Correct Utilization of Exercise Electrocardiographic Leads in Differentiation of Men With Coronary Artery Disease from Patients With a Low Likelihood of Coronary Artery Disease Using Peak Exercise ST-Segment Depression

Jari Viik, LicTech, Rami Lehtinen, PhD, Väinö Turjanmaa, MD, Kari Niemelä, MD, and Jaakko Malmivuo, PhD

In this study we compared the diagnostic characteristics of the individual exercise electrocardiographic leads, 3 different lead sets comprising standard leads and the effect of the partition value in the detection of coronary artery disease (CAD). The diagnostic variable used was ST-segment depression at peak exercise, and the study population consisted of 101 patients with CAD and 100 patients with a low likelihood of the disease. The lead system used was the Mason-Likar modification of the standard 12-lead system and exercise tests were performed on a bicycle ergometer. The comparisons were performed by means of receiver-operating characteristic analysis and by determining sensitivities at a fixed 95% specificity. These properties, defined here as diagnostic capacity, were the most efficacious in leads I, aVR, V4, V5, and V6. Diagnostic capacities in leads aVL, aVF, III, V1, and V2 were quite poor; statistical comparisons indicated significant differences between these leads and lead V5 (p < 0.0001 in each case). Use of the maximum value of ST-segment depression at peak exercise derived from all 12 leads produced a considerable decrease in the diagnostic capacity of the exercise electrocardiogram compared with lead V5. The exclusion of leads aVL, V1, and III improved the diagnostic capacity compared with the 12-lead set, but it was still smaller than that of lead V5. With use of a lead set with the 5 best leads increased the diagnostic capacity over other lead sets and over any individual lead. Further improvement was noted when a 50% smaller partition value was applied to leads I and aVR than for the other leads (p = 0.041). In conclusion, this study suggests that use of leads I, aVR, V4, V5, and V6 is the most influential when differentiating between patients with CAD and patients with a low likelihood of disease using peak exercise ST-segment depression. The effective use of leads I and aVR requires the partition value applied for these leads to be 50% smaller than that used for the lateral precordial leads.

METHODS

Study population: Subjects were selected from a group of 1,507 consecutive patients who underwent routine clinical exercise testing. All patients had been referred for exercise testing at Tampere University Hospital, Finland, and there were no volunteer subjects. All patients selected were men. Patients with left or right bundle branch block and recent myocardial infarction (<8 weeks) were excluded.
The CAD group was selected from patients who underwent selective coronary angiography (performed using the Judkins technique) within 180 days of exercise testing. In all cases, each coronary artery was imaged in multiple views. The degree of stenosis was defined as the greatest percent reduction in luminal diameter in any view compared with the nearest normal segment. CAD was considered significant when ≥50% luminal narrowing of the diameter of the major epicardial coronary arteries was present. After these exclusions there were 101 men included in the CAD group. Of these, 44 had significant stenosis in all 3 major coronary arteries or in the left main coronary artery, 25 had 2-vessel disease, and 32 had 1-vessel disease.

The reference group was selected on the basis of previous clinical history. The inclusion criteria for the reference group were no history of any cardiac disease, a normal resting electrocardiogram, and no anginal type chest pain or cardiac medication. After these restrictions, 100 male patients were available for use in the study. With probabilistic assessment, the reference group was estimated to have a low likelihood (p <0.05) of CAD.7

**Exercise electrocardiographic test:** The exercise test was performed on a bicycle ergometer using a computerized recording system (SYSTEM II EXES, Siemens-Elema, Solna, Sweden). The graded protocol followed a standard clinical routine with an initial workload of 50 W and an increment of 50 W every 4 minutes. The exercise tests were sign- and symptom-limited maximal tests using recommended criteria for termination; fatigue and chest pain were the reason for termination in most cases. The lead system used was the Mason-Likar modification of the standard 12-lead system.9 Computer-determined ST-segment amplitudes measured to the nearest 10 μV were obtained at a point 60 ms after the J junction,10,11 with the end of the PR segment considered as the isoelectric line. The ST-segment amplitude, heart rate, and workload data were stored for further processing and analysis. The ST-segment value at the peak exercise (STpeak) was used as a diagnostic classifier.

**Electrocardiographic leads, lead sets, and partition values used in the analysis:** All individual leads of the 12-lead system were used separately in the detection of ischemic responses. Lead aVR was inverted so that it had a lead vector directed to the left and downward. Thus, a positive response to the exercise test also corresponded to ST-segment depression. The importance of the number of leads was studied with 3 different lead sets. The maximum STpeak value, determined from the leads of each lead set, was used as the diagnostic classifier. The 3 lead sets used (1) all 12 leads, (2) 9 leads (aVL, III and V1 were excluded), and (3) 5 leads (I, -aVR, V4, V5, and V6) of the standard 12-lead electrocardiogram, and were denoted by A12, A9, and A5, respectively. Selection of the leads for leads sets A9 and A5 was based on the diagnostic capacities of the individual leads. Different partition values were applied for the particular leads in order to define the effect of more detailed criteria on the sensitivity and specificity of the test.

**Data analysis and statistical methods:** Continuous variables are described as mean ± SD. Significant differences among study groups with respect to cardiac medication, chest pain, and previous myocardial infarction were examined using the chi-square test with Yates’ correction. Quantitative variables were analyzed using a 2-tailed Student’s t test. Comparisons of the sensitivity of individual leads and different lead sets at a fixed 95% specificity were performed using McNemar’s modification of the chi-square method for paired proportions. Due to the fact that the sensitivity and specificity are dependent on the partition values chosen for test positivity, the diagnostic accuracy of the study variables without any partition value were also compared by means of receiver-operating characteristic (ROC) analysis. The area under the ROC curve represents the overall diagnostic performance, i.e., the probability that a random pair of patients with and without CAD will be correctly diagnosed.12 Statistical differences between the areas under 2 ROC curves were compared using a nonparametric analysis of correlated ROC curves13 with a routine written by Vida14 (version 2.5). In statistical comparisons, individual leads were compared with lead V5. Lead sets A12, A9, and A5 were compared with each other and with lead V5. When defining the alpha level of the tests, the Bonferroni correction was taken into account in all statistical comparisons.15

**RESULTS**

Group characteristics and exercise performance:

Group characteristics and statistical differences according to clinical status are listed in Table I. Because
of the different exclusion criteria used in the selection of the groups, highly significant differences (p < 0.0001) were achieved with respect to age, maximal workload, maximum heart rate, chest pain, previous myocardial infarction, and cardiac medication, with the exception of digitalis.

**Individual leads:** The descriptive statistics (mean, SE, and SD) of the ST peak in each lead are shown in Figure 1. The order of the limb leads is in accordance with lead direction in the frontal view. The ST peak values of the CAD group and the reference group are illustrated side by side, starting with the CAD group. The mean values of ST peak were lower in the CAD group than in the reference group in each lead, with the exception of lead V1.

The ROC curves for each individual lead are shown in Figure 2. The circles indicate −0.10 mV ST peak (1.0 mm ST-segment depression) and the diamonds indicate the nearest partition values providing 95% specificity. The value of <0.500 in lead V1 indicated that lead V1 should have been inverted in this study population (the ROC area for −V1 would be 0.577). Statistical comparison of the leads showed that the areas under the ROC curves in leads aVL, aVF, III, V1, and V2 were significantly smaller than in lead V5 (p ≤ 0.0001 in all cases). With use of a corrected alpha level of 0.0045 for the univariate Z test, significant differences were not observed between the ROC curves of leads V5 and V2 (p = 0.0007). Moreover, no significant differences were detected when comparing leads I, −aVR, V4, and V6 with lead V5.

The sensitivity values obtained at 95% specificity showed statistical differences when comparing leads III, aVL, aVF, V1, and V2 with lead V5 (in all cases p < 0.0001), lead II with V5 (p = 0.0004) and lead V3 with V5 (p = 0.0007), but not in the case of leads I, −aVR, V4, and V6.

**Lead sets:** The ROC curves determined using the maximum values from lead sets A12, A9, and A5 are shown in Figure 3. The circles indicate −0.10 mV ST peak and the diamonds indicate the nearest partition values providing 95% specificity. By comparing the areas under the ROC curves, significant differences (alpha level of 0.017) were found between A5 and A9 (p = 0.0152), A5 and A12 (p < 0.0001), and A9 and A12 (p = 0.0001). When comparing the different lead sets with lead V5, a significant difference was only observed in the case of V5 and A12 (p = 0.0023). No significant differences were observed between A5 and V5 (p = 0.0906) or V5 and A9 (p = 0.4181).

By comparing the sensitivities at fixed specificity between lead sets, significant differences (alpha level of 0.017) were detected between A5 and A12 (p < 0.0001), A5 and A9 (p = 0.0060), and A9 and A12 (p = 0.0008). A significant difference was also found when comparisons were performed between lead set A5 and lead V5 (p = 0.0133) and between V5 and A12 (p = 0.0003), but no significant difference was observed between V5 and A9 (p = 0.4227). The partition values corresponding to a specificity of 95% were dissimilar. The partition value decreased proportionally when the number of leads in the set decreased (from −0.14 mV to −0.06 mV).

The specificity and sensitivity values of lead sets A12, A9, and A5 at different partition values are shown in Figure 4 and are denoted by A to D. For cases A and C the specificity and sensitivity values were determined using different partition values for leads I and −aVR (first value in the positive test criteria keybox) and for the other leads (second value); for example, for cases labeled C in Figure 4, a partition value of −0.05 mV for test positivity was applied to leads I and −aVR and a partition value of −0.10 mV was applied to the other leads. The distribution of

**FIGURE 1.** Standard deviations (SD), standard errors (SE) and means of the ST-segment values at peak exercise lead by lead. Shaded symbols indicate the CAD group and open symbols patients with a low likelihood of the disease. ST peak = ST-segment value at peak exercise.
the ST_{peak} values in each lead in the CAD group and reference group (SD in Figure 1) was taken into account when defining lead specific partition values. Using the −0.05 mV criterion for leads I and −aVR, the sensitivity improved by 5 percentage points. Further improvement in sensitivity was achieved when the partition value for test positivity was reduced, but at the same time the specificity of the test decreased. This decrease in specificity was smaller in lead set A5 than in lead sets A9 or A12. The specificity of the lead sets at different partition values indicated that lead set A5 was superior to lead sets A9 and A12. Comparing the traditional criterion D (ST_{peak} from all leads $\leq -0.10$ mV) and criterion C (ST_{I,-aVR} $\leq -0.05$ mV or ST_{other} $\leq -0.10$ mV), a significant difference was observed in lead set A5 ($p = 0.0412$).

**DISCUSSION**

**Individual leads:** Many studies\textsuperscript{16–19} have shown lead V$_5$ to be capable of detecting most ischemic responses when a positive test criteria of $\geq 0.10$ mV ST-segment depression is used. According to this study a fixed global partition value applied to each standard lead does not treat individual leads equally. Figure 1 shows that it is easy to see that larger partition values are most suitable for the lateral precordial leads (V$_4$, V$_5$, and V$_6$). The highest sensitivities at a partition value of $-0.10$ mV were found in leads V$_5$ and V$_6$ (Figure 2), supporting the previous statement. Using the same $-0.10$ mV criterion, leads I and −aVR attained extremely poor sensitivities. A model study\textsuperscript{20} and our previous clinical study\textsuperscript{21} also support these findings. However, the areas under ROC curves of these leads indicated diagnostic performances equally as good as those of the lateral precordial leads, and the sensitivities at a fixed 95% specificity did not significantly differ from lead V$_5$. The only difference is that the partition value for these leads should be 50% smaller.

Considering the diagnostic capacity of individual leads, the results of this study support incontestably the previous finding\textsuperscript{6,21} that leads aVL and V$_1$ are unreliable in the general detection of CAD. The areas under the ROC curves were <0.600 and the sensitivities at 95% specificities were the most unsatisfactory. Additionally, the diagnostic capacities of leads aVF, III, and V$_2$ were also deficient and significantly smaller than in lead V$_5$.

**Multiple leads:** Several studies have demonstrated an improvement in the detection of CAD using multiple leads during the exercise test.\textsuperscript{16,22–26} Increasing
the number of leads used in the analysis of ischemic response detection increases the sensitivity, but often a problem arises due to an increase in false-positive responses, and the diagnostic accuracy of the lead set does not necessary increase. When using the maximum value from all 12 leads (lead set A12) in the analysis of CAD detection, the diagnostic capacity of the ST peak diminished compared with lead V5 or lead sets A9 and A5. When using the same fixed partition value for different lead sets (e.g., $-0.10 \text{ mV}$ in Figure 3), it can be seen that the deterioration in diagnostic capacity was caused by a decrease in specificity (the sensitivity did not change). Thus, use of a larger partition value (from $-0.06 \text{ mV}$ to $-0.14 \text{ mV}$) is recommended if high specificity is desired. The exclusion of the poorest leads (aVL, III, and V1) from the analysis increased the area under the ROC curve as well as the sensitivity at 95% specificity, but still the diagnostic capacity of lead set A9 was approximately the same as the diagnostic capacity of the best individual leads. This study indicated that the use of lead set A5 improved the diagnostic capacity of the ST peak significantly compared with lead sets A9 and A12. Furthermore, the area under the ROC curve for lead set A5, as well as its sensitivity at a fixed 95% specificity, was more competent than the corresponding values for any individual lead.

**Study limitation:** There are several limitations to this study. The study population was restricted to men. Thus, the conclusions may not be directly applicable to women. The CAD patients had angiographically proven CAD, but the reference patients were defined only by clinical history. In an ideal study, the entire study population would have been examined by angiography. However, taking into account the more frequent use of exercise ECG testing as a screening test in large populations, this type of approach with a bipartite study population is appropriate. Cardiac medication, the type of exercise test, and the protocol influenced the results. However, these influences can be assumed to be identical for each lead and thus the comparative performance of different leads would be the same regardless of medication, exercise modality, or protocol.

**Acknowledgment:** We thank Simon Walker, PhD, and James Rowland for their valuable suggestions and assisting in preparation of this manuscript.

3. Detrano R, Gianrossi R, Froelicher V. The diagnostic accuracy of the exercise


