Using Computerized ECG Measurements in the Bayesian Analysis of Heart Disease

Robert A. Warner, MD*

Tigard Research Institute
12228 SW Chandler Drive
Tigard, OR, 97224 USA
hillwarner@frontier.com

Keywords: systolic dysfunction, ECG data, prior probability

Regular Research Paper

Abstract

The study utilized 1096 sets of diagnostic data from patients who reported to emergency departments with acute shortness of breath. Left ventricular dysfunction (LVSD) was defined as an echocardiographic left ventricular ejection fraction <50% and the diagnostic sensitivities for LVSD at 98% specificity for two tests for LVSD – brain natriuretic peptide (BNP) and the electronically recorded third heart sound (S3) were determined. For both the BNP and the S3 tests, an automated ECG criterion - duration of the QRS complex >120 ms. - identified subgroups with higher prior probabilities of LVSD than the total population. As expected from Bayesian principles, both the BNP and S3 tests exhibited better performances for detecting LVSD in the higher prior probability subgroups than in the total population. Computerized measurement of ECG QRS duration can assess the prior probability of LVSD and help improve the performances of specific tests for LVSD.

1.0 Introduction

The electrocardiogram (ECG) is an inexpensive, convenient and widely available test that is often used in patients with known or suspected heart disease. The nature of the diagnostic information provided by the ECG is very useful for detecting arrhythmias and ischemic heart disease and for suggesting the presence of cardiac chamber enlargement. In contrast, ECG data are considered to be unsuitable for directly detecting hemodynamic abnormalities such as left ventricular systolic dysfunction (LVSD), a condition that is often associated with disabling and potentially lethal heart failure.

However, in keeping with the principles of Bayesian statistics, I conjectured that ECG data could be used to augment tests that are specifically intended to detect LVSD. This is because heart disease often has multiple manifestations whose presence can be suggested by different types of tests, including the ECG. Conversely, most patients who are free of heart disease have normal findings on multiple cardiological tests. Therefore, the prevalence of LVSD is likely to be higher in patients who have even nonspecific ECG evidence for heart disease than it is in patients who lack such evidence. In the terminology of Bayesian analysis, the presence of ECG abnormalities
in a patient increases the prior probability that LVSD is also present in that patient.

In this study, I selected the duration of the QRS complex as the ECG parameter to determine the prior probability that LVSD is present in each member of a population of patients. The QRS complex is the portion of the ECG signal inscribed during electrical activation, i.e. depolarization, of the ventricles. The QRS duration in ms. is routinely reported in ECG interpretations and is a highly reproducible measurement whether it is made visually or algorithmically. The upper limit of normal of the QRS duration is 100 ms., borderline values are 100 to 120 ms. and abnormal values exceed 120 ms. The use of QRS duration is a particularly relevant parameter in the present context because a commonly used type of treatment for LVSD with heart failure – cardiac resynchronization therapy - has been found to be particularly effective if the patient’s QRS duration is abnormally prolonged.1

A test that is often used to identify LVSD directly is the echocardiogram. The most useful echocardiographic parameter for this purpose is the left ventricular ejection fraction. The left ventricular ejection fraction expresses the percentage of the left ventricular end-diastolic blood volume that is ejected from the ventricle during a single left ventricular systolic contraction. The mean normal value of this parameter is about 65% and any value of the ejection fraction less than 50% is considered to indicate that LVSD is present.2

Although the echocardiogram is a definitive test for LVSD, it is expensive, not always readily available for the evaluation of patients and requires special expertise to record and interpret. Therefore, other less expensive and more readily available tests are routinely used to detect LVSD and the heart failure that is often associated with it. One of these is a blood test that measures brain natriuretic peptide (BNP). BNP is elevated in LVSD with heart failure and, in the appropriate clinical context, values of BNP >500 pg/ml are considered to be strong evidence of LVSD with heart failure.3 Also often associated with LVSD is the presence of a third heart sound (S3).4 The S3 is a low pitched sound that originates in the ventricle and occurs in early diastole. The S3 can be heard with a stethoscope or can be detected and recorded using an electronic sound sensor applied to the left side of the chest.5 An advantage of recording the S3 electronically is that its strength can be quantified using the amplitude and frequency that the sound exhibits in a given patient.

Based on the above considerations, I tested the following hypotheses:

- The prevalence, i.e. the prior probability, of LVSD is greater in subgroups of patients with prolonged ECG QRS complexes than it is in the general population.

- In accordance with the principles of Bayesian statistics, the diagnostic performances of the BNP and S3 tests for detecting LVSD are better in populations of patients with a higher prevalence of LVSD than in populations with a lower prevalence of LVSD.

2.0 Materials and Methods

2.1 Selection of Patients

I studied a total of 1096 sets of data from a convenience sample of patients (mean age 61 years, 32% women) who had presented with acute shortness of breath to the emergency department of one of several metropolitan hospitals.

2.2 Diagnostic Tests

In each case, the left ventricular ejection fraction was measured by echocardiography within 24 hours of each patient’s
arrival at the hospital. At the time of arrival at the hospital, 432 patients also had an ECG and electronically recorded heart sounds (Audicor™, Inovise Medical, Inc. Portland, Oregon, USA) and in 374 patients, BNP was measured. The unit of acoustical strength of the S3 was the “display value” - a proprietary parameter that utilizes both the amplitude and the frequency of the recorded sound. Blood levels of BNP were expressed in pg./ml.

2.3 Analysis of the Data

LVSD was considered present if the patient’s left ventricular ejection fraction was <50%. Receiver-operating characteristic curves were used to determine the diagnostic performances of the S3 and BNP in the entire group of patients as well as in the subgroup with ECG QRS duration >120 ms. For both the S3 and the BNP data, chi square analysis was used to compare diagnostic sensitivities at 98% specificity in the entire group of patients vs. the patients with QRS duration >120 ms.

3.0 Results

3.1 BNP Blood Test

Table 1 illustrates the importance of considering the duration of the ECG QRS complex when using BNP to detect LVSD. Table 1 shows that in the entire group of patients in which BNP was measured, the prevalence of LVSD was 48.9% and in the subgroup with prolonged QRS duration, the prevalence of LVSD was 82.8%. In the entire, lower prevalence group of patients, the diagnostic sensitivity of BNP for detecting LVSD was 10.9%. In the subgroup with prolonged QRS duration, the diagnostic sensitivity of BNP for detecting LVSD was 46.8% and this difference was highly statistically significant. Table 1 also shows that in the entire group of 374 patients compared to the subgroup with prolonged QRS duration, the threshold values of BNP needed to attain 98% diagnostic specificity were 1740 and 407 pg./ml., respectively. Furthermore, the ratio of true positive to false positive BNP test results was 5.2 in the entire group and increased to 112.5 in the subgroup with prolonged QRS duration.

3.2 S3 Heart Sound

Table 2 shows the importance of considering the duration of the ECG QRS complex when using electronically recorded heart sound data to detect LVSD. Table 2 reveals that electronically recorded heart sounds were recorded from 432 patients and of this total number of patients, 107 had ECG QRS durations >120 ms. In the entire group, the prevalence of LVSD was 49.5% and in the subgroup with prolonged QRS duration, the prevalence of LVSD was 79.4%. In the entire group of patients, the diagnostic sensitivity of the electronically recorded S3 for detecting LVSD at 98% specificity was 21.0%. In the subgroup with prolonged QRS duration, the diagnostic sensitivity of the recorded S3 for detecting LVSD was 31.8% and this improvement in diagnostic performance reached statistical significance. Table 2 also shows that in the entire group of 432 patients compared to the subgroup with prolonged QRS duration, the thresholds of the recorded S3 display values needed to attain 98% specificity were 5.66 and 4.99, respectively. In addition, the ratio of true positive to false positive S3 test results was 10.3 in the entire group and rose to 27.0 in the subgroup with prolonged QRS duration.

As expected, the prevalence data in the third columns of both Tables 1 and 2 show that the prevalence of LVSD is greater in the subgroup of patients with QRS duration >120 ms. than it is in the entire population...
Table 1
BNP Data for Detecting LVSD

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Prevalence of LVSD</th>
<th>Sensitivity @ 98% Specificity</th>
<th>Chi Square</th>
<th>P Value</th>
<th>Threshold Value*</th>
<th>TP/FP Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Patients</td>
<td>374</td>
<td>48.9%</td>
<td>10.9%</td>
<td></td>
<td></td>
<td>1740</td>
<td>5.2</td>
</tr>
<tr>
<td>QRS &gt;120</td>
<td>183</td>
<td>82.8%</td>
<td>46.8%</td>
<td>41.2</td>
<td>1.4x10^{-10}</td>
<td>407</td>
<td>112.5</td>
</tr>
</tbody>
</table>

*in pg./ml.

BNP = blood natriuretic peptide, FP = false positive, LVSD = left ventricular systolic dysfunction, QRS = ECG QRS complex duration (ms.), TP = true positive

Table 2
S3 Data for Detecting LVSD

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Prevalence of LVSD</th>
<th>Sensitivity @ 98% Specificity</th>
<th>Chi Square</th>
<th>P Value</th>
<th>Threshold Value*</th>
<th>TP/FP Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Patients</td>
<td>432</td>
<td>49.5%</td>
<td>21.0%</td>
<td></td>
<td></td>
<td>5.66</td>
<td>10.3</td>
</tr>
<tr>
<td>QRS &gt;120</td>
<td>107</td>
<td>79.4%</td>
<td>31.8%</td>
<td>3.8</td>
<td>0.05</td>
<td>4.99</td>
<td>27.0</td>
</tr>
</tbody>
</table>

*in proprietary recorded heart sound display values

FP = false positive, LVSD = left ventricular systolic dysfunction, QRS = ECG QRS complex duration (ms.), TP = true positive

Table 3
Comparison of BNP, S3 and QRS Duration Data as Parameters For Detecting LVSD

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Group</th>
<th>LVSD Present</th>
<th>Sensitivity @ 98% Specificity</th>
<th>Threshold Value</th>
<th>Chi Square</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>QRS Duration</td>
<td>All</td>
<td>215</td>
<td>20.0</td>
<td>151</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BNP</td>
<td>All</td>
<td>183</td>
<td>10.9</td>
<td>1740</td>
<td>6.11</td>
<td>0.01</td>
</tr>
<tr>
<td>BNP</td>
<td>QRS&gt;120</td>
<td>77</td>
<td>46.8</td>
<td>407</td>
<td>20.6</td>
<td>5.7x10^{-6}</td>
</tr>
<tr>
<td>S3 DV</td>
<td>All</td>
<td>214</td>
<td>21.0</td>
<td>5.66</td>
<td>.07</td>
<td>NS</td>
</tr>
<tr>
<td>S3 DV</td>
<td>QRS&gt;120</td>
<td>85</td>
<td>31.8</td>
<td>4.99</td>
<td>4.71</td>
<td>0.03</td>
</tr>
</tbody>
</table>

1compared to QRS Duration, 2 in ms., 3 in pg./ml.

BNP = blood natriuretic peptide, DV = proprietary display values, LVSD = left ventricular systolic dysfunction, S3 = third heart sound

of patients. In other words, the presence of QRS duration >120 ms. increases the prior probability that LVSD is present. In keeping with the principles of Bayesian statistics, these higher prior probabilities of LVSD are associated with improved diagnostic performances of both the BNP and the recorded heart sound tests.6

The observed association of prolonged QRS duration with improved diagnostic performances of both BNP and recorded heart sounds raises the question of whether QRS duration itself is a useful parameter for detecting LVSD and the data shown in Table 3 address this possibility. Table 3 indicates that in the entire population of patients, the diagnostic sensitivity at 98% specificity of
QRS duration is significantly better (20.0% vs. 10.9%) than that of the BNP test when the latter is used in the lower prevalence total population. To reach 98% specificity for LVSD, the QRS duration must be markedly prolonged to 151 ms. However, in patients whose QRS durations exceed 120 ms, the sensitivity of the BNP test highly significantly exceeds that of both QRS duration alone and the BNP test applied without regard to QRS duration. Table 3 also shows that QRS duration and the S3 heart sound test have similar diagnostic performances when both parameters are used in the entire population of patients. However, when applied to patients whose QRS duration exceeds 120 ms, the performance of the S3 heart sound test is significantly better than that of the QRS duration. Thus, for both the BNP blood test and the S3 heart sound test for LVSD, the best diagnostic results are obtained when the ECG QRS duration has been used to identify subgroups that have higher prevalences of LVSD than those of the entire population.

4.0 Conclusions

The results of the present study support the hypothesis that a precise and readily available ECG parameter – the duration in ms. of the QRS complex – can be used to identify subgroups of patients in which the prevalence of LVSD is higher than the prevalence of LVSD in the entire population of patients. This is particularly remarkable because in contrast to an asymptomatic “screening” population; the prevalence of LVSD in adults reporting to emergency departments with shortness of breath is already high. Thus, the use of the QRS duration to identify subgroups with a higher prevalence of LVSD has incremental value over using just the clinical circumstances of the population being evaluated. The findings of the present study also support the hypothesis that the results of tests specifically intended to detect LVSD exhibit better diagnostic performances in the subpopulations with higher prevalences of LVSD than in the populations that have lower prevalences of this abnormality. Tables 1 and 2 show these results for the BNP data and for the S3 heart sound data, respectively. A way to emphasize the importance of these findings is from the perspective of considering the meaning of a “positive” result of a test for LVSD in a given patient. The last column of Table 1 shows that for the BNP test, the ratio of true positive to false positive results is 5.2 in the lower prevalence total group compared to 112.5 in the higher prevalence subgroup. This means that a given positive result on the BNP test in the higher prevalence subgroup compared is 112.5/5.2 = 21.6 times more likely to be correct than it is in the lower prevalence total group. The last column of Table 2 shows that a given positive result on the S3 heart sound test in the higher prevalence subgroup is 27.0/10.3 = 2.6 times more likely to be correct than it is in the lower prevalence total group.

The data in Table 3 indicate that using ECG QRS complex duration in ms. itself can be used as a parameter for detecting LVSD, although the diagnostic sensitivity at 98% specificity is only 20.0%. Nevertheless, when applied to the lower prevalence populations, the diagnostic sensitivity of QRS duration alone is better than the BNP test (20.0% vs. 10.9%) and equivalent to the S3 heart sound test in the low prevalence population (20.0% vs. 21.0%). However, in the high LVSD prevalence subpopulations, using QRS duration alone is diagnostically inferior to both BNP and the S3 heart sounds (sensitivities at 98% specificity of 46.8% and 31.8%, respectively). These data show that for detecting LVSD, using the ECG QRS duration to identify subgroups with high prior probabilities of LVSD is superior to using the QRS duration as a standalone diagnostic test.
A previous study from this laboratory showed that one may also assess the prior probability of LVSD using ECG evidence of prior myocardial infarction. However, the use of the QRS duration on the ECG has important advantages. First, there are many different sets of diagnostic criteria for myocardial infarction and the accuracy of each of them is imperfect. Second, the presence or absence of myocardial infarction is irrelevant in the many patients whose LVSD has been caused by pathology other than coronary artery disease. In contrast, the measurement of the QRS complex is ubiquitous, extremely accurate and highly reproducible. QRS duration in ms. is routinely reported in both computerized and visual ECG interpretations. Additional work will be required to determine whether ECG findings other than criteria for myocardial infarction or measurements of QRS duration are superior for assessing the prevalence of LVSD in a given population.

In everyday life, we often informally consider prior probability when trying to decide whether something is true or false. This involves thinking about the background circumstances in which the truth or falsity of an assertion is being evaluated. For example, we tend to be less likely to accept uncritically a salesman’s description of a product if we know that he is trying to sell that product to us. Prior probability is also considered in medical diagnosis, even though such considerations are also often informal. For example, if a particular disease is especially common in elderly women, a physician knows that a positive result on a diagnostic test for that disease is less likely to be correct if the test had been performed on a young male. The present study shows that computerized ECG data can be used in a formal, quantitative way to affect the prior probability that a certain disease is present. This in turn can significantly improve the diagnostic performances of non-ECG tests for that disease. The specific findings of the present study are especially important because LVSD can result from a wide variety of types of underlying heart disease (e.g. coronary artery disease, hypertension, infection, valvular heart disease, cardiomyopathy and congenital heart disease) and it often leads to premature disability and death. Therefore, improvements in the accuracy of diagnostic tests for LVSD can increase the likelihood that patients receive timely and effective treatment for this important condition.

5.0 References

7. Warner RA. Using the principles of bayesian statistics to improve the performances of medical diagnostic tests. Proceedings of the 2014 International Conference on Computa-