Best leads in the standard electrocardiogram for the emergency detection of acute coronary syndrome.

Green, Michael; Ohlsson, Mattias; Lundager Hansen, Jakob; Björk, Jonas; Edenbrandt, Lars; Ekelund, Ulf

Published in:
Journal of Electrocardiology

DOI:

2007

Link to publication

Citation for published version (APA):
Best leads in the standard electrocardiogram for the emergency detection of acute coronary syndrome

Michael Green, MSc,a,* Mattias Ohlsson, PhD,a Jakob Lundager Forberg, MD,b Jonas Björk, PhD,d Lars Edenbrandt, MD, PhD,c Ulf Ekelund, MD, PhDb

aDepartment of Theoretical Physics, Lund University, Lund, Sweden
bDepartment of Emergency Medicine, Lund University, Lund, Sweden
cDepartment of Clinical Physiology, Lund University, Lund, Sweden
dCompetence Centre for Clinical Research, Lund, Sweden

Received 5 July 2006; accepted 15 December 2006

Abstract

Background and Purpose: The purpose of this study was to determine which leads in the standard 12-lead electrocardiogram (ECG) are the best for detecting acute coronary syndrome (ACS) among chest pain patients in the emergency department.

Methods: Neural network classifiers were used to determine the predictive capability of individual leads and combinations of leads from 862 ECGs from chest pain patients in the emergency department at Lund University Hospital.

Results: The best individual lead was aVL, with an area under the receiver operating characteristic curve of 75.5%. The best 3-lead combination was III, aVL, and V2, with a receiver operating characteristic area of 82.0%, compared with the 12-lead ECG performance of 80.5%.

Conclusions: Our results indicate that leads III, aVL, and V2 are sufficient for computerized prediction of ACS. The present results are likely important in situations where the 12-lead ECG is impractical and for the creation of clinical decision support systems for ECG prediction of ACS.

Keywords: Acute coronary syndrome; Myocardial Infarction; Electrocardiography; Artificial neural networks

Introduction

In the emergency department (ED), the electrocardiogram (ECG) is crucial in the evaluation of a possible acute myocardial infarction (AMI) or unstable angina pectoris, that is, acute coronary syndrome (ACS). The standard 12-lead ECG may, in this situation, convey as much diagnostic information as all other clinical data taken together.1 For the ED diagnosis of ACS, it is conceivable that all of the standard 12 leads are not equally important. Myocardial ischemia and infarction are more frequent in some parts of the heart, and there are also “blind spots” in the standard ECG for certain regions of the heart, for example, that supplied by the left circumflex artery.2 If a few leads, or combinations of leads, would have as good or almost as good performance for ACS as the complete standard 12-lead ECG, this would be of interest both in situations where the 12-lead ECG is impractical, as in prehospital triage or in ECG monitoring of possible ACS, and for the creation of ECG decision support software. Selection of the best leads from a 12-lead ECG has previously been attempted for detection of coronary artery disease3 and for the assessment of QT prolongation.4

Artificial neural networks (ANNs) represents a machine learning tool that has proved useful for complex pattern recognition problems and is widely used for medical applications.5 The networks learn by associating different ECG patterns with the desired classification, not by being fed a set of predefined diagnostic criteria. Data from a large group of observations are presented to the networks, together with the desired classification, during a so-called training session. Neural networks have already
been applied to different aspects of automated interpretation of ECGs, for example, in the diagnosis of myocardial infarction. These studies have demonstrated a significantly improved performance over both conventional ECG criteria and experienced ECG readers. Neural networks have also been used for ACS prediction in patients with acute chest pain and have been compared to standard statistical methods such as multiple logistic regression. These studies indicate that networks are well suited as a tool for analyzing ECGs in suspected ACS patients.

The aim of this study was to elucidate, with the use of neural networks, which of the standard ECG leads or which combination of these leads have the largest predictive capability for the emergency diagnosis of ACS when being used together with a machine learning tool.

Methods

Study population

This retrospective study was based on the first ECGs recorded in the ED of Lund University Hospital on patients with a principal complaint of chest pain—from July 1997 to March 1999. Electrocardiograms were recorded 5 minutes to 1 hour after the patient arrived at the ED. Only ECGs for which the electronic ECG data could be retrieved were included, excluding ECGs with severe technical deficiencies and ECGs from pacemaker patients. Each ECG was classified as either “ACS” or “non-ACS,” depending on the hospital discharge diagnosis of the patient. A diagnosis of ACS was defined as a discharge diagnosis of AMI or unstable angina pectoris, and the criteria for these diagnoses were the ones used during the ECG recording period. Acute myocardial infarction was defined by the World Health Organization criteria where the biochemical criterion was at least one measurement of creatine kinase–MB below 10 µg/L or troponin T below 0.1 µg/L. The criteria for unstable angina were ischemic symptoms (chest pain >15 minutes, syncope, acute heart failure or pulmonary edema) together with at least one of the following: (a) ECG changes—transient or persisting ST-segment depression (≥1 mm) and/or T-wave inversion (≥1 mm) without developing Q waves or loss of R wave height, or (b) biochemical markers—creatine kinase–MB 5 to 10 µg/L or troponin T 0.05 to 0.1 µg/L.

All discharge diagnoses were made by the senior ward physician or the ED physician (in cases discharged from the ED), reviewed by a senior research nurse, and when ambiguous, further reviewed by a senior cardiologist. In the review of diagnoses for cases discharged from the ED, available data from the patient records indicated that the rate of missed diagnosis of ACS, compared to the above-described criteria, was low (≤2%).

The final data set consisted of 862 patients, 345 with diagnosis of ACS and 517 with diagnosis of no ACS. Among the non-ACS cases, 123 patients were diagnosed as stable angina pectoris, 114 as suspected angina pectoris, and the remaining 280 patients belonged to the category “other diagnoses.” The mean age within the ACS and non-ACS group was 69(13) and 62(18) years, respectively, and the numbers in parenthesis are SDs. In addition, the ACS group consisted of 227 men and 118 women, and the corresponding numbers for the non-ACS group were 291 and 226.

This study was approved by the Lund University Research Ethics Committee.

Electrocardiography

The 12-lead ECGs were recorded by the use of computerized electrocardiographs (Siemens-Elema AB, Solna, Sweden), and the following 12 measurements taken from each of the 12 leads were selected for further analysis: QRS duration, QRS area, Q duration, Q amplitude, R duration, R amplitude, ST-J amplitude, ST slope (the slope at the beginning of the ST segment), ST amplitude 2/8, ST amplitude 3/8, positive T amplitude, and negative T amplitude. All durations and amplitudes are measured in milliseconds and microvolts, respectively. The ST amplitude 2/8 and ST amplitude 3/8 were obtained by dividing the interval between ST-J point and the end of the T wave into 8 parts of equal duration. The amplitudes at the end of the second and the third intervals were denoted ST amplitude 2/8 and ST amplitude 3/8. In total, 144 measurements from each 12-lead ECG were collected. To reduce the number of input measurements for the neural networks, a principal component analysis (PCA) on the 12 measurements within each lead was used. Using only the first 6 principal components in each lead resulted in a total of 72 measurements when considering all 12 leads. The number of selected principal components was chosen as to include at least 90% of the variance in each lead. The variance captured in each lead varied within a range of 91.1% to 94.9%. The PCA analysis was based on the correlation matrix.

Artificial neural networks

In this work, we built ACS prediction classifiers using neural network ensembles with the bagging technique. A general presentation of ANNs can be found in the work of Cross et al. An ensemble size of 50 was chosen, which has been found to be sufficient in numerical studies. The ensemble prediction was computed as the average over the output of each of the individual networks. All 6 principal components from the PCA step was fed to the ANN as continuous variables.

The model selection consisted of selecting the best architecture and regularization parameter for each neural network ensemble with respect to the area under the receiver operating characteristic (ROC) curve. The ROC area is commonly used as a performance measure and can be interpreted as the probability that a randomly chosen patient with ACS has a higher risk output than a randomly chosen patient without ACS. We used K-fold cross-validation to estimate the best ensemble parameters. To accomplish this, the training data were split into K random equally sized disjoint parts. One part was selected for the validation of the neural network ensemble, which was constructed on the other K-1 parts. This procedure was repeated for all K parts.
Table 1
The test ROC areas for the individual leads

<table>
<thead>
<tr>
<th>Selected ECG lead</th>
<th>Test ROC area (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>74.1 (67.9, 81.8)</td>
</tr>
<tr>
<td>II</td>
<td>68.6 (61.6, 76.2)</td>
</tr>
<tr>
<td>III</td>
<td>75.0 (68.2, 80.6)</td>
</tr>
<tr>
<td>aVR</td>
<td>67.9 (62.2, 75.3)</td>
</tr>
<tr>
<td>aVL</td>
<td>75.5 (65.8, 82.6)</td>
</tr>
<tr>
<td>aVF</td>
<td>72.0 (63.7, 78.1)</td>
</tr>
<tr>
<td>V1</td>
<td>67.8 (60.6, 75.7)</td>
</tr>
<tr>
<td>V2</td>
<td>74.3 (67.7, 82.5)</td>
</tr>
<tr>
<td>V3</td>
<td>73.7 (65.4, 81.3)</td>
</tr>
<tr>
<td>V4</td>
<td>72.3 (66.1, 79.6)</td>
</tr>
<tr>
<td>V5</td>
<td>71.5 (65.6, 79.3)</td>
</tr>
<tr>
<td>V6</td>
<td>73.7 (65.1, 81.6)</td>
</tr>
</tbody>
</table>

The ROC area is presented as median (2.5, 97.5 percentiles) over the 100 test sets.

The K-fold cross-validation was repeated N times, and the total validation result was taken as the mean of the N × K validation results. We used N = 10 and K = 5 for the model selection.

To estimate the generalization performance of the neural network ensemble, an outer cross-validation loop was used. The data were randomly split into 5 disjoint parts. Each part was selected as a test set with the rest of the parts as the corresponding training set. The outer cross-validation loop was repeated 20 times, resulting in 100 training and test sets. The total test result was evaluated as the median over the 100 test results.

Statistics

We used the area under the ROC curve to assess the performance of the neural networks. When comparing 2 different neural network classifiers, on a given test set, we used their corresponding outputs to evaluate whether they produced significantly different ROC areas or not. Statistical significance was evaluated using a permutation test\(^{20}\) where we considered a \(P < 0.05\) as statistically significant.

Results

All results are presented as medians over the 100 ROC areas produced by the outer cross-validation loop. The results for the neural network classifiers fed with single leads as input are presented in Table 1. The 3 best limb leads I, III, and aVL had similar performance with ROC areas of 74.1%, 75.0%, and 75.5%, respectively. Leads II, aVR, and aVF did not match that performance. For the precordial leads, the best performance was obtained using lead V2 with an ROC area of 75.3%. However, leads V3 and V6 were almost as good with ROC areas of 73.7%. Statistical evaluations showed that a significant difference between the best-performing (aVL) and the worst-performing (V1) leads was found in 36 of the 100 test sets.

The performance of the neural networks classifiers fed with inputs from different combinations of leads are presented in Table 2. The 2 (III and aVL) best individual leads were combined, and this combination obtained an ROC area of 78.9%. Any 2 lead combinations of the six limb leads resulted in similar ROC areas with a median area of 77.9% (range, 74.5%-78.9%). Adding 1 precordial lead to the best 2 lead combinations almost always increased the performance (see Table 2). The best three lead combinations was III, aVL, and V2 with an area under the ROC curve of 82%.

Table 2 also shows the results for the combination of all limb leads (denoted 6-lead ECG), the 2 best combinations of the 6-lead ECG and 1 precordial lead, and the full 12-lead ECG. The performance of the neural network when using the 12-lead ECG was 80.5%. A statistical comparison of the best 3-lead combination (III-aVL-V2) and the full 12-lead ECG resulted in only 10 of the 100 test splits being significantly different, indicating that performance of these 2 combinations of leads are comparable.

The ROC curves for the best single lead, the best 3-lead combination, and for the 12-lead ECG are shown in Fig. 1. A comparison with traditional ECG criteria for AMI detection resulted in a specificity and sensitivity of 95.6% and 24.3%, respectively. The sensitivity of the AMI subgroup was 34.1%, and the corresponding result for the unstable angina subgroup was 5.2%.

Discussion

In the present study we attempted to establish the best lead, or combination of leads, for the ED diagnosis of ACS. The results showed that the best individual lead was aVL (ROC area of 75.5%) and that the 6 limb leads together with either V2 (80.2%) or V3 (80.7%) had principally the same performance for ACS as the complete 12-lead ECG (80.5%). Somewhat surprisingly, using only leads III, aVL, and V2 gave similar discriminatory power for ACS (82.0%). It thus seems that these 3 leads together contain all the ACS-predicting information present in the standard 12-lead ECG, at least in the present patient material. This can partially be explained by the fact that any 2 limb leads can be used to derive the other 4 limb leads when using the raw ECG lead recording. Thus, given that our representation of the ECG is good enough, the ANN will be able to extract information about all 6 limb leads even if only 2 of them are fed to the network as inputs.

The present results are compatible with previous studies on optimal leads for detection of ST segment deviations in acute myocardial ischemia. During coronary occlusion
induced by balloon angioplasty, the largest ST changes have been observed in leads V₂ to V₄ (occlusions of the left anterior descending or circumflex arteries) and in leads III and aVF (right coronary artery), and these leads have therefore been suggested to be optimal for ischemia detection during balloon angioplasty. For identification of ST changes in established AMI, leads III and V₂ have been suggested to be optimal. However, these results are not immediately applicable to ED patients with suspected ACS. First, many ED patients with ACS do not have ST-segment changes at all but, rather, T-wave inversions, new Q waves, or no ECG changes at all, and the ECG changes may, in turn, be due to subtotal and varying occlusion of branches of the large coronary arteries. Because we considered not only the ST segment but several other ECG variables (QRS duration, QRS area, Q duration, R duration, and R and T amplitudes), it is not surprising that our results differ from those in studies focusing only on the ST segment. For instance, aVL was the single best lead for ACS prediction in our study, whereas during balloon angioplasty, ST in aVL was too low to be of any use for ischemia detection. Second, in the present study, only 1 ECG from each patient was considered. The neural networks thus only had access to absolute measures in the ECG and not to any relative changes induced by ischemia in the ACS patients. It may be that preexisting ECG changes unrelated to current ischemia contributed to ACS detection by the neural networks in our patients.

In the present results, good ACS discriminating power with only 3 leads was observed. Electrocardiographic registration with reduced lead sets is practical for many reasons. Few leads interfere less with the everyday care of the patient, with diagnostic tests such as echocardiography and with emergency procedures such as defibrillation. To detect acute ischemia by ST deviation, however, current consensus is that all 12 leads of the standard ECG are necessary. Indeed, many ischemic events were missed when only the usual telemetry leads (V₁ and II) were used, or even the 3 single best leads for detection of ST deviation. Ischemia detection with reduced lead sets have, in fact, so far, only been successful when the omitted leads have been calculated or when a derived 12-lead ECG has been used. Thus, in reduced lead sets, it seems that ischemia detection will not be satisfactory if only the ST segment is monitored. To our knowledge, detection of ischemia using multiple ECG variables in reduced lead sets has not previously been tested. Our finding that leads III, aVL, and V₂ together predicted ACS as well as the standard 12-lead ECG may thus be explained by the fact that we included several ECG measurements in addition to the ST segment. We have not, however, investigated the relative importance among the ECG measurements within each included lead.

In this study, we used neural networks as the method for ACS prediction with a varying number of input leads. This choice of classification method was guided by previous work where neural networks have proved to be useful for ACS and AMI prediction. Standard linear statistical methods, such as multiple logistic regression, would not have been sufficient because there are nonlinear relationships among the lead measurements, used by the networks, that are important for predicting ACS. The PCA preprocessing of the ECGs has been used previously and can be motivated by the fact that measurements for each lead showed large correlations. Furthermore, it is always advantageous to keep the number of inputs to the network models as low as possible since the problem of overtraining usually increases with an increasing number of inputs. Using PCA for this reduction is a...
commonly used method. Care was taken to obtain as reliable estimates as possible of the generalization performance for each lead selection. Although the study population was relatively small, which may have influenced the absolute values for the ROC areas, we believe that the obtained selection of important leads is valid.

**Clinical implications**

The present results have their main implications for the creation of future clinical decision support systems (CDSS) for ECG interpretation. For a CDSS to produce as robust ACS predictions as possible, it is essential that it is allowed to work only with the ECG elements crucial for ACS prediction and that other information is left out. With more robust ACS predictions, the CDSS will, of course, be more valuable to the patients and the physicians using them. The identified leads III, aVL, and V2, together with clinical patient data such as chest pain history and blood pressure, could be used to develop a neural network-based CDSS that would potentially be useful in situations where the standard 12-lead ECG is impractical, as in, for example, prehospital triage or in telemedicine settings. For true clinical usefulness, such a CDSS should also include an ANN able to detect ST-elevation myocardial infarction in need of urgent reperfusion therapy. Before clinical implementation, the CDSS would of course need to be validated prospectively, preferably at multiple centers.

**Limitations of the study**

The results from this study are probably not applicable to the manual interpretation of ECGs by physicians. It is not at all evident, and perhaps even unlikely, that leads III, aVL, and V2 together would be as useful as the 12-lead ECG to the physician trying to establish whether the patient has ACS or not. Some of the variables used in the present study are not part of the standard ECG interpretation routine and are not easy to appreciate by eye. Furthermore, if ANNs such as those in the present study are to be used in CDSS for physicians, a problem is that the ANN is unable to explain to the user the reasons for the suggested decisions. Current research is trying to overcome this problem.

The results were obtained using ECG data collected from a limited number of patients during a limited period and at 1 center only. Other populations might, of course, produce different results. Likewise, we cannot exclude the possibility that another set of ECG variables than the ones chosen would produce other results. However, we believe it is unlikely that results in other populations or with other variables would differ substantially because only the relative performances of the different leads and combinations of leads were analyzed in this study.

The ECGs in the present study were collected in the late 1990s, and old definitions of AMI and unstable angina were used. More recent definitions of AMI have lower cutoff values for biochemical markers, and for the diagnosis of unstable angina no marker elevation is currently needed. A few patients classified as non-ACS in the present study may thus be classified as having an ACS with current diagnostic criteria.

**Conclusions**

The aim of this study was not to find the best neural network classifiers for prediction of ACS but, rather, to compare the information content of the different leads and of the different combinations of leads. We found that the lead aVL was the single best lead for ACS detection and that the leads III, aVL, and V2 together yielded similar performance as the full 12-lead ECG for predicting ACS. It thus seems that these 3 leads together contain all the ACS predicting information present in the standard 12-lead ECG, at least in our patient population. These findings may be useful for the creation of ECG decision support software to be used in situations where the 12-lead ECG is impractical.

**References**


---

Artifact mimicking ventricular fibrillation. The patient was using an electrical toothbrush. Note that there is no change in the QRS rate, and the QRS complexes can be marched out throughout the strip by the S-wave deflection (arrows).

doi:10.1016/j.jelectrocard.2007.03.005