ORIGINAL ARTICLE

A New 12-Lead ECG Prognostic Score

Muhammad Soofi, B.S., B.A.,* Nikhil A. Jain, B.A.,† Jonathan Myers, Ph.D.,†‡ and V.F. Froelicher, M.D.†‡

From the *Ohio State University, College of Medicine, Columbus, OH; †Stanford Cardiovascular Institute/The Division of Cardiovascular Medicine, Department of Medicine, Stanford University School of Medicine, Stanford, CA; and ‡Veterans Affairs Palo Alto Health Care System, Palo Alto, CA

Background: The Diagonal Line Rule (DLR) is an educational pattern recognition tool for the classification of 12-lead electrocardiograms (ECGs) as normal or abnormal.

Objective: To develop a score using the DLR for the risk stratification of abnormal ECGs.

Methods: Patients at the Palo Alto, VA, referred for a routine resting ECG between 1987 and 1999 were retrospectively considered. The initial ECGs of 43,798 patients were scored according to the DLR. The main outcome variable was cardiovascular mortality with 8 years of follow-up.

Results: ECG abnormalities identified by the DLR were combined with age, gender, and ST depression to develop an additive score. The score was validated and shown to stratify levels of cardiac risk with a 0.78 area under the receiver operating characteristic (ROC) curve for identifying those who had a cardiovascular death.

Conclusion: The score effectively stratifies ECGs for prediction of risk of cardiovascular death.

Ann Noninvasive Electrocardiol 2015;20(6):554–560

cardiovascular disease; electrocardiography; scores; prognosis

The Diagonal Line Rule (DLR) is a pattern recognition tool proposed by Sibbitt for teaching interpretation of the 12-lead electrocardiogram (ECG).¹ The DLR requires that a diagonal line be drawn through leads III, aVL, and V₁ to indicate that diagnostic Q waves and inverted P or T waves can be normal findings in these leads, while abnormal in all other leads. An accompanying box is drawn around aVR indicating that P and T waves and the QRS complex should be net negative (Fig. 1).

The DLR has been shown in a previous study to be effective in rapidly classifying ECGs as either normal or abnormal.² It was also demonstrated to have characteristics that could be used to stratify ECGs in gradations of risk rather than simply classify them in binary "normal" and "abnormal"

groups. For example, DLR abnormalities in the high lateral leads carried a higher risk of cardiac mortality relative to abnormalities in the anterior leads. This feature could be used to stratify ECGs by risk, with high lateral lead abnormalities being weighted at a higher risk level than inferior lead abnormalities. The predictive power of the DLR could also be improved by including other markers of cardiac risk, including age and gender as well as ECG characteristics such as ST depression.

In this study, we sought to develop a comprehensive weighted score combining the DLR criteria of abnormal Q, P, or T waves with other proven cardiac risk markers, namely, age, gender, and ST depression.^{3,4} We hypothesized that the resulting score would be able to simply and effectively stratify ECGs as to risk.

Address for correspondence: Nikhil Jain, B.A., 19902 Seagull Way Saratoga, CA 95070. Tel: 408-489-8403; E-mail: nikhiljain@wustl.edu

There were no contributors besides the authors. No financial support was used for this study. These data have not been presented previously.

Conflict of interest: Dr. Froelicher is co-owner of Cardea Associates, Inc., developers of hardware and software for screening for CV risk. Dr. Myers, Muhammad Soofi, and Nikhil Jain have no conflicts to disclose.

© 2015 Wiley Periodicals, Inc.
DOI:10.1111/anec.12261

Figure 1. Illustration of the diagonal line rule (DLR), applied only to ECGs with a QRS duration ≤ 120 ms.

METHODS

Clinical Population

We performed a retrospective study of 45,829 recorded ECGs between March 1987 and December 1999 at the Veterans Affairs Palo Alto Health Care System. Since 1987, the Veterans Affairs Palo Alto Health Care System has used a centralized computerized ECG system for collection, storage, and analysis of ECGs (GE Healthcare, Wauwatosa, WI, USA). This system has been approved by the United States Food and Drug Administration and the European Union, and is widely used around the world. Race was determined by self-report at the time of ECG acquisition. We considered the first ECG obtained for each patient, and excluded ECGs with one or more of the following findings: paced rhythms ($n = 290$), ventricular preexcitation ($n = 42$), QRS duration ≥ 120 ms without right bundle branch block ($n = 1699$), leaving 43,798 patients for inclusion in our development population. Similar exclusion criteria were applied in a separate population to yield the validation population. ECGs were ordered at the discretion of the ambulatory clinic health care provider and were most often gathered for preventive medicine purposes without symptomatology or diagnosis provided in the order. Unfortunately, additional clinical data were not included since the electronic data system used went through multiple iterations in the study time period and paper medical records were no longer available.

ECG Analysis

ECGs measurements and diagnostic statements were made by the 12-SL software (GE Healthcare). All the ECGs were overread by the senior author (VF) as part of clinical practice with the onset and offset points of the waveforms confirmed. The amplitude and duration measurements of the Q and T waves, based on the computer algorithms, were used to code the Q and T waves. This program is one of the most widely used automated programs and has been well validated.⁵ In those with QRS duration less than 120 ms or with right bundle branch block, T-wave inversion was coded for abnormal if ≤ -0.1 mV and Q waves for abnormal if ≥ 40 ms in all leads except aVR where abnormal was considered a T wave ≥ 0.1 mV and an R wave ≥ 40 ms (the inverse of the other leads). ST depression was considered ≤ -0.5 mm flat or down slope in V_5 as well as the more strict criteria of contiguous ST Depression in V_4 , V_5 , or V_6 . Leads were grouped as V_2 and V_3 (anterior); aVF and II (inferior); V_4 , V_5 , and V_6 (lateral); and I and aVL (high lateral).

Outcomes

The primary outcome variable was time to cardiovascular death. The California Health Department Service and Social Security Death Indices were used to ascertain the vital status of each patient as of December 31, 2002. Accuracy of

Table 1. General Characteristics of the Total Population

	Total Population	Development Population	Validation Population
n	45,497	43,798	20,799
Mean age (years)	56.7 ± 14.7	56.3 ± 14.6	43.3 ± 8.0
Mean height (inches)	68.69 ± 3.7	68.7 ± 3.7	69.2 ± 3.6
Mean weight (lbs)	182.5 ± 40.1	182.4 ± 40.1	187.6 ± 41.7
Women	10.0% (4571/45,497)	10.2% (4474/43,798)	9.5 % (1969/20,799)
African American	12.0% (5334/45,497)	12.3% (5212/42,525)	16.4% (3409/20,799)
Atrial fibrillation	2.7% (1253/45,497)	2.6% (1158/43,798)	0.6% (114/20,799)
Right bundle branch block	3.4% (1601/45,497)	3.7% (1601/43,798)	0.9% (193/20,799) (RBBB)
Left bundle branch block	1.3% (600/45,497)	0.0% (0/43,798)	0.0% (0/20,799) (LBBB)
Diagnostic anterior Q waves	2.6% (1186/45,497)	2.5% (1083/43,798)	1.3% (279/20,799)
Diagnostic inferior Q waves	8.3% (3803/45,497)	8.1% (3547/43,798)	4.5% (925/20,799)
Diagnostic lateral Q waves	0.9% (394/45,497)	0.7% (320/43,798)	0.5% (95/20,799)
ST depression in V ₅ (≤ −0.5 mm flat or down slope)	7.0% (3227/45,497)	5.9% (2595/43,798)	0.5% (106/20,799)
ST depression in V ₄ , V ₅ , or V ₆ contiguous (≤ −0.5 mm flat or down slope)	1.9% (901/45,497)	1.5% (651/43,798)	0.4% (76/20,799)
Annual all-cause mortality	3.4%	3.3%	1.3%
Annual CV mortality	1.2%	1.1%	0.3%

all causes of death was confirmed using the Veterans Affairs Computerized Patient Record System (CPRS) by two physicians blinded to the ECG results. The CPRS record system enables access to data from all other VA's to account for situations in which the veteran moves. Clinical visits and testing, problem lists, autopsies, and discharge/death summaries were used to specify cardiovascular cause of death.

point of cardiovascular death were calculated for the risk groups derived from the score. The score was further tested as a continuous variable using receiver operating characteristic (ROC) statistics. The developed score was validated in a younger veteran population (n = 20,799, mean age 53) with a longer follow-up time (mean of 16.9 ± 6.04 years).

RESULTS

Statistical Analysis

NCSS software 2007 (NCSS, Kaysville, UT, USA) was used for all statistical analyses. Survival analysis was performed using Cox Proportional Hazards Models and the method of Kaplan and Meier for univariate plots and annual survival. Score development was completed through analysis of correlation coefficients and Cox Hazard models including age, gender, ST depression, and abnormal Q or T waves in the different lead groupings. When calculating the hazards of the risk groups derived from the score, the Cox Proportional Hazards model was not adjusted for age or gender because the developed algorithm contained both variables and including them again in the Cox Hazard model would be statistically redundant. The proportional hazards assumption was confirmed using scaled Schoenfeld residuals. Kaplan-Meier plots and Cox models with the end

Population Characteristics

Table 1 compares the general population with the two populations used to develop and validate the score (titled development and validation population, respectively). The development population resembled the general population in demographic characteristics and disease prevalence. There were no statistical differences between the development and general population, except for prevalence of left bundle branch block (LBBB), which was low in the development population due to the exclusion criteria. The development population was 10.2% women and 12.3% African American. The validation group was statistically different than the development group in demographic characteristics and disease prevalence. This difference was desired to establish the external reliability of the score in different groups. Compared to the development group, the validation group was

Figure 2. Flow chart for DLR score calculation.

Figure 3. Plot of annual cardiovascular (CV) mortality for score points.

younger (mean age: 43.3 ± 8.0) and had far lower prevalence of all measured disease characteristics. The validation group had correspondingly low annual all cause and cardiovascular mortality rates.

Score Development

A score including age, gender, ST depression, and abnormal Q or T waves in the different lead groupings was developed from the regression coefficients and hazard ratios of the Cox model (data not shown). The derived score was $(3 \times (\text{Age Bracket})) + (3 \times (\text{Gender} = "M")) + (2 \times (\text{High Lateral})) + (2 \times (\text{Inferior})) + (3 \times (\text{Lateral})) + (\text{ST Depression}) + (\text{Anterior})$. Here, the presence of abnormalities in any of the lead groupings would be coded as "1" and the absence would be coded as "0." This yields a maximum value of 27 and a minimum of 3 (Fig 2.).

The correlation between score value and cardiovascular mortality was studied to evaluate the

predictive power of the score. As shown in Figure 3, score value was able to closely predict cardiovascular mortality and the score value increased relative to cardiac mortality rate. The relationship was best modeled by a second-degree polynomial with an R^2 of 0.98. This plot can be used to determine the annual cardiovascular mortality for any score value.

Risk Groups

The 24 strata defined by the score were further classified into no, low, medium, and high-risk groups through survival analysis with Kaplan-Meier plots and Cox Hazards ratios. Table 2 reports the characteristics of each risk group. The "No Risk" group had a Cox Hazard of 1.00 and higher risk groups were associated with incrementally greater Cox Hazards, with the "High Risk" group having a Cox Hazard of 38 (CI: 30.3–48.0, P value < 0.0001). Both all-cause and cardiac annual

Table 2. Characteristics of the Four Different Risk Groups: Cox Hazard and Annual Mortality Rates Associated with the Diagonal Line Rule Score

	n (%)	Cox Hazard Ratio		Annual Mortality Rate		
		Cox Hazard (CI)	P Value	Cardiovascular Mortality	All-Cause Mortality	Cardiovascular Death (%)
No Risk	7044 (16.1%)	1.0 (0.8–1.4)	<0.0001	0.1%	0.9%	77 (1.09%)
Low Risk	18,150 (41.4%)	4.4 (3.5–5.6)	<0.0001	0.6%	2.1%	714 (3.93%)
Medium Risk	10,214 (23.3%)	15.6 (12.4–19.6)	<0.0001	1.8%	4.6%	1292 (12.65%)
High Risk	8390 (19.2%)	38.2 (30.3–48.0)	<0.0001	3.8%	6.9%	1876 (22.36%)

Figure 4. Kaplan-Meier survival curves for the four different risk groups.

mortality rates increased as the risk group increased and the "High Risk" population had rates significantly greater than that of the general population.

Figure 4 shows the Kaplan-Meier results with end point of cardiovascular death for the survival analysis of the four risk groups. The "No Risk" group has an expected survival curve with little decay over the follow-up and the "High Risk" group has a separation of over 55%. There is clear separation between the four strata, further supporting the classification of the score into four risk groups.

Sensitivity Analyses

The sensitivity, specificity, and discriminatory power of our score in predicting cardiovascular death were evaluated through the ROC shown in Figure 5. The ROC had an area under the curve (AUC) of 0.78, indicating that its discriminatory power was comparable to other scores estimating cardiovascular death used in medicine.^{6–8}

Validation

The accuracy of the developed score was validated in a secondary population. The secondary

Figure 5. Receiver operating curve (ROC) for the developed score.

sample pool had similar demographic characteristics as the primary population (9.5% women, 16.4% African American). However, it had a younger mean age (43.3 ± 8 years) and had correspondingly lower disease prevalence. The score and the four risk classifications were able to effectively stratify the risk of the population and Kaplan-Meier analysis showed clear separation between risk groups, validating the discriminatory power of the score. The ROC curve had an AUC of 0.68 (Fig. 6).

DISCUSSION

The resting ECG is readily available, noninvasive, and inexpensive, making it one of the most widely used cardiovascular diagnostic tests. Many attempts have been made to study the utility of the ECG for screening and risk stratification.^{9–11} In the present study, we developed a linear, additive score for the effective risk stratification of ECGs using ECG abnormalities coded by the DLR and known cardiac risk markers (i.e., age, gender, and ST depression). Score values were highly proportional to cardiovascular mortality rates and could be used to explicitly calculate the expected cardiovascular mortality rate for those risk factors. The 24 strata can be grouped into four major risk

Figure 6. Receiver operating curve (ROC) for the developed score in a secondary population for validation.

groups of no, low, medium, or high risk. The separation of these groups in Kaplan-Meier survival analysis plots and the incremental increase in Cox Hazard Ratios and cardiovascular mortality rates indicate that the classification into these groups is justified. The discriminatory power of the score rivals others used in medicine as demonstrated by a 0.68–0.78 area under the ROC curve.^{6–8}

Validation of the developed score through Kaplan-Meier and ROC analysis in a secondary population confirmed its discriminatory power and supports its use in other populations. The incremental rise in Cox hazard ratios when comparing different risk strata mirrors the significant separation of survival curves observed in Kaplan-Meier plots and a parallel trend in mortality rates. However, the markedly high Cox hazard ratio in the High Risk group remains an outlier. The cardiovascular mortality in the No Risk group is shown to be 0.1%, compared to 3.8% in the High Risk group. Given the small magnitude of risk carried by the No Risk group and the significant difference in risk levels between the No Risk and High Risk groups, it is possible that the magnitude of risk in the High Risk group may be artificially inflated. Thus, while the High Risk group hazard is likely smaller, our findings support a significant hazard carried by this group.

The developed score is applicable as an educational tool for learning the ECG and augments

the DLR by providing greater diagnostic power and risk discrimination in stratifying abnormal ECGs. The DLR dichotomizes ECGs as either "normal" or "abnormal." Our findings add clinical applicability to the DLR by providing an algorithm through which users can ascertain the level of risk of an abnormal ECG and calculate the specific rate of mortality. Various ECG scores and criteria have been used to predict cardiovascular mortality, including the Minnesota code,¹² Cardiac Injury Index damage (CIIS) score,¹³ and other indices.¹⁰ Unfortunately, these scores are relatively complicated, thus limiting their practical utility in the clinical setting. Simplification of ECG interpretation would be a great benefit for health care professionals. The DLR proposed by Sibbitt has been validated and provides a pattern recognition tool to quickly identify abnormal 12-lead ECGs. Our developed score extends the utility of the DLR by providing a method to further stratify the risk level of abnormal ECGs identified by the DLR and directly calculate the specific cardiovascular mortality rate associated with that stratum.

The gradations of risk identified by the DLR reflect the anatomy and physiology of the heart. Leads III, aVL, and V₁ view the electrical activity of the heart from above the rim of the left ventricle with no myocardium underneath to generate an R wave or subsequent upright T wave. These three leads are correspondingly less informative and form the diagonal line of the DLR. However, all the other leads overlies myocardial tissue that can generate a net positive QRS complex and T wave.² The lateral and high lateral leads are noted to carry the greatest risk because their vector angles capture most of the left ventricular myocardium. The developed and validated score in this study could be valuable in medical education, and allow those beginning their ECG training to approach the 12-lead ECG with confidence. Using this score, ECG students would be able to not only identify abnormal ECGs, but also predict the risk level of a patient.¹⁴

One potential drawback of diagnostic scores is their reliance on the uniformity of the populations in which they were developed and tested. As is typical of data from Veterans Administration systems, our data represent a predominantly male population and thus, our findings may have limited external validity in female populations. A limitation in the study was the lack of clinical data on associated cardiovascular comorbidities that could

potentially be confounding variables. However, the exclusion of these variables does not negate the risk associated with documented abnormal ECG findings since other studies including these variables have still demonstrated the risk of ECG variables.⁶

CONCLUSION

Multivariate clinical scores represent a tool to aid clinical assessment and evidence-based clinical decision making. We propose another ECG score that can be applied using the DLR, a tool that can be helpful in teaching electrocardiography.

REFERENCES

1. Sibbitt W. The diagonal line rule: A useful mnemonic in ECG interpretation. *Am Heart J* 1984;108(6):1591.
2. Soofi M, Yong C, Froelicher V. Justification of an introductory ECG teaching mnemonic by demonstration of its prognostic value. *Am J Med* 2014; 127(12): 1202-7.
3. Mendis S, Puska P, Norrving B. Global Atlas on Cardiovascular Disease Prevention and Control. World Health Organization (in collaboration with the World Heart Federation and World Stroke Organization), Geneva, 2011.
4. Muramoto D, Singh N, Aggarwal S, et al. Spectrum of ST amplitude: Athletes and an ambulatory clinical population. *J Electrocardiol* 2013;46(5):427-433.
5. Kligfield P, Badilini F, Rowlandson I, et al. Comparison of automated measurements of electrocardiographic intervals and durations by computer-based algorithms of digital electrocardiographs. *Am Heart J* 2014;167(2):150-159.
6. Ashley EA, Raxwal V, Froelicher V. An evidence-based review of the resting electrocardiogram as a screening technique for heart disease. *Prog Cardiovasc Dis* 2001;44(1):55-67.
7. Geissler HJ, Holz P, Marohl S, et al. Risk stratification in heart surgery: Comparison of six score systems. *Eur J Cardiothorac Surg* 2000;17(4):400-406.
8. Lip GY, Nieuwlaat R, Pisters R, et al. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: The Euro Heart Survey on atrial fibrillation. *CHEST* 2010;137:263-272.
9. Tan SY, Sungar GW, Myers J, et al. A simplified clinical electrocardiogram score for the prediction of cardiovascular mortality. *Clin Cardiol* 2009;32:82-86.
10. Ashley EA, Raxwal VK, Froelicher VF. The prevalence and prognostic significance of electrocardiographic abnormalities. *Curr Probl Cardiol* 2000;25:1-72.
11. Richardson K, Engel G, Yamazaki T, et al. Electrocardiographic damage scores and cardiovascular mortality. *Am Heart J* 2005;149:458-463.
12. Prineas RJ, Crow RS, Blackburn H. The Minnesota Code Manual of Electrocardiographic Findings. Boston, MA, John Wright PSG Inc., 1982, pp. 223-229.
13. Rautaharju PM, Warren JW, Jain U, et al. Cardiac infarction injury score: An electrocardiographic coding scheme for ischemic heart disease. *Circulation* 1981;64:249-256.
14. Loewen P. SPARC—Stroke Prevention in Atrial Fibrillation Risk Tool." Web. 5 Aug. 2014. <<http://www.sparctool.com/>>.