Comparative Studies on the Diagnostic Accuracy between the Two Diagnostic Logic in Electrocardiographic Interpretation

MASASHI YOKOI*, SHOJI YASUI*, MITSUHARU OKAJIMA**, AND YASUSHI MIZUNO*

For the purpose of comparison in the diagnostic accuracy between the decision tree method and the joint probability method, they were applied to the diagnosis of the orthogonal electrocardiograms obtained from 101 patients with myocardial infarction (Infarction) and 601 patients without infarction (Non-infarction).

As significant parameters, 14 measurements, including the duration and amplitude of Q wave, the deviation of S-T segment and the maximal T amplitude, were selected.

A small-sized digital computer (NEAC-2203) was used for this study.

When the correct diagnosis was made over 80 per cent in Infarction with the decision tree method, it was unavoidable for the rate of “false positive” diagnosis in Non-infarction to exceed 10 per cent.

The accuracy of pattern classification in Infarction was 95.0 per cent with the joint probability method, and frequency of “false positive” diagnosis in Non-infarction was as low as 4.8 per cent.

Thus, the better results in discrimination were obtained with joint probability method than with the logical decision tree method.

These results indicate that the degree of diagnostic accuracy may be largely dependent on the classification method used and the joint probability method would be the better categorizer than the logical decision tree method in the automatic electrocardiographic interpretation.

Recently a remarkable progress has been made in the field of medical data processing by means of computer technique. Especially, many trials have been taken for applying digital computers to the pattern recognition of electrocardiograms (ECG’s). An objective and quantitative interpretation of ECG can be obtained in a brief time by computer, provided that an appropriate program is set up. The diagnostic procedure would be correctly executed according to the programs by the computer since it possesses inherent characteristics such as a vast capacity of memory and high accuracy and high speed in logical and analytical operations.

However, one of the crucial problems debated currently in this field is to choose the most adequate logical process to be used in the ECG interpretation. In this paper, two commonly used diagnostic logics were selected and compared in its capability in diagnosis of ECG, i.e., the logical decision tree and the joint probability. The discriminative power with the myocardial infarction and other groups was chosen as the index to evaluate these two logical procedures.

**MATERIALS**

One hundred and one patients with myocardial infarction (Infarction) and 601 subjects without infarction (Non-infarction) were selected, whose diagnoses were made on the basis of clinical and labo-

(Received for Publication November 7, 1968)

* First Department of Internal Medicine, Nagoya University School of Medicine, Nagoya
** Research Institute of Environmental Medicine, Nagoya University, Nagoya

Japanese Circulation Journal Vol. 33, January 1969 51
ratory findings.
Non-infarction group consisted of 201 normals (Normal), 127 cases with left ventricular hypertrophy (LVH), 174 cases with right ventricular hypertrophy (LVH), 174 cases with right ventricular hypertrophy (RVH) and 99 cases with combined ventricular hypertrophy (CVH).

The diagnosis of "Infarction" was confirmed by the clinical findings such as typical precordial pain lasting over one hour, elevation of serum enzyme activities such as glutamic oxaloacetic transaminase and lactic dehydrogenase, and serial ECG findings supportive of the history of an acute myocardial infarction. Patients of "Non-infarction" were also diagnosed by clinical findings, including history, blood pressure, 12-lead ECG's, X-ray films, phonocardiograms and, in some cases, the right heart catheterization. The majority of patients with LVH suffered from essential hypertension, and materials of RVH and CVH consisted mainly of patients with congenital and rheumatic heart diseases.

Recording and Measurement

The X, Y and Z scalar electrocardiograms of Frank lead system were recorded simultaneously on the heat writing paper at a speed of 100mm per second.

Thirty parameters were measured manually in each ECG, namely 1) the amplitude of Q, R, S and R' in lead X, Y and Z, 2) the deviation of S-T segment at 0.04 second after S-T junction in lead X, Y and Z, 3) the maximal amplitude of T wave in lead X, Y and Z, 4) the width of Q, R and S waves and the interval from the onset of QRS complex to the top of R wave in lead X and Z, and the width of R, S and R' waves in lead Y, 5) the time interval between the beginning the earliest deflection of the QRS wave noted in any of the three simultaneous records of X, Y and Z leads and the end of QRS complexes. In this study, left (X) to the patients, upward (Y) and anterior (Z) were designated as positive.

A preliminary analysis by means of T-test was performed with these 30 parameters to test the ability of differentiating Infarction group from Non-infarction group. Fourteen out of 30 parameters were proved to be significant (p<0.01) for differentiating two groups. They consisted from the amplitudes (A) and durations (D) of Q wave in lead X and Z (abbreviated as QAX, QAZ, QDX, QDZ respectively), those of R wave in lead Y and Z (RAY, RAZ, RDY, RDZ), the S-T deviation at 0.04 second after S-T junction and the maximal T amplitude in lead X, Y and Z (S-TX, S-TY, S-TZ, TX, TY, TZ). The remainders of the 30 measurements were not significant and discarded in the later analysis.

Method of Computation

A. The logical decision tree method

Here, the diagnostic procedure was carried out through the six steps of pre-set threshold of diagnostic criteria (Fig. 1). Step 1 and 2 were criteria

![Diagram](image)

**Fig. 1.** Flow diagram for differentiation between infarction group and others by the logical decision tree method.

Step 1, 2 are thresholds of discrimination for a lateral infarction.
Step 3, 4 for an inferior infarction.
Step 5, 6 for an anterior infarction.

The process of discrimination was carried out through the Step 1 to the Step 6.
LVH: left ventricular hypertrophy,
RVH: right ventricular hypertrophy,
CVH: combined ventricular hypertrophy.
Infar: myocardial infarction.

...
Table I: Threshold Levels for Classification between Infarction Group and Others. These thresholds consisted of five levels of severity ranging from the most strict grade (Grade 1) to the mildest (Grade 5).

<table>
<thead>
<tr>
<th>Grade</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>QDX</td>
<td>40</td>
<td>30</td>
<td>28</td>
<td>26</td>
<td>24</td>
</tr>
<tr>
<td>QAz</td>
<td>0.3</td>
<td>0.20</td>
<td>0.18</td>
<td>0.14</td>
<td>0.10</td>
</tr>
<tr>
<td>QDX</td>
<td>30</td>
<td>28</td>
<td>26</td>
<td>24</td>
<td>20</td>
</tr>
<tr>
<td>QAz</td>
<td>0.30</td>
<td>0.20</td>
<td>0.18</td>
<td>0.14</td>
<td>0.10</td>
</tr>
<tr>
<td>STx</td>
<td>0.15</td>
<td>0.05</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Tx</td>
<td>-0.20</td>
<td>-0.15</td>
<td>-0.05</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>RDy</td>
<td>40</td>
<td>30</td>
<td>28</td>
<td>26</td>
<td>24</td>
</tr>
<tr>
<td>RAy</td>
<td>0.30</td>
<td>0.20</td>
<td>0.18</td>
<td>0.14</td>
<td>0.10</td>
</tr>
<tr>
<td>RDy</td>
<td>30</td>
<td>28</td>
<td>26</td>
<td>24</td>
<td>20</td>
</tr>
<tr>
<td>RAy</td>
<td>0.30</td>
<td>0.20</td>
<td>0.18</td>
<td>0.14</td>
<td>0.10</td>
</tr>
<tr>
<td>STy</td>
<td>-0.15</td>
<td>-0.05</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Ty</td>
<td>0.20</td>
<td>0.15</td>
<td>0.10</td>
<td>0.05</td>
<td>0.00</td>
</tr>
<tr>
<td>QDX</td>
<td>40</td>
<td>30</td>
<td>28</td>
<td>26</td>
<td>24</td>
</tr>
<tr>
<td>QAz</td>
<td>0.04</td>
<td>0.08</td>
<td>0.12</td>
<td>0.16</td>
<td>0.16</td>
</tr>
<tr>
<td>STz</td>
<td>0.15</td>
<td>0.13</td>
<td>0.11</td>
<td>0.09</td>
<td>0.07</td>
</tr>
<tr>
<td>Tz</td>
<td>-0.20</td>
<td>-0.15</td>
<td>-0.10</td>
<td>-0.05</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Parameter values are above or below the corresponding threshold. Five different runs were performed with each ECG by setting the threshold value for the discrimination at five different levels for each parameter in all of the 6 steps. For example, in the Step 1 the thresholds of QDX were set at 40, 30, 28, 26 and 24 milliseconds (Grade 1 – Grade 5) (Table I). Diagnostic accuracy of the five runs with the five grades of pre-set level of threshold were compared. In order to obtain better discrimination, these threshold values were changed several times. Relatively good results were obtained with the threshold levels of Table I.

B. The joint probability method

Since Infarction group contained 101 ECG records, 78 cases were used in establishing the probability density function and the others were used to test the stability and validity of the classification method. The former group was called "internal sample" and the latter infarction group were named "external sample".

As a preliminary analysis of the moment coefficients of skewness and kurtosis, the pattern of the distribution of parameter values in Infarction group was examined. The distributions of all parameters concerned were proved to be described by multivariate normal curves. Then, the score of the probability of Infarction for a parameter value was computed from the normal distribution function.

When X̄i is the mean of parameter i and σi is its variance in Infarction group, then the score of the probability Pi of Infarction was calculated by the following formula:

\[ P_i = \int_{-\infty}^{\frac{X_i - \bar{X}_i}{\sigma_i}} f(x)dx \]

where the function f was the normal distribution function and X̄i was the value of parameter i measured in each ECG (Fig. 2).

The probability for a value of Q wave in lead X or Z was computed as left tailed probability. This came from the fact that, in the specificity of Infarction, the greater the Q wave, the higher the possibility of being Infarction. The scores of the probability for other parameters were calculated by the similar principle.

The score of the joint probability was calculated assuming that these phenomena are mutually independent.

The discrimination was made according to the difference between the score of the joint probability...
for Infarction and the threshold set for the probability (Fig. 3).

The score of the joint probability of lateral infarction ($A_1$), that of inferior infarction ($A_2$) and those of anterior infarction ($A_3$ and $A_4$) were computed. Based on the confidence limits of the joint probability in the infarction group were determined at 90 per cent as the strict level, while 99 per cent as the mild threshold for the classification of the independent group. Then, each score of the joint probability was compared with these two different levels of threshold, the strict grade and mild one.

When one of these scores of the joint probability was greater than the strict threshold, the ECG was classified into Infarction. If all of the probabilities for three locations of Infarction were smaller than the mild threshold, the diagnosis given was Non-infarction. When one or two of the scores of probability were between these thresholds, another score of the joint probability $P_{QD} \times P_{QA}$ was calculated and the differentiation was made depending on whether this score was above or below the preset threshold.

**Digital Computer**

The NEAC 2203 digital computer at the Computation Center of Nagoya University was used for analytical operation in this study.

This small scale computer had only 400-word core memory and 12,000-word magnetic drum memory, and the computation time of an addition or a subtraction was about 3 milliseconds.

**Results**

**A. The Logical Decision Tree Method**

For each of the five Grades of threshold, respectively the percentages of correct diagnosis in Infarction group were 37.7, 63.4, 64.3, 70.2 and the percentages of the false positive cases in Non-infarction group, in which a misdiagnosis of Infarction was given, were 0.5, 3.8, 4.5, 8.7 and 17.0 (Fig. 4). Thus, a relatively good discrimination was obtained by the threshold level of Grade 4. Correct classification in Infarction group varied with the location and extent of the infarcted area. For the threshold level of Grade 4, the correct diagnosis was made in 26 cases out of 40 anteroseptal infarction (65.0 per cent), in 13 cases out of 18 anterolateral infarction (72.2 per cent), in 19 cases out of 23 inferolateral infarction (82.6 per cent), and in 12 cases out of 20 inferior infarction (60.0 per cent). Inferolateral infarction was recognized at highest rate. The frequency of false positive cases in Non-infarction group was 4.4 per cent in Normal, 13.1 per cent in LVH, 11.0 per cent in RVH and 13.2 per cent in CVH for this threshold. Relatively high incidence of false positive was obtained in LVH and CVH. In the milder threshold levels such as Grade 4 and

---

*Japanese Circulation Journal  Vol. 33, January 1969*
5, less false negatives in Infarction group were observed, while more false positives in Non-infarction group occurred. Correct diagnosis in Infarction group was made mostly because of the presence of abnormal Q wave. However, at fairly large percentage, false positives in Non-infarction group occurred because of QX or RY.

B. The Joint Probability Method

Using this method, the rates of correct diagnosis in Infarction group were achieved as high as 94.8 per cent in the internal sample and 95.2 per cent in the external one. The false positive diagnosis in Non-infarction group reached 4.8 per cent. Anteroserial infarction was correctly recognized in 97.5 per cent, anterolateral in 94.4 per cent, inferolateral in 100.0 per cent and inferior in 85.0 per cent. A higher rate of misdiagnosis was observed in inferior infarction than in the other locations of Infarction. False positive diagnosis was made in 4.0 per cent of Normal, 6.3 per cent of LVH, 4.6 per cent of RVH and 5.0 per cent of CVH. That is, the most frequent misdiagnosis of Non-infarction was made in LVH. False positive diagnosis in cases of Normal, LVH and CVH were mostly caused by the probability for lateral infarction $(A_1)$. False diagnosis as Infarction in cases of RVH was mainly due to the probability for the inferior infarction $(A_2)$. As compared with the decision tree method, the better results were obtained by the joint probability method.

DISCUSSION

In the computer diagnosis of electrocardiogram, the diagnostic accuracy is influenced by the following three stages; 1) data acquisition and conversion into a suitable form for the computer, 2) parameter measurement by computer, and 3) automated interpretation of ECG pattern. The third factor will be discussed in this paper.

A. Selection of Parameters for ECG Interpretation

In the previous reports of computer application to ECG diagnosis, there were used various parameters such as amplitudes and durations of ECG waves, instantaneous vectors, spatial components, polar vectors and Fourier's coefficients. Diagnostic accuracy was found to depend upon the number of parameters and the information contained in each parameter.

By means of statistical analysis, the 14 parameters used in this study were selected as significant ones for differentiating Infarction group from others ($P<0.01$).

B. Classification Method

There were few reports in which the diagnostic accuracy of the various classification methods or categorizers was discussed.

i) The Decision Tree Method

In the several studies, the decision tree method was applied to the automatic diagnostic system of ECG's.

Using this method, the classification of ECG's could be easily made by "yes or no" logic depending on the threshold which corresponded to the clinical diagnostic criteria. This categorizer can be used even for a small scale automatic diagnostic system, because of the simple analytical process.

However, in the logical decision tree method, a decision is made simply on the basis of whether the parameter values are above or below the threshold, although the distribution of the parameter values in Infarction group and Non-infarction group, for example, is normally overlapped widely each other. When one of the parameters has a value below the threshold, the diagnosis should be Non-infarction, even though other parameters have value sufficiently above their thresholds of Infarction.

Consequently, when the continuous values like amplitudes of ECG waves are used as parameter, it might be difficult to determine the point for the threshold which will permit the best discrimination. Moreover considerable amount of the information may not be utilized in this method.

In this study, therefore five grades of threshold sets were used for the discrimination in order to avoid these shortcomings.

ii) The Joint Probability Method

As compared with the logical decision tree method, the better discrimination was obtained by the joint probability method, in which the more impartial weighting and the better use
of the efficient information could be expected in the evaluation of different diagnostic parameters.

Rikli et al. used the probability density function method the ECG interpretation. In this study, however, the integral value of the probability density function, i.e. the tailed probability, was computed in the following consideration.

When, in a certain case, the amplitude of Q wave is larger than the mean value of the Infarction group, the probability density score becomes smaller than that of the mean value of the Infarction group, although the possibility of Infarction becomes more definite by the large amplitude of Q wave. In the preliminary analysis on the same material of this study by the probability density function method, the correct classification in the Infarction group was 89.1 per cent, while the false positive classification in the Non-infarction group was 9.4 per cent, whereas correct classification was made in 95.0 per cent for the Infarction group and in 4.8 per cent false positive diagnosis for the Non-infarction group respectively by the joint probability method (Table II).

The following problems must be kept in mind to make use of this method as the categorizer.

First, the distribution pattern of the ECG

<table>
<thead>
<tr>
<th>Table II</th>
<th>Comparison of Diagnostic Accuracy between the Probability Density Function Method and the Joint Probability Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physician's Diagnosis</td>
<td>No. of Subjects</td>
</tr>
<tr>
<td>Infarction</td>
<td>101</td>
</tr>
<tr>
<td>Non-infarction</td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>201</td>
</tr>
<tr>
<td>LVH</td>
<td>127</td>
</tr>
<tr>
<td>RVH</td>
<td>174</td>
</tr>
<tr>
<td>CVH</td>
<td>99</td>
</tr>
</tbody>
</table>

items has to be examined in order to avoid the parameters which are not necessarily distributed normally. The probability of being Non-infarction was not computed, because of the distorted distributions of the parameter values in Non-infarction group, which were used to differentiate Infarction from other conditions. Second problem has to be emphasized as to the mutual dependency among various parameters. The amplitude value of Q wave, for example, is highly correlated to the duration of Q wave (Table IIIA, Table IIIB).

In this study, however, the score of the joint probability of Infarction was calculated on the

Table IIIA Correlation Coefficients among Various Electrocardiographic Items of anterior Infarction Subjects

<table>
<thead>
<tr>
<th>STX</th>
<th>TX</th>
<th>RAY</th>
<th>STV</th>
<th>TV</th>
<th>QAZ</th>
<th>RAZ</th>
<th>STZ</th>
<th>TZ</th>
<th>QDX</th>
<th>RDV</th>
<th>QDZ</th>
<th>RDZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>QAX</td>
<td>-.11</td>
<td>.22</td>
<td>.01</td>
<td>-.01</td>
<td>.18</td>
<td>-.01</td>
<td>-.01</td>
<td>.24</td>
<td>.45</td>
<td>.80</td>
<td>-.01</td>
<td>.00</td>
</tr>
<tr>
<td>STX</td>
<td>.47</td>
<td>.01</td>
<td>-.16</td>
<td>-.14</td>
<td>-.13</td>
<td>.01</td>
<td>-.26</td>
<td>.00</td>
<td>-.01</td>
<td>.15</td>
<td>-.14</td>
<td>.16</td>
</tr>
<tr>
<td>TX</td>
<td>-.17</td>
<td>-.32</td>
<td>-.68</td>
<td>-.38</td>
<td>-.14</td>
<td>-.33</td>
<td>.00</td>
<td>-.17</td>
<td>-.01</td>
<td>-.31</td>
<td>.19</td>
<td></td>
</tr>
<tr>
<td>RAY</td>
<td>.00</td>
<td>.27</td>
<td>.01</td>
<td>-.22</td>
<td>.00</td>
<td>.01</td>
<td>.24</td>
<td>.79</td>
<td>.22</td>
<td>.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STV</td>
<td>.50</td>
<td>.13</td>
<td>.16</td>
<td>-.01</td>
<td>.01</td>
<td>-.18</td>
<td>.01</td>
<td>.01</td>
<td>.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TY</td>
<td>.27</td>
<td>-.17</td>
<td>.26</td>
<td>.13</td>
<td>.18</td>
<td>.28</td>
<td>.24</td>
<td>-.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QAZ</td>
<td>-.20</td>
<td>.25</td>
<td>.11</td>
<td>-.18</td>
<td>.01</td>
<td>.90</td>
<td>-.45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAZ</td>
<td>-.62</td>
<td>-.14</td>
<td>-.14</td>
<td>-.19</td>
<td>-.15</td>
<td>.41</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STZ</td>
<td>.29</td>
<td>.17</td>
<td>-.01</td>
<td>.01</td>
<td>-.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TZ</td>
<td>.36</td>
<td>-.01</td>
<td>.01</td>
<td>-.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QDX</td>
<td>-.01</td>
<td>-.21</td>
<td>.15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RDV</td>
<td>.23</td>
<td>.21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QDZ</td>
<td>-.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

STX, STV and STZ: the deviation of S-T segment at 0.04 second after S-T junction in lead X, Y and Z. TX, TY and TZ: the maximal T amplitude in lead X, Y and Z. RAY and RAZ: the amplitude of R wave in lead Y and Z. QAX and QAZ: the amplitude of Q wave in lead X and Z. QDX and QDZ: the duration of Q wave in lead X and Z. RDV and RDZ: the duration of R wave in lead Y and Z.
assumption of mutual independency among the parameters, because of the complexity in computation. The following procedures may decrease this shortcomings 1) to avoid either one of the two highly correlated parameters, 2) to take a simple approximation regarding to the dependency\[^9\].

On the third problem, discrimination among the internal samples which have been used for the determination of the thresholds, could be made successfully, but good discrimination may not be always expected among the new independent samples.

In this study, however, both internal and external samples were well recognized in about 95 per cent of the cases.

Considering the clinical importance of myocardial infarction, the threshold levels for discrimination between Infarction group and Non-infarction group should be determined so as to minimize the false negative diagnosis in Infarction group. From this point of view, the two sets of the thresholds (the strict one and the mild one) were determined and compared with the score of the joint probability of being Infarction.

**Conclusion**

The diagnostic accuracy in the automated system of ECG interpretation may be largely dependent on the classification method. Based on the results of this comparative study, it may be concluded that the joint probability method seems to be better than the logical decision tree method in the electrocardiographic interpretation.

**Acknowledgement**

We wish to thank Emeritus Prof. Susumu Hibino, Prof. Kazuo Yamada and Dr. Noboru Okamoto for kind advice and helpful criticism.

**REFERENCES**