal. Likewise, in only 63 of 8,632 (.7%) tracings of which the computer called normal did the physicians edit this tracing to read abnormal."[107]

Table 44. Overall Classification via Large Database

Representative test population . . . . . Large hospital
Total number of test ECGs 39,238
Method(s) used to verify diagnosis . . . . . Physician diagnosis

Verified Diagnosis	N	Sensitivity (%)	Specificity (%)	PPV(%)
Normal ECG	8,632	99.9	100	99.9
Abnormal ECG	25,987	99.9	99.9	99.9

- As tested on 26,734 male and 3,737 female veterans, a classification of a normal ECG by the 12SL analysis program "is associated with extremely good survival".[81]
- "Three ECG computer programs–Hewlett Packard analog program (HP), Telemed analog program (T) and Marquette 12SL digital program (MAC)–were evaluated and their accuracy of ECG reading compared with the reading of 4 experienced interpreters on 140 ECGs of patients with various clinical abnormalities. Major disagreement with effect on patient management, and minor disagreement were defined at a joint session with a senior (consensus). The computers identified all normal ECGs correctly (sensitivity 100%). The percentage of major agreements (full agreements and minor disagreements) between consensus and computer was 79% for HP, 90% for T and 93% for MAC."[172]

- "A total of 2194 ECGs were included for analysis in the study. One hundred twenty two ECGs with a disagreement between the two cardiologists were excluded from analysis. Out of 2072 remaining cases, 776 (37.5%) were read by the computer as normal ... There were no discordances in the ECGs read as normal."[100]
- The computer correctly interpreted all normal ECGs.[89]
- "The quality of computer-assisted ECG interpretation was comparable to that of review provided by a cardiology service."[29] As a result, the overall result of the computerized interpretation is comparable in performance to the average cardiologist. (See IEC 60601-2-51 clause 50.102.)

Serial Comparison

The Serial Comparison program compares ECGs over time, appending interpretive statements to the report generated by GE's Marquette 12SL analysis program. The Serial Comparison program is only available via the MUSE system and is described in the 12SL physician's guide.

The Serial Comparison program compares statements, measurements and waveforms.[2] The purpose of the program is to detect a significant clinical change and describe the change in terminology familiar to the cardiologist. Note that interpretive statements can change across serial ECGs, even though there is no significant clinical change in the ECGs. In this case, the program will not state a change.

The Serial Comparison program will compare ECGs that are analyzed by different versions of the 12SL program. This is because the Serial Comparison program re-analyzes historical ECGs. Furthermore, it compares the actual waveforms of the stored median complexes. However, it is critical this comparison be done on medians and fiducial point measurements generated by the same signal processing 12SL methodology, otherwise there will be a poor superimposition of the waveforms. This is important if an institution is going to compare and evaluate repolarization changes throughout the continuum of care, as recently demonstrated in a study that used 12SL measurements and waveforms to measure the potential significance of spontaneous and interventional ST-changes in patients transferred for primary percutaneous coronary intervention.[173]

GE Healthcare has developed specialized tools[85, 86, 174-178] for the collection, trending and comparison of serial 12-lead ECGs analyzed by the 12SL analysis program for the assessment of the acute coronary syndrome patient as they migrate from the prehospital setting through to intervention and the CCU.

Conclusion

This document has presented the performance of GE's Marquette 12SL analysis program. The evidence came from the scientific literature and it is, indeed, extensive. Nevertheless, gold standard data continues to be collected and the performance of the program evaluated.

Collection of data is an unending pursuit, for several reasons. The first, and most obvious, is that the program needs to be tested as improvements are made to it. However, equally important, is that new gold standards become available that can fundamentally change our understanding of the ECG. Sometimes, ECG criteria that are well accepted and have been used for decades can be rejected, as recently demonstrated for atrial enlargement.[179] In addition, changes in clinical practice, can change the meaning of a gold standard, as in the case of evaluating Q-waves in an environment of aggressive treatment for STEMI. Clinical practice can also alter the use of the ECG or generate new manifestations of the ECG, as in the case of artificial pacing. The challenge is to keep abreast of these changes and, yet, have an interpretive program that is understandable to the practicing physician.

GE is committed to continuous improvement of the program and obtaining the highest performance in the industry. GE recognizes that data collection is key to this improvement and, as a result, collaborates across the globe with several centers in the collecting of ECGs correlated with gold standard data or other clinical input. Given the capabilities of the MUSE system, most centers can investigate the performance of the program in a systematic fashion. GE welcomes this activity and is interested in collaborating with those who are equally committed to the advancement of computerized electrocardiography. Feel free to contact us with your comments and insights.

Bibliography

1. Rowlandson, I. “Computerized electrocardiography. A historical perspective”. *Annals of the New York Academy of Science*, 1990. 601: p. 343-352.
2. Rowlandson, I. “Strategy for Serial Comparison”. *Proceedings of the 1986 Engineering Foundation Conference, Computerized Interpretation of the Electrocardiogram XI*. New York: Engineering Foundation., 1986: p. 106-109.
3. Rowlandson, I. “New Techniques in Criteria Development”. *Proceedings of the 1985 Engineering Foundation Conference, Computerized Interpretation of the Electrocardiogram X, 1985* (New York: Engineering Foundation 1985): p. 177-184.
4. Davignon A, R.P. Boiselle E, et al., “Normal ECG standards for infants and children”. *Pediatric Cardiology*, 1979. 1: p. 123-52.
5. Rowlandson, I., P.J. Kudenchuk, and P.P. Elko. “Computerized recognition of acute infarction. Criteria advances and test results”. *Journal of Electrocardiology*, 1990. 23 Suppl: p. 1-5.
6. K903644 Series 1500 12-Lead Analysis Option – 510K. 1991, U.S. Food and Drug Administration, Center for Devices and Radiological Health.
7. Reddy, B.R., et al. “Data compression for storage of resting ECGs digitized at 500 samples/second”. *Biomed Instrum Technol*, 1992. 26(2): p. 133-149.
8. K921669 Marquette SL Series Transport Remote Acquisition – 510K. 1993, U.S. Food and Drug Administration, Center for Devices and Radiological Health.
9. Aufderheide, T.P., et al. “Test of the acute cardiac ischemia time-insensitive predictive instrument (ACI-TIPI) for prehospital use”. *Annals of Emergency Medicine*, 1996. 27(2): p. 193-198.
10. Aufderheide, T.P., et al. “QT dispersion and principal component analysis in prehospital patients with chest pain”. *Computers in Cardiology*, 1997: p. 665-668.
11. Xue, Q. and B.R. Reddy. “Clinical research workstation”. *U.S.P. Office, Editor*. 2002 GE Medical Systems Information Technologies, Inc. (Waukesha, WI): United States.
12. K981024 – QT dispersion and T wave analysis system (QT Guard System) – 510K. 1998, U.S. Food and Drug Administration, Center for Devices and Radiological Health.
13. Xue, Q. and S. Reddy. “Algorithms for computerized QT analysis”. *Journal of Electrocardiology*, 1998. 30 Suppl: p. 181-6.
14. Reddy, B.R., et al. “Prospective evaluation of a microprocessor-assisted cardiac rhythm algorithm: results from one clinical center”. *Journal of Electrocardiology*, 1998. 30 Suppl: p. 28-33.
15. Xue, J., et al. “A new method to incorporate age and gender into the criteria for the detection of acute inferior myocardial infarction”. *Journal of Electrocardiology*, 2001. 34 Suppl: p. 229-34.

16. Taha, B. and B.R. Reddy. “Method and apparatus for automatically detecting and interpreting paced electrocardiograms”. *U.S.P. Office, Editor*. 2001, GE Medical Systems Information Technologies, Inc. (Waukesha, WI): United States.

17. Batchvarov, V., K. Hnatkova, and M. Malik. “Assessment of noise in digital electrocardiograms”. *Pacing and Clinical Electrophysiology*, 2002. 25(4 Pt 1): p. 499-503.

18. Batchvarov, V. “Short and Long Term Reproducibility of the QT/RR Relationship in Healthy Subjects”. 2001. 37(2_Supp_A): p. 1A-648A.

19. Batchvarov, V. “Bazett Formula is not Suitable for Assessment of the Circadian Variation of the Heart Rate Corrected QT Interval”. 2001. 37(2_Supp_A): p. 1A-648A.

20. Hnatkova, K., et al. “Precision of QT Interval Measurement by Advanced Electrocardiographic Equipment”. *Pacing and Clinical Electrophysiology*, 2006. 29(11): p. 1277-1284.

21. Farrell, R. and B. Young. “Effect of Lead Quality on Computerized ECG Interpretation”. *Computers in Cardiology*, 2004. 31: p. 173-176.

22. K042782 – SEER MC – 510K. 2005, U.S. Food and Drug Administration, Center for Devices and Radiological Health.

23. K060833 – 12SL ECG Analysis Program; Computerized detection of right ventricular involvement in acute inferior myocardial infarction – 510K. 2006, U.S. Food and Drug Administration, Center for Devices and Radiological Health.

24. de Bruyne, M.C., et al. “Diagnostic interpretation of electrocardiograms in population-based research: computer program research physicians, or cardiologists?” *Journal of Clinical Epidemiology*, 1997. 50(8): p. 947-952.

25. Savelieva, I., et al. “Agreement and reproducibility of automatic versus manual measurement of QT interval and QT dispersion”. *The American Journal of Cardiology*, 1998. 81(4): p. 471-477.

26. Okin, P.M., et al. “Assessment of QT interval and QT dispersion for prediction of all-cause and cardiovascular mortality in American Indians: The Strong Heart Study”. *Circulation*, 2000. 101(1): p. 61-66.

27. International Standard IEC 60601-2-51:2003 “Medical electrical equipment. Particular requirements for safety, including essential performance, of recording and analysing single channel and multichannel electrocardiographs”. *I.E.C. (IEC), Editor*. 2003, International Electrotechnical Commission (IEC). p. 86.

28. Willems, J.L., et al. “Effect of combining electrocardiographic interpretation results on diagnostic accuracy”. *European Heart Journal*, 1988. 9(12): p. 1348-1355.

29. Woolley, D., M. Henck, and J. Luck. “Comparison of electrocardiogram interpretations by family physicians, a computer, and a cardiology service”. *The Journal of Family Practice*, 1992. 34(4): p. 428-432.

30. Crevasse, L. and M. Ariet. “Clinical usefulness of computerized ECG systems”. *Journal of Medical Systems*, 1987. 11(1): p. 21-24.

31. Brailer, D.J., E. Kroch, and M.V. Pauly. “The impact of computer-assisted test interpretation on physician decision making: the case of electrocardiograms”. *Medical Decision Making*, 1997. 17(1): p. 80-86.

32. Goodacre, S., A. Webster, and F. Morris. “Do computer generated ECG reports improve interpretation by accident and emergency senior house officers?” *Postgraduate Medical Journal*, 2001. 77(909): p. 455-457.

33. Tsai, T.L., D.B. Fridsma, and G. Gatti. “Computer decision support as a source of interpretation error: the case of electrocardiograms”. *Journal of the American Medical Informatics Association*, 2003. 10(5): p. 478-483.

34. Huffman, D. “A Method for the Construction of Minimum-Redundancy Codes”. *Proceedings of the Insitute of Radio Engineers*, 1952. 40: p. 1098-1110.

35. Zywietz, C., J.L. Willems, and R. Degani. “Data compression for computerized electrocardiography”. in *Proc of the 1st Working Conf of the SCP-ECG Project AIM 1015*. 1990. Belgium: Leuven ACCO Press.

36. ANSI/AAMI, Standard communications protocol – Computer assisted electrocardiography, EC71:2001, Editor. 2002. p. 156.

37. JA, A., D. LS, and A. M, 10th Bethesda Conference: optimal electrocardiography. *The American Journal of Cardiology*, 1978. 41: p. 111-191.

38. Jain, A.K., R.P.W. Duin, and J. Mao. “Statistical Pattern Recognition: A Review”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000. 22:1: p. 4-37.

39. Zywietz, C., D. Celikag, and G. Joseph. “Influence of ECG measurement accuracy on ECG diagnostic statements”. *Journal of Electrocardiology*, 1996. 29 Suppl: p. 67-72.

40. Willems, J.L., et al. “A reference data base for multilead electrocardiographic computer measurement programs”. *Journal of the American College of Cardiology*, 1987. 10(6): p. 1313-1321.

41. ANSI/AAMI, Diagnostic electrocardiographic devices, EC11:1991/(R)2001, Editor. 1992. p. 39.

42. Bailey, J.J., 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): p. 730-739.

43. Bogun, F., et al. “Misdiagnosis of atrial fibrillation and its clinical consequences”. *The American Journal of Medicine*, 2004. 117(9): p. 636-642.

44. Ahlstrom, M.L. and W.J. Tompkins. “Digital filters for real-time ECG signal processing using microprocessors”. *IEEE Trans Biomed Eng*, 1985. 32(9): p. 708-713.

45. Brodnick, D.E. “Method of removing signal interference from sampled data and apparatus for effecting the same”. U.S.P. Office, Editor. 2002, GE Medical Systems Information Technologies, Inc. (Waukesha, WI) United States.

46. Tayler, D.I. and R. Vincent. “Artefactual ST segment abnormalities due to electrocardiograph design”. *British Heart Journal*, 1985. 54(2): p. 121-8.

47. Schmitt, O. “Inter- and Intra-Individual Variability in ECG Electrode-Skin Impedance”. *Biophysical Society Abstracts*, 1970. Feb. 25-27.

48. Wiese, S.R., et al. “Electrocardiographic motion artifact versus electrode impedance”. *IEEE Trans Biomed Eng*, 2005. 52(1): p. 136-139.

49. Willems, J.L., et al. “Influence of noise on wave boundary recognition by ECG measurement programs. Recommendations for preprocessing”. *Computers and Biomedical Research*, 1987. 20(6): p. 543-562.

50. Zywietz, C., et al. “Stability of computer ECG amplitude measurements in the presence of noise. The CSE Working Party”. *Computers and Biomedical Research*, 1990. 23(1): p. 10-31.

51. Reddy, B.R., Q. Xue, and C. Zywietz. “Analysis of interval measurements on CSE multilead reference ECGs”. *Journal of Electrocardiology*, 1996. 29 Suppl: p. 62-66.

52. Farrell, R. and G.I. Rowlandson. “The Effects of Noise on Computerized ECG Measurements”. *Journal of Electrocardiology*, 2006.

53. “Recommendations for measurement standards in quantitative electrocardiography. The CSE Working Party”. *European Heart Journal*, 1985. 6(10): p. 815-825.

54. Willems, J.L., et al. “Common standards for quantitative electrocardiography: goals and main results. CSE Working Party”. *Methods of Information in Medicine*, 1990. 29(4): p. 263-271.

55. Willems, J.L., et al. “Assessment of the performance of electrocardiographic computer programs with the use of a reference data base”. *Circulation*, 1985. 71(3): p. 523-534.

56. Moody, G.B., M. WI, and M. RG. “A noise stress test for arrhythmia detectors”. *Computers in Cardiology*. 1984. Long Beach (CA): IEEE Computer Society.

57. ANSI/AAMI, Ambulatory electrocardiographs, EC38:1998, Editor. 1999. p. 58.

58. Mulcahy, D., et al. “Can a computer assisted electrocardiograph replace a cardiologist for ECG measurements?” *Irish Journal of Medical Science*, 1986. 155(12): p. 410-414.

59. Graham. “User Evaluation of a Commercially Available Computerized Electrocardiographic Interpretation System”. in *Computer ECG analysis – towards standardization : proceedings of the IFIP-IMIA Working Conference on Computer ECG Analysis, Towards Standardization*, J.L. Willems, J.H. VanBemmel, and C. Zywietz, Editors. 1986. p. 399.

60. Kors, J.A. and G. van Herpen. “The coming of age of computerized ECG processing: can it replace the cardiologist in epidemiological studies and clinical trials?” *Medinfo*, 2001. 10(Pt 2): p. 1161-7.

61. Kalahasti, V., et al. “QRS duration and prediction of mortality in patients undergoing risk stratification for ventricular arrhythmias”. *The American Journal of Cardiology*, 2003. 92(7): p. 798-803.

62. Dhingra, R., et al. “Electrocardiographic QRS duration and the risk of congestive heart failure: the Framingham Heart Study”. *Hypertension*, 2006. 47(5): p. 861-867.

63. Turner, M.S., et al. “Electrical and Mechanical Components of Dyssynchrony in Heart Failure Patients With Normal QRS Duration and Left Bundle-Branch Block: Impact of Left and Biventricular Pacing”. *Circulation*, 2004. 109(21): p. 2544-2549.

64. Sanderson, J.E. “Assessment of ventricular dyssynchrony: global or regional function?” *European Heart Journal*, 2006. 27(20): p. 2380-2381.

65. Bleeker, G.B., et al. “Relationship Between QRS Duration and Left Ventricular Dyssynchrony in Patients with End-Stage Heart Failure”. *Journal of Cardiovascular Electrophysiology*, 2004. 15(5): p. 544-549.

66. Murkofsky, R.L., et al. “A prolonged QRS duration on surface electrocardiogram is a specific indicator of left ventricular dysfunction”. *Journal of the Americal College of Cardiology*, 1998. 32(2): p. 476-482.

67. Elhendy, A., et al. “Relation of QRS duration on the surface 12-lead electrocardiogram with mortality in patients with known or suspected coronary artery disease”. *The American Journal of Cardiology*, 2005. 96(8): p. 1082-1088.

68. Luliano, S., et al. “QRS duration and mortality in patients with congestive heart failure”. *American Heart Journal*, 2002. 143(6): p. 1085-1091.

69. Shenkman, H.J., et al. “Congestive heart failure and QRS duration: establishing prognosis study”. *Chest*, 2002. 122(2): p. 528-534.

70. Desai, A.D., et al. “Prognostic Significance of Quantitative QRS Duration”. *The American Journal of Medicine*, 2006. 119(7): p. 600-606.

71. Okin, P.M., et al. “Computerized ST depression analysis improves prediction of all-cause and cardiovascular mortality: the strong heart study”. *Annals of Noninvasive Electrocardiology*, 2001. 6(2): p. 107-116.

72. Okin, P.M., et al. “Quantitative assessment of electrocardiographic strain predicts increased left ventricular mass: the Strong Heart Study”. *Journal of the American College of Cardiology*, 2002. 40(8): p. 1395-1400.

73. Eskola, M.J., et al. “Comparative accuracy of manual versus computerized electrocardiographic measurement of J-, ST- and T-wave deviations in patients with acute coronary syndrome”. *The American Journal of Cardiology*, 2005. 96(11): p. 1584-1588.

74. Guidance for Industry: E14 Clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs. ICH, Editor. 2005, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH), Food and Drug Administration (FDA), Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER).

75. Malik, M., et al. “Sample size, power calculations, and their implications for the cost of thorough studies of drug induced QT interval prolongation”. *Pacing and Clinical Electrophysiology*, 2004. 27(12): p. 1659-1669.

76. Darpo, B., et al. “Man versus machine: is there an optimal method for QT measurements in thorough QT studies?” *The Journal of Clinical Pharmacology*, 2006. 46(6): p. 598-612.

77. Rautaharju, P.M., et al. “Normal standards for QT and QT subintervals derived from a large ethnically diverse population of women aged 50 to 79 years (the Women's Health Initiative [WHI])”. *The American Journal of Cardiology*, 2006. 97(5): p. 730-737.

78. Dmitrienko, A.A., et al. “Electrocardiogram Reference Ranges Derived from a Standardized Clinical Trial Population”. *Drug Information Journal*, 2005. 39: p. 395-405.

79. Xue, Q. and B.R. Reddy. “New algorithms for QT dispersion analysis”. *Computers in Cardiology*. 1996. Indianapolis, IN, USA.

80. Okin, P.M., et al. “Principal Component Analysis of the T Wave and Prediction of Cardiovascular Mortality in American Indians: The Strong Heart Study”. *Circulation*, 2002. 105(6): p. 714-719.

81. Froelicher, V., R. Marcus, and P. Heidenreich. “Prognostic Value of Computer Electrocardiography in Veteran Outpatients”. *Federal Practitioner*, 2004. 21(3): p. 11-12, 17-20.

82. Zabel, M., et al. “Assessment of QT dispersion for prediction of mortality or arrhythmic events after myocardial infarction: results of a prospective, long-term follow-up study”. *Circulation*, 1998. 97(25): p. 2543-2550.

83. Zabel, M., et al. “Analysis of T-wave morphology from the 12-lead electrocardiogram for prediction of long-term prognosis in male US veterans”. *Circulation*, 2002. 105(9): p. 1066-1070.

84. Potter, L., et al. “QT Dispersion in Acute Myocardial Infarction: A Tool for Identification of Patients with Initially Normal Electrocardiograms”. *Academic Emergency Medicine*, 1999. 6:5: p. 424.

85. Kesek, M., et al. “Non-dipolar content of the T-wave as a measure of repolarization inhomogeneity in ST-elevation myocardial infarction”. *Clinical Physiology and Functional Imaging*, 2006. 26(6): p. 362-370.

86. Kesek, M., et al. “Principal Component Analysis of the T Wave in Patients with Chest Pain and Conduction Disturbances”. *Pacing and Clinical Electrophysiology*, 2004. 27(10): p. 1378-1387.

87. Kowalski, M., et al. “Temporal Relationship between QTc Prolongation and ST-segment Elevation in Early Transmural Myocardial Ischemia Measured by a Continuous 12-lead ECG Monitor”. *American Heart Association*. 2006

88. Miller, M.D., C. Porter, and M.J. Ackerman. “Diagnostic accuracy of screening electrocardiograms in long QT syndrome”. *Pediatrics*, 2001. 108(1): p. 8-12.

89. Snyder, C.S., et al. “The emergency department versus the computer: which is the better electrocardiographer?” *Pediatric Cardiology*, 2003. 24(4): p. 364-368.

90. Grunkemeier, G.L. and N. Payne. “Bayesian analysis: A new statistical paradigm for new technology”. *The Annals of Thoracic Surgery*, 2002. 74(6): p. 1901-1908.

91. Mieres, J.H., et al. “Role of Noninvasive Testing in the Clinical Evaluation of Women With Suspected Coronary Artery Disease: Consensus Statement From the Cardiac Imaging Committee, Council on Clinical Cardiology, and the Cardiovascular Imaging and Intervention Committee, Council on Cardiovascular Radiology and Intervention, American Heart Association”. *Circulation*, 2005. 111(5): p. 682-696.

92. Reddy, B.R., et al. “Detection of P waves in resting ECG: A preliminary study”. Los Alamitos, CA: IEEE Computer Society Press, 1992. *Computers in Cardiology*, 1992: p. 87-90.

93. Reddy, B.R., P.P. Elko, and G.I. Rowlandson. “Contextual information in atrial arrhythmia analysis”. *Journal of Electrocardiology*, 1993. 26 Suppl:100.

94. Reddy, B.R., et al. “Detection of complex atrial arrhythmias in resting ECG”. Los Alamitos, CA: IEEE Computer Society Press, 1994. *Computers in Cardiology*, 1994: p. 777-780.

95. Reddy, B.R., P. Elko, and S. Swiryn. “A new arrhythmia analysis program, MAC-RHYTHM, for resting ECG. in International Congress on Electrocardiology. 1996. Singapore.

96. Takatori. “Arrhythmia diagnosis by the Marquette 12SL ECG analysis program. Computerized interpretation of the electrocardiogram”. *Proceedings of the Engineering Foundation Conference*, 1983. VII: p. 63-74.

97. Swiryn, S. and J. Jenkins. “Arrhythmia diagnosis by computer. The Marquette 12SL program. Computerized interpretation of the electrocardiogram”. *Proceedings of the Engineering Foundation Conference*, 1983: p. 35-54.

98. Farrell, R., J. Xue, and B. Young. “Enhanced Rhythm Analysis for Resting ECG Using Spectral and Time-Domain Techniques”. *Computers in Cardiology*, 2003. 30: p. 733-736.

99. Poon, K., P.M. Okin, and P. Kligfield. “Diagnostic performance of a computer-based ECG rhythm algorithm”. *Journal of Electrocardiology*, 2005. 38(3): p. 235-238.

100. Guglin, M.E. and D. Thatai. “Common errors in computer electrocardiogram interpretation”. *International Journal of Cardiology*, 2006. 106(2): p. 232-237.

101. Bossone, E., et al. “The interpretation of the electrocardiogram in patients with pulmonary hypertension: the need for clinical correlation”. *Italian Heart Journal*, 2003. 4(12): p. 850-854.

102. Hongo, R.H. and N. Goldschlager. “Overreliance on computerized algorithms to interpret electrocardiograms”. *The American Journal of Medicine*, 2004. 117(9): p. 706-708.

103. Nessmith Jr, M.G., T.W. Simmons, and W.K. Haisty Jr. “Computerized Detection of Paced Rhythms in the Resting 12-Lead Electrocardiogram”. 2001, Wake Forest University Medical Center.

104. Kazi W., M. and R. Mauerman S.K. “New algorithm improves detection accuracy of paced rhythms in the resting ECG”. 2002, Division of Cardiology/Electrophysiology. Medical College of Wisconsin-Milwaukee.

105. Guglin, M.E. and N. Datwani. “Electrocardiograms with pacemakers: accuracy of computer reading”. *Journal of Electrocardiology*, 2006. In press. Corrected proof.

106. Chiu, C.C., et al. “Evaluation of computerized interpretation of the pediatric electrocardiogram”. *Journal of Electrocardiology*, 2006.

107. Swartz, M.H. and L.E. Teichholz. “Marquette 12SL ECG analysis program: evaluation of physician changes”. *Computers in Cardiology*, 1982: p. 437-440.

108. Hammill, S.C. “Comparison of a Computerized Interpretive ECG Program Using XYZ Leads with a Program using Scalar Leads”. *Journal of Electrocardiology*, 1988. 88(Supp).

109. Hammill, S. “Accuracy of automated electrocardiographic detection as performed at the Mayo Clinic; personal communication”. 1989.

110. Willems, J.L., et al. “The diagnostic performance of computer programs for the interpretation of electrocardiograms”. *The New England Journal of Medicine*, 1991. 325(25): p. 1767-73.

111. Schwartz, P.J., et al. “Guidelines for the interpretation of the neonatal electrocardiogram. A task force of the European Society of Cardiology”. *European Heart Journal*, 2002. 23(17): p. 1329-1344.

112. Alpert, J.S., et al. “Myocardial infarction redefined – a consensus document of The Joint European Society of Cardiology/American College of Cardiology Committee for the redefinition of myocardial infarction”. *Journal of the American College of Cardiology*, 2000. 36(3): p. 959-969.

113. Pope, J.E. “Performance of the Marquette 12SL ECG analysis program. Computerized interpretation of the electrocardiogram”. *Proceedings of the Engineering Foundation Conference*, 1986. X: p. 185-193.

114. Haisty, W.K. “Multivariate Q-Wave Criteria for Inferior Myocardial Infarction”. *Computerized Interp. of the ECG*. 1987.

115. Elko, P. and R.A. Warner. “Using directly acquired digital ECG data to optimize the diagnostic criteria for anterior myocardial infarction”. *Journal of Electrocardiology*, 1994. 27 Suppl: p. 10-13.

116. Selvester, R.H., G.S. Wagner, and N.B. Hindman. “The Selvester QRS scoring system for estimating myocardial infarct size. The development and application of the system”. *Archives of Internal Medicine*, 1985. 145(10): p. 1877-1881.

117. Bounous, E.P., Jr., et al. “Prognostic value of the simplified Selvester QRS score in patients with coronary artery disease”. *Journal of the American College of Cardiology*, 1988. 11(1): p. 35-41.

118. Rautaharju, P.M., et al. “Cardiac infarction injury score: an electrocardiographic coding scheme for ischemic heart disease”. *Circulation*, 1981. 64(2): p. 249-256.

119. Pope, J.E., et al. “Development and validation of an automated method of the Selvester QRS scoring system for myocardial infarct size”. *The American Journal of Cardiology*, 1988. 61(10): p. 734-738.

120. Haisty, W.K., Jr., et al. “Performance of the automated complete Selvester QRS scoring system in normal subjects and patients with single and multiple myocardial infarctions”. *Journal of the American College of Cardiology*, 1992. 19(2): p. 341-346.

121. Johanson, P., et al. “A dynamic model forecasting myocardial infarct size before, during, and after reperfusion therapy: an ASSENT-2 ECG/VCG substudy”. *European Heart Journal*, 2005. 26(17): p. 1726-1733.

122. Khan, M.A., et al. “A normal electrocardiogram precludes the need for left ventriculography in the assessment of coronary artery disease”. *Heart*, 1998. 79(3): p. 262-267.

123. Richardson, K., et al. “Electrocardiographic damage scores and cardiovascular mortality”. *American Heart Journal*, 2005. 149(3): p. 458-463.

124. Engblom, H., et al. “Size and transmural extent of first-time reperfused myocardial infarction assessed by cardiac magnetic resonance can be estimated by 12-lead electrocardiogram”. *American Heart Journal*, 2005. 150(5): p. 920.

125. Wagner, G.S., et al. “A method for assembling a collaborative research team from multiple disciplines and academic centers to study the relationships between ECG estimation and MRI measurement of myocardial infarct size”. *Journal of Electrocardiology*, 2001. 34 Suppl: p. 1-6.

126. Engel, G., et al. “Electrocardiographic arrhythmia risk testing”. *Current Problems in Cardiology*, 2004. 29(7): p. 365-432.

127. Engblom, H., et al. “Quantitative clinical assessment of chronic anterior myocardial infarction with delayed enhancement magnetic resonance imaging and QRS scoring”. *American Heart Journal*, 2003. 146(2): p. 359-366.

128. Antman, E.M., et al. “ACC/AHA guidelines for the management of patients with ST-elevation myocardial infarction: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Revise the 1999 Guidelines for the Management of Patients with Acute Myocardial Infarction)”. *Circulation*, 2004. 110(9): p. e82-292.

129. Garvey, J.L., et al. “Pre-Hospital 12-Lead Electrocardiography Programs: A Call for Implementation by Emergency Medical Services Systems Providing Advanced Life Support – National Heart Attack Alert Program (NHAAP) Coordinating Committee; National Heart, Lung, and Blood Institute (NHLBI); National Institutes of Health”. *Journal of the American College of Cardiology*, 2006. In press. Corrected proof.

130. Grim, P., et al. “Cellular telephone transmission of 12-lead electrocardiograms from ambulance to hospital”. *The American Journal of Cardiology*, 1987. 60(8): p. 715-720.

131. Aufderheide, T.P., et al. “A prospective evaluation of prehospital 12-lead ECG application in chest pain patients”. *Journal of Electrocardiology*, 1992. 24 Suppl: p. 8-13.

132. Kereiakes, D.J., et al. “Time delays in the diagnosis and treatment of acute myocardial infarction: a tale of eight cities. Report from the Pre-hospital Study Group and the Cincinnati Heart Project”. *American Heart Journal*, 1990. 120(4): p. 773-780.

133. Foster, D.B., et al. “Prehospital recognition of AMI using independent nurse/paramedic 12-lead ECG evaluation: impact on in-hospital times to thrombolysis in a rural community hospital”. *The American Journal of Emergency Medicine*, 1994. 12(1): p. 25-31.

134. Millar-Craig, M.W., et al. “Reduction in treatment delay by paramedic ECG diagnosis of myocardial infarction with direct CCU admission”. *Heart*, 1997. 78(5): p. 456-461.

135. Aufderheide, T.P., et al. “The diagnostic impact of prehospital 12-lead electrocardiography”. *Annals of Emergency Medicine*, 1990. 19(11): p. 1280-1287.

136. Weaver, W.D., et al. 'Myocardial Infarction Triage and Intervention Project – phase I: patient characteristics and feasibility of prehospital initiation of thrombolytic therapy”. *Journal of the American College of Cardiology*, 1990. 15(5): p. 925-931.

137. Elko, P.P. and I. Rowlandson. “A statistical analysis of the ECG measurements used in computerized interpretation of acute anterior

myocardial infarction with applications to interpretive criteria development". *Journal of Electrocardiology*, 1992. 25 Suppl: p. 113-119.

138. Elko, P.P., et al. "The dilemma of sensitivity versus specificity in computer-interpreted acute myocardial infarction". *Journal of Electrocardiology*, 1992. 24 Suppl: p. 2-7.

139. Tighe, M., et al. "The early diagnosis of acute myocardial infarction. Comparison of a simple algorithm with a computer program for electrocardiogram interpretation". *Irish Journal of Medical Science*, 1996. 165(3): p. 159-163.

140. Otto, L.A. and T.P. Aufderheide. "Evaluation of ST segment elevation criteria for the prehospital electrocardiographic diagnosis of acute myocardial infarction". *Annals of Emergency Medicine*, 1994. 23(1): p. 17-24.

141. O'Rourke, M.F., et al. "Accuracy of a portable interpretive ECG machine in diagnosis of acute evolving myocardial infarction". *Australian and New Zealand Journal of Medicine*, 1992. 22(1): p. 9-13.

142. K99190 – Lifepak 12 – 510K. 1999, U.S. Food and Drug Administration, Center for Devices and Radiological Health.

143. K991556 – Zoll M Series 12SL Option – 510K. 1999, U.S. Food and Drug Administration, Center for Devices and Radiological Health.

144. K002209 – GE Marquette ECG Analysis Program – 510K. 2002, U.S. Food and Drug Administration, Center for Devices and Radiological Health.

145. Garza, M.A. "Cellular technology. Revolutionizing EMS communications". Jems, 1990. 15(5): p. 46-7, 50-3, 55-56.

146. Ewy, G.A. and J.P. Ornato. "Emergency cardiac care: introduction". *Journal of the American College of Cardiology*, 2000. 35: p. 825-880.

147. Martin, J.S., P.E. Litwin, and W.D. Weaver. "Early recognition and treatment of the patient suffering from acute myocardial infarction. A description of the Myocardial Infarction Triage and Intervention Project". *Critical Care Nursing Clinics of North America*, 1990. 2(4): p. 681-688.

148. Kudenchuk, P.J., et al. "Accuracy of computer-interpreted electrocardiography in selecting patients for thrombolytic therapy. MITI Project Investigators". *Journal of the American College of Cardiology*, 1991. 17(7): p. 1486-1491.

149. Bradley, E.H., et al. "Strategies for Reducing the Door-to-Balloon Time in Acute Myocardial Infarction". *The New England Journal of Medicine*, 2006.

150. McClelland, A.J., et al. "Comparison of the 80-lead body surface map to physician and to 12-lead electrocardiogram in detection of acute myocardial infarction". *The American Journal of Cardiology*, 2003. 92(3): p. 252-257.

151. Braunwald, E., et al. "ACC/AHA 2002 guideline update for the management of patients with unstable angina and non-ST-segment elevation myocardial infarction – summary article: a report of the American College of Cardiology/American Heart Association task force on practice guidelines (Committee on the Management of Patients With Unstable Angina)". *Journal of the American College of Cardiology*, 2002. 40(7): p. 1366-1374.

152. Massel, D., J.A. Dawdy, and L.J. Melendez. "Strict reliance on a computer algorithm or measurable ST segment criteria may lead to errors in

thrombolytic therapy eligibility". *American Heart Journal*, 2000. 140(2): p. 221-226.

153. Ornato, J.P. "Chest pain emergency centers: improving acute myocardial infarction care". *Clinical Cardiology*, 1999. 22(8 Suppl): p. IV3-9.

154. Wright, R.S. "Women with Acute Anterior Myocardial Infarction Have Less Precordial ST Elevation Than Men Independent of Age of Presentation". 37(2001): 361A. . *Journal of the American College of Cardiology*, 2001. 37(2001): 361A.

155. Wellens, H.J. "The value of the right precordial leads of the electrocardiogram". *The New England Journal of Medicine*, 1999. 340(5): p. 381-383.

156. Zehender, M., et al. "Right ventricular infarction as an independent predictor of prognosis after acute inferior myocardial infarction". *The New England Journal of Medicine*, 1993. 328(14): p. 981-988.

157. Saw, J., et al. "Value of ST elevation in lead III greater than lead II in inferior wall acute myocardial infarction for predicting in-hospital mortality and diagnosing right ventricular infarction". *The American Journal of Cardiology*, 2001. 87(4): p. 448-450, A6.

158. Wong, T.W., et al. "New electrocardiographic criteria for identifying the culprit artery in inferior wall acute myocardial infarction-usefulness of T-wave amplitude ratio in leads II/III and T-wave polarity in the right V5 lead". *The American Journal of Cardiology*, 2004. 94(9): p. 1168-1171.

159. Zimetbaum, P.J., et al. "Usefulness of ST-segment elevation in lead III exceeding that of lead II for identifying the location of the totally occluded coronary artery in inferior wall myocardial infarction". *The American Journal of Cardiology*, 1998. 81(7): p. 918-919.

160. Braat, S.H., et al. "Right ventricular involvement with acute inferior wall myocardial infarction identifies high risk of developing atrioventricular nodal conduction disturbances". *American Heart Journal*, 1984. 107(6): p. 1183-1187.

161. Kinch, J.W. and T.J. Ryan. "Right ventricular infarction". *The New England Journal of Medicine*, 1994. 330(17): p. 1211-1217.

162. Rodrigues, E.A., et al. "Diagnosis and prognosis of right ventricular infarction". *British Heart Journal*, 1986. 56(1): p. 19-26.

163. Morgera, T., et al. "Right precordial ST and QRS changes in the diagnosis of right ventricular infarction". *American Heart Journal*, 1984. 108(1): p. 13-18.

164. Zalenski, R.J., et al. "Assessing the diagnostic value of an ECG containing leads V4R, V8, and V9: the 15-lead ECG". *Ann Emerg Med*, 1993. 22(5): p. 786-93.

165. Zalenski, R.J., et al. "Value of posterior and right ventricular leads in comparison to the standard 12-lead electrocardiogram in evaluation of ST-segment elevation in suspected acute myocardial infarction". *The American Journal of Cardiology*, 1997. 79(12): p. 1579-85.

166. Taha, B., et al. "Normal limits of ST segment measurements in posterior ECG leads". *Journal of Electrocardiology*, 1998. 31 Suppl: p. 178-9.

167. Farrell, R.M. and I. Rowlandson. “Computerized detection of right ventricular involvement in acute inferior infarction”. *Journal of Electrocardiology*, 2006. 39(4, Supplement 1): p. S128-S129.

168. Selker, H.P., J.L. Griffith, and R.B. D’Agostino. “A tool for judging coronary care unit admission appropriateness, valid for both real-time and retrospective use. A time-insensitive predictive instrument (TIPI) for acute cardiac ischemia: a multicenter study”. *Med Care*, 1991. 29(7): p. 610-27.

169. Pozen, M.W., et al. “A predictive instrument to improve coronary-care-unit admission practices in acute ischemic heart disease. A prospective multicenter clinical trial”. *The New England Journal of Medicine*, 1984. 310(20): p. 1273-8.

170. Selker, H.P., R.B. D’Agostino, and M.M. Laks. “A predictive instrument for acute ischemic heart disease to improve coronary care unit admission practices: a potential on-line tool in a computerized electrocardiograph”. *Journal of Electrocardiology*, 1988. 21 Suppl: p. S11-7.

171. Selker, H.P., et al. “Use of the acute cardiac ischemia time-insensitive predictive instrument (ACI-TIPI) to assist with triage of patients with chest pain or other symptoms suggestive of acute cardiac ischemia. A multicenter, controlled clinical trial”. *Annals of Internal Medicine*, 1998. 129(11): p. 845-55.

172. Saner, H., et al. “Comparison of 3 ECG computer programs with interpretation by physicians”. *Schweiz Med Wochenschr*, 1987. 117(27-28): p. 1035-9.

173. Terkelsen, C.J., et al. “Potential significance of spontaneous and interventional ST-changes in patients transferred for primary percutaneous coronary intervention: observations from the ST-Monitoring in Acute Myocardial Infarction study (The MONAMI study)”. *European Heart Journal*, 2006. 27(3): p. 267-275.

174. Fesmire, F.M. “Which chest pain patients potentially benefit from continuous 12-lead ST-segment monitoring with automated serial ECG?” *The American Journal of Emergency Medicine*, 2000. 18(7): p. 773-8.

175. Jernberg, T., B. Lindahl, and L. Wallentin. “The combination of a continuous 12-lead ECG and troponin T; a valuable tool for risk stratification during the first 6 hours in patients with chest pain and a non-diagnostic ECG”. *European Heart Journal*, 2000. 21(17): p. 1464-72.

176. Zairis, M.N., et al. “C-reactive protein and ST-segment monitoring by continuous 12-lead electrocardiogram in patients with primary unstable angina pectoris”. *The American Journal of Cardiology*, 2003. 91(5): p. 600-3.

177. Zairis, M.N., et al. “Continuous 12-lead electrocardiographic ST monitoring adds prognostic information to the thrombolysis in myocardial infarction risk score in patients with non-ST-elevation acute coronary syndromes”. *Clinical Cardiology*, 2005. 28(4): p. 189-92.

178. Zairis, M.N., et al. “Continuous 12-lead electrocardiographic ST monitoring and long-term prognosis after successful coronary stenting”. *American Heart Journal*, 2006. 151(4): p. 892-7.

179. Lee, K.S., et al. “Relation of electrocardiographic criteria for left atrial enlargement to two-dimensional echocardiographic left atrial volume measurements”. *The American Journal of Cardiology*, 2007. 99(1): p. 113-8.

For your notes

World Headquarters
GE Medical Systems
Information Technologies, Inc.
8200 West Tower Avenue
Milwaukee, WI 53223 USA
Tel: + 1 414 355 5000
1 800 558 5120 (US only)
Fax: + 1 414 355 3790

European Representative
GE Medical Systems
Information Technologies GmbH
Munzinger Straße 3-5
D-79111 Freiburg
Germany
Tel: + 49 761 45 43 - 0
Fax: + 49 761 45 43 - 233

Asia Headquarters
GE Medical Systems
Information Technologies Asia; GE (China) Co., Ltd.
24th Floor, Shanghai MAXDO Center,
8 Xing Yi Road, Hong Qiao Development Zone
Shanghai 200336, P.R. China
Tel: + 86 21 5257 4650
Fax: + 86 21 5208 2008

www.gehealthcare.com