Effect of Lead Quality on Computerized ECG Interpretation

RM Farrell, BJ Young

GE Healthcare, Milwaukee, USA

Abstract

Electrocardiograms of poor technical quality present challenges to accurate interpretation. Poor technical quality ECGs may result from muscle tremor, AC power interference, electrode motion, or baseline shifts.

A system which assesses the ECG lead quality is described. An analysis algorithm assessed the ECG for artifact and assigned the ECG a lead quality level of green (good), yellow (marginal), or red (poor).

A total of 120,698 physician-confirmed ECGs were retrospectively analyzed by the lead quality algorithm and the GE Healthcare 12SL analysis program. The discordance rates (percentage of ECGs in which the rhythm from the 12SL analysis differed from the rhythm of the confirmed ECG) were reported by lead quality level. While less than 5% of the ECGs were of marginal or poor lead quality, the discordance rate was found to markedly increase as the lead quality degraded.

1. Introduction

Electrocardiograms (ECGs) of poor technical quality present challenges to accurate interpretation by electrocardiographers and by computerized interpretation programs. Artifact appearing on the ECG may be due to physiologic or non-physiologic sources. Physiologic sources include muscle tremor, shivering, hiccups, and baseline wander caused by respiration. Non-physiologic sources of artifact include electrical interference and electrode movements. Electrical interference is most often caused by 50 or 60 Hz alternating current (AC) power sources, although other electric or electronic devices may also contribute to the interference. Poor electrode contact may produce baseline instabilities, which may occur gradually or abruptly [1]. Large swings in the ECG baseline can occur with position changes, coughing, ambulation, or movement of the leadwires. These effects originate from the site of the electrode contact with the skin. As the epidermis is stretched at the site of the electrode, a change of up to several millivolts may be observed [2].

Any of these types of artifact may be present on any combination of leads, depending on the electrodes involved. Artifact affecting only the left arm electrode, for example, will have the greatest effect on lead I, but will not affect lead II, while all other leads will be somewhat affected due to the relation between this electrode and the derivation of those leads [3]. On the other hand, artifact that affects only a single chest electrode will manifest itself only on its corresponding V lead.

The most common artifacts are caused by skeletal muscle tremor, electrical interference, and electrode movements [4]. Artifact may mimic pathologic electrocardiographic findings, ranging from arrhythmia to ischemic change [2,5,6]. The American College of Cardiology and the American Heart Association advise of the importance of minimizing and of recognizing patient related artifact such as muscle tremor and movement [7].

Failure to minimize and recognize artifact while recording, and failure to recognize artifacts during interpretation, may result in an incorrect diagnosis of arrhythmias and may lead to unnecessary interventions and treatment [5,7].

2. Methods

A new analysis algorithm assessed the ECG for artifact and assigned the ECG a lead quality level of green, yellow, or red. Green indicated generally acceptable lead quality, yellow indicated artifact on one or more leads, and red denoted electrode disconnections or deflections that significantly exceeded the expected range of the normal ECG.

Physician-confirmed ECGs were analyzed by the lead quality algorithm and the GE Healthcare 12SL resting ECG analysis program. The 12SL rhythm analysis performance was assessed as a function of the ECG lead quality.

2.1. Lead quality analysis

Leads I, II, and V1 – V6 of each ECG were analyzed for muscle tremor, baseline wander, powerline (AC) interference, and electrode noise. Leads III, aVR, aVL, and aVF were not analyzed because these leads are not acquired, rather they are derived from leads I and II using the standard equations of Einthoven and Goldberger [1,4].

Muscle tremor artifact was detected by counting the number of deflections exceeding a fixed threshold per

second. If this count exceeded a limit in any one-second window, then the lead was said to contain muscle artifact.

Powerline, or AC, interference was detected by running a "frequency hunting" filter over each lead of the 10-second ECG [8]. This filter was an adaptive filter that learned both the frequency and the phase of the interference signal. In this case, the adaptive filter only learned the interference signal, but did not modify the input signal, as any user-configurable AC filtering had previously been completed before this stage. The method had the advantage of assessing AC interference on the actual signal that will be analyzed by the computerized interpretation program and by the human electrocardiographer.

Baseline wander was characterized as either "saturation" or "sway". Saturation was defined as signals which exceeded the thresholds of +/- 4.8 millivolts for greater than 100 continuous milliseconds or ECGs with signals that exceeded both of these limits for any amount of time over the 10-second record. Sway was determined by tracking the minimum and maximum of a low-pass filtered version of the ECG signals. If the difference between these exceeded a threshold, the ECG lead was said to be affected by sway. An additional test considered the slope between the QRS onsets of consecutive QRS complexes.

Electrode noise was determined by examining QRS complexes for false QRS detections. This algorithm examined the individual lead energy content of the QRS, the RR intervals of QRS complexes, and a measure of the correlation of the QRS across all leads.

After each of these individual artifact tests, the overall ECG was assigned a lead quality as follows. If saturation was present in any lead, or if any lead was disconnected for any part of the 10-second record, the ECG was assigned a lead quality of "red". Detection of any other type of artifact resulted in a lead quality assignment of "yellow". If no artifact was found from any of the tests, the ECG was assigned a lead quality of "green".

2.2. ECG test sets

Two separate ECG databases were analyzed. The first was a database of 46385 ECGs from chest pain patients from two hospitals. The second database contained 74313 randomly selected ECGs from four different hospitals. Because the ECGs from the second database were randomly selected, they may be considered as a representative sampling of the ECGs that would be interpreted in clinical practice. Parts of the first database were used in the development of the artifact detection algorithms, while the second database was not used in any way during this development.

All ECGs were originally interpreted by earlier versions of the 12SL program at the time of acquisition, and were then confirmed, with or without editing, by a physician.

2.3. Analysis

All ECGs were assessed for artifact and assigned a lead quality level of green, yellow, or red. Except as noted below, if the ECG was marked as confirmed (a flag in the stored ECG record) and if a primary rhythm statement was found in the confirmed interpretation, the ECG was then analyzed by the most recent version of the 12SL algorithm and the primary rhythm of the 12SL interpretation was compared to the physician-confirmed primary rhythm interpretation. The discordance rate was defined as the percent of ECGs where the physician-confirmed primary rhythm interpretation differed from the primary rhythm of the 12SL interpretation.

ECGs with a confirmed rhythm of atrial, ventricular, or AV sequential pacing were analyzed for lead quality, but were excluded from the 12SL rhythm performance analysis. This was because proper pace spike detection relies on information not available to the 12SL program during a retrospective analysis from previously acquired ECG records. Because they lacked a reference rhythm interpretation, ECGs with a confirmed rhythm of "Undetermined Rhythm" were also excluded from the 12SL performance analysis. Likewise, ECGs for which 12SL gave a rhythm interpretation of "Undetermined Rhythm" were also excluded from the 12SL performance analysis. That is, an interpretation of "Undetermined Rhythm" was considered neither right nor wrong.

The primary rhythms were categorized as follows: sinus, AV block, ectopic atrial, atrial fibrillation, atrial flutter, supraventricular tachycardia (SVT), and junctional. Sinus included all sinus-originating rhythms with 1:1 AV conduction and constant PR interval. AV block included all ECGs with 2nd degree atrial-ventricular block (Mobitz I, Mobitz II, or 2 to 1), complete heart block (i.e., 3rd degree AV block), or AV dissociation. Ectopic atrial included ECGs with unusual P axis, but with 1:1 AV conduction. Junctional included all rhythms in which P waves are not present or not observed, including junctional rhythm, wide QRS rhythm, and idiopathic ventricular rhythm, but excluding the separate category of SVT.

The chi-square statistic was used for tests of difference of proportions. Fisher's exact test was used when the assumptions of the chi-square test were not met [9].

3. Results

All 120,698 ECGs were assessed for lead quality. However, 1168 ECGs from the first database and 3771 ECGs from the second database were excluded from the rhythm performance assessment because the ECG was not marked as confirmed, no primary rhythm statement could be found in the confirmed interpretation, or the confirmed interpretation was that of a paced ECG. An additional 410 ECGs from the first database and 1010 ECGs from the second database were excluded because primary rhythm statement was "Undetermined Rhythm" in either the confirmed or the 12SL retrospective analysis interpretation. This left a total of 44807 ECGs from the first database and 69532 ECGs from the second database remaining for the rhythm interpretation analysis.

Lead quality distributions and rhythm discordance rates for the combined test set (all 120,698 ECGs) are shown in Table 1.

Table 1. Lead quality and rhythm discordance for combined test set (N = 120,698).

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 ($p < 0.005$ for each successive change in proportion).

These measurements are reported by the respective databases in Tables 2 and 3. Figure 1 shows the discordance rates graphically as a function of the lead quality.

The first database showed lead quality distributions of 96.7%, 3.2%, and 0.1% for green, yellow, and red, with respective discordance rates of 3.7%, 7.5%, and 11.8% ($p < 0.001$ for green to yellow, not significant for yellow to red). In the second database, 94.57% of ECGs were green, 4.96% were yellow, and 0.46% were red, with discordance rates of 4.1%, 7.4%, and 12.1%, respectively ($p < 0.01$ for each successive change in proportion).

The difference in discordance rates between the two different databases were not significant for yellow or for red. Statistically significant differences were observed in the proportions of ECGs categorized as green, yellow, and red between the two databases ($p < 0.001$).

Table 2. Lead quality and rhythm discordance for first database (chest pain ECGs, N = 46385).

Lead quality	N	Percent of total	Discordance rate
Green	44848	96.67%	3.7%
Yellow	1482	3.19%	7.5%
Red	55	0.12%	11.8%

Table 3. Lead quality and rhythm discordance for second database (randomly selected ECGs, N = 74313).

Lead quality	N	Percent of total	Discordance rate
Green	70280	94.57%	4.1%
Yellow	3688	4.96%	7.4%
Red	345	0.46%	12.1%

Baseline wander was found to be the most frequent cause of non-green lead quality, followed by muscle tremor, and then electrode noise.

4. Discussion and conclusions

Although less than five percent of the ECGs were rated as yellow (marginal) or red (poor) lead quality, the effect on the performance of the computerized interpretation program was marked. In both databases, the 12SL rhythm analysis performance degraded as the lead quality degraded: the discordance rates nearly doubled as the quality went from green to yellow, and increased by as much again as the quality went from yellow to red.

Comparison of the lead quality prevalence in Tables 2 and 3 reveals an important observation: lead quality is a bigger problem at some institutions than at others. In the first database, which contained ECGs from two hospitals, only 3.3% of the ECGs were of non-green lead quality. Contrast this to the second database, which contained ECGs from four other institutions, where 5.4% of ECGs were of non-green lead quality. In particular, the proportion of red ECGs nearly quadrupled.

On the other hand, even though the prevalence of poor lead quality ECGs was greater in the second database, the 12SL performance was found to be consistent with regard to the lead quality level between the two databases.

A limitation of this study with regard to the 12SL rhythm analysis performance (discordance rate) is the assumption that the rhythm statement in the confirmed ECG was the true rhythm. It is known, however, that this is not always the case. The accuracy of the interpretation in the confirmed ECG depends largely on the overreading physician. Often the physician may not wish to spend time correcting the original 12SL interpretation for statements that may seem marginal or irrelevant to patient care. Different institutions may have varying quality

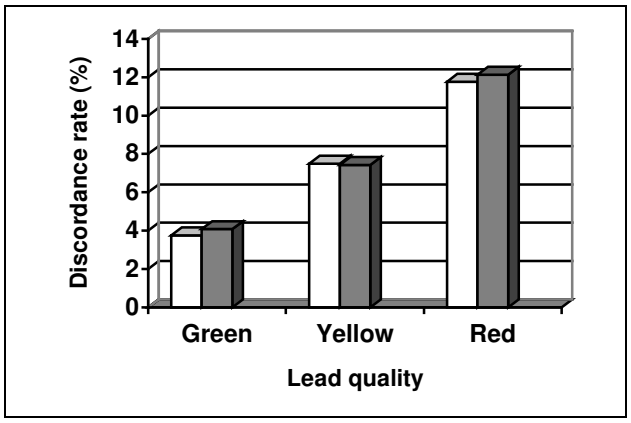

Figure 1. Primary rhythm discordance rates as a function of lead quality. White bars are the first database, dark bars are the second database. Rhythm analysis performance degrades as the lead quality degrades.

control procedures regarding the process of confirming the ECG, or there may be an institutional bias in overreading. Selecting ECGs randomly from four institutions was the best way to mitigate this limitation.

Another limitation, somewhat related to the first, is the extent to which the overreading physician may be influenced by the original 12SL interpretation.

The strength of this study is the large number of ECGs analyzed, as well as the fact that the second (larger) database provided a representative sampling of the ECGs that computerized interpretation algorithms are faced with on a daily basis. Because the ECGs used in this analysis are a representative sampling drawn from four institutions, we expect that the results of the second database will be reflective of the performance in actual clinical practice.

To our knowledge, this is the first report to quantify (1) the prevalence of poor technical quality ECGs in routine clinical practice and (2) the effect on poor technical quality on a computerized interpretation program.

The importance of lead quality is well known. Yet, ECGs are taken every day with poor lead quality. The reasons why this is so are many. One trend that may be contributing to this phenomenon is the use of lesser-trained ECG technicians. The primary purpose of the present work was to devise an easy-to-understand measure of lead quality for ECG technicians at the time of the ECG acquisition. As implemented in the electrocardiograph, a small on-screen indicator is used to denote the current lead quality with a text message appearing on the screen explaining any yellow or red indicator (e.g., muscle tremor, baseline wander, etc., and on which leads). In preliminary user evaluation of such cardiograph software, we have also been told that this scheme is a useful teaching aid for ECG technicians.

In conclusion, we have developed an algorithm that can quantify signal quality in a way that is directly related to the accuracy of computer interpretation of the data. Implementation of such an algorithm in ECG acquisition and analysis devices can provide a useful tool to improve the quality of acquired ECGs, reduce the effort to overread and edit the resulting records, and improve the quality of care based on interpretation these ECGs.

Acknowledgements

The authors wish to thank Don Brodnick for his original implementations of the AC, muscle artifact, and baseline wander artifact detection algorithms.

References

[1] Wagner GS. Marriott's practical electrocardiography. 10th ed. Philadelphia: Lippincott Williams & Wilkins; 2001.

[2] Chase C, Brady WJ. Artifactual electrocardiographic change mimicking clinical abnormality on the ECG. Am J Emerg Med 2000;18:312-316.

[3] Association for the Advancement of Medical Instrumentation. ANSI/AAMI EC11:1991/(R)2001 Diagnostic cardiographic devices. Arlington, VA: AAMI; 2001.

[4] Surawicz B, Knilans TK. Chou's electrocardiography in clinical practice: adult and pediatric. 5th ed. Philadelphia: W.B. Saunders Company; 2001.

[5] Knight BP, Pelosi F, Michaud GF, Strickberger SA, Morady F. Clinical consequences of electrocardiographic artifact mimicking ventricular tachycardia. N Engl J Med 1999;341:1270-1274.

[6] Krasnow AZ, Bloomfield DK. Artifacts in portable electrocardiographic monitoring. Am Heart J 1976;91:349-357.

[7] Kadish AH, Buxton AE, Kennedy HL, et al. ACC/AHA clinical competence statement on electrocardiography and ambulatory electrocardiography: a report of the ACC/AHA/ACP-ASIM task force on clinical competence. Circulation. 2001;104:3169 –3178.

[8] Brodnick DE, inventor; GE Medical Systems Information Technologies, assignee. Method of removing signal interference from sampled data and apparatus for effecting the same. US patent 6 351 664. 2002 Feb 26.

[9] Rosner BA. Fundamentals of biostatistics. 3rd ed. Boston: PWS-Kent; 1990.

Address for correspondence

Robert Farrell, Ph.D.
GE Healthcare Information Technologies
Diagnostic Cardiology
W126 N7449 Flint Road, Suite 300
Menomonee Falls, WI 53051
Robert.Farrell@med.ge.com