Quantitative electrocardiographic measures and long-term mortality in exercise test patients with clinically normal resting electrocardiograms

Eiran Z. Gorodeski, MD, MPH,[a] Hemant Ishwaran, PhD,[b] Eugene H. Blackstone, MD, FACC,[a] and Michael S. Lauer, MD, FACC, FAHA[c] ***Cleveland, OH, and Bethesda, MD***

Background Currently, the only function of the resting electrocardiogram (ECG) in patients referred for exercise testing is to determine whether imaging is mandated. It is unknown if subtle ECG findings in those patients with clinically normal resting ECGs have prognostic significance.

Methods We performed a single-center cohort study of 18,964 patients without known cardiovascular disease who had a clinically normal resting ECG and who underwent treadmill exercise testing for evaluation of suspected coronary artery disease. Eleven quantitative ECG measures relating to heart rate, conduction, left ventricular mass, or repolarization were collected digitally. The primary outcome was all-cause mortality. The prognostic importance of a composite ECG score was assessed by measuring its impact on the c-index (analogous to area under receiver operating characteristic curve) and by measures of reclassification.

Results During a median follow-up of 10.7 years, 1,585 patients died. The 4 most predictive digital ECG variables were higher ventricular rate, more leftward QRS axis, more downward ST-segment deviation, and longer QT interval. The ECG score was independently associated with mortality (75th vs 25th percentile hazard ratio 1.36, 95% confidence interval 1.25-1.49, $P < .0001$). The ECG score had modest impact on discrimination (change in c-index 0.04) and reclassification of risk (3.0% decrease of relative integrated discrimination improvement, $P < .001$).

Conclusions Subtle ECG findings relating to heart rate, conduction, left ventricular mass, or repolarization in patients with clinically normal ECGs referred for exercise testing may provide modest additional prognostic information over and above clinical and exercise measures. (Am Heart J 2009;158:61-70.e1.)

The primary purpose of the resting electrocardiogram (ECG) before exercise stress testing is to identify patients who have baseline abnormalities that preclude interpretation of exercise-induced ST-segment changes.[1] It is unknown whether, among these patients who have clinically normal ECGs, more precise ECG measures may improve estimation of prognosis.

Prior studies demonstrated that subtle ECG findings relating to heart rate (ventricular rate), conduction (P wave duration, PR interval, QRS duration), left ventricular mass (Sokolow-Lyon voltage, Cornell voltage, QRS axis), and repolarization (ST-segment deviation, ST-segment slope, QT interval, T-wave amplitude) have prognostic significance in observational cohorts,[2-6] patients with hypertension[7] or heart failure,[8] patients resuscitated after cardiac arrest,[9] patients referred for coronary artery bypass grafting,[10] and in patients enrolled in clinical trials.[7,11,12] Data are sparse regarding the potential importance of subtle ECG findings in patients referred for exercise testing, especially those with clinically normal ECGs.

Exercise testing is widely recognized as a powerful yet simple prognostic tool. We previously published and externally validated a prognostic model[13] composed of 12 demographic (age, gender, history of current or recent smoking), clinical (history of non–insulin-treated diabetes mellitus, insulin-treated diabetes mellitus, hypertension, and typical angina), and exercise test (exercise capacity, ST-segment depression during test, test angina, abnormal heart rate recovery, frequent ventricular ectopy in recovery) measures that performed well in predicting mortality. In the current study, we examine in a subset of the same cohort whether subtle ECG measures can augment this model's performance.

From the [a]Heart and Vascular Institute, Cleveland Clinic, Cleveland, OH, [b]Department of Quantitative Health Sciences, Cleveland Clinic, Cleveland, OH, [c]Division of Prevention and Population Sciences, National Heart, Lung, and Blood Institute, Rockledge Center II, Bethesda, MD.

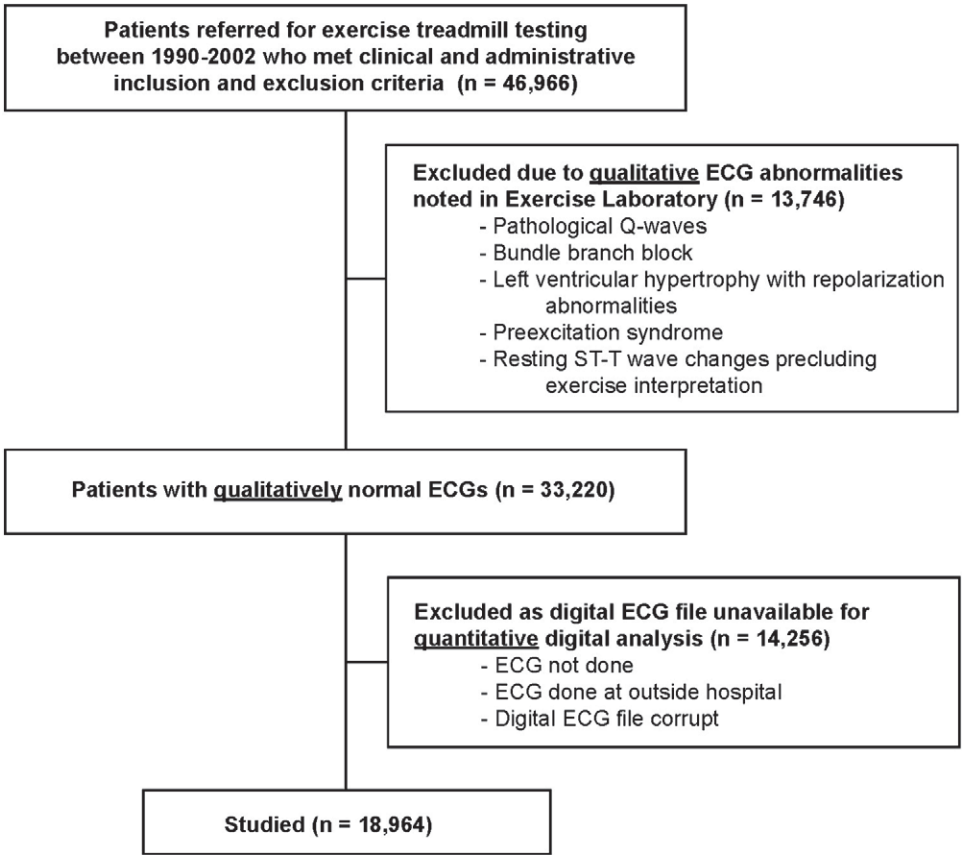
Submitted January 21, 2009; accepted April 17, 2009.

Reprint requests. Michael S. Lauer, MD, FACC, FAHA, Division of Prevention and Population Sciences, National Heart, Lung, and Blood Institute, Rockledge Center II, 6701 Rockledge Drive, Room 10122, Bethesda, MD 20892.
E-mail: lauerm@nhlbi.nih.gov

0002-8703/$ - see front matter

doi:10.1016/j.ahj.2009.04.015

Figure 1



Flow diagram of patient recruitment and participation.

We prospectively studied patients without known cardiovascular disease (CVD) who were referred for treadmill exercise testing. We hypothesized that (1) subtle ECG findings as identified by digital electrocardiography portend a worse prognosis, and (2) identification of these findings may improve risk stratification beyond 12 established demographic, clinical, and exercise test risk factors[13] previously shown to predict all-cause mortality.

Methods

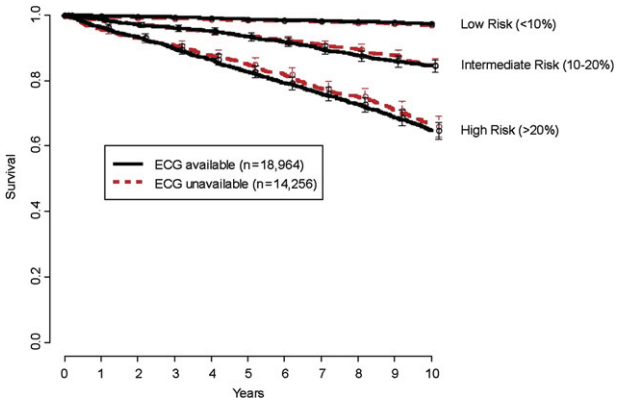
Study population

We performed a single-center (Cleveland Clinic, Cleveland, OH) cohort study. Between September 1990 and December 2002, 46,966 patients without known CVD or end-stage renal disease, 30 years or older, were referred for symptom-limited treadmill exercise stress testing at our institution (Figure 1). None of the patients had known coronary disease (as defined by a history of myocardial infarction, percutaneous coronary intervention, or coronary artery bypass grafting), heart failure, documented left ventricular systolic dysfunction, cardiomyopathy, valvular or congenital disease, previous organ transplantation, atrial fibrillation, digitalis use, pacemaker or defibrillator placement, or end-stage renal disease. All patients had a US Social Security number. Only the first test performed for patients who had >1 stress test was included.

At our institution, board-certified cardiologists and trained exercise physiologists qualitatively review ECGs before the start of stress testing. Based on their interpretation, we excluded 13,746 patients with the following *qualitative* baseline abnormalities: pathological Q waves, bundle branch block, left ventricular hypertrophy (LVH) with ST-segment deviation, preexcitation syndrome, or resting ST-T wave changes unrelated to LVH precluding exercise interpretation (≥1 mm of horizontal or downsloping ST-segment deviation) (Figure 1).

From the remaining 33,220 patients, we excluded 14,256 who did not have a digital 12-lead resting ECG file performed outside the stress laboratory with standard lead placement within 3 months before referral for exercise testing (Figure 1). All ECGs in our hospital are standard 10-second recordings obtained with the Marquette General Electric MUSE system (GE Healthcare, Menomonee Falls, WI). Our ECG laboratory has standardized quality improvement protocols for assuring that lead placement errors and motion artifact are minimized. This

Figure 2



To address possible selection bias, plotted is observed mortality according to availability of digital ECG data and categories of predicted risk by our previously validated prognostic model, Lauer et al,[13] that did not include ECG variables.

resulted in 18,964 patients who had an ECG file that was analyzed digitally. Patients who had digital ECG files available were somewhat more likely to be male and less likely to have hypertension as compared with those excluded; otherwise, there were no major differences (Appendix A). Furthermore, when stratified according to risk groups based on our previously validated prognostic model, there were no differences in mortality among those who did and did not have digital ECG files available (Figure 2).

Digital electrocardiography

Eleven quantitative ECG measures that are commonly considered predictors of poor outcomes were extracted for each patient[2-12] and are believed to be correlates of heart rate, conduction, hypertrophy, or repolarization. These have not been previously scrutinized in patients with clinically normal ECGs undergoing exercise testing.

Digital ECGs were analyzed with General Electric's Magellan Software System (Menomonee Falls, WI), which provided detailed data on the duration and amplitudes of all segments of the P wave, QRS complex, ST segment, and T wave in all 12 leads, with amplitudes recorded to the nearest 100th of a millivolt and times recorded to the nearest millisecond.

In this study, the P wave duration, PR interval, QRS duration, and uncorrected QT interval were the median values from all 12 leads. ST-segment deviation was measured at the end of the segment in lead V5 because this lead has been identified as a superior marker of coronary artery disease.[14] In an analogous manner, we measured T-wave amplitude, and it was the highest value measured in lead V5. Sokolow-Lyon voltage was calculated by adding S wave amplitude in lead V1 to that of the maximum R wave in lead V5 or V6.[15] Cornell voltage was calculated by adding R-wave amplitude in lead AVL to that of the S wave in lead V3.[15,16] ST slope was calculated as the difference between ST-segment deviation at the end of the segment and at the J point in lead V5.[10,14]

Clinical data

Descriptions of prospectively collected demographic and clinical data, as well as details of exercise stress testing protocols in our laboratory, have been published.[13,17-19] Briefly, heart rate recovery was considered to be abnormal if after the first minute of exercise the heart rate fell ≤12 beat/min in patients undergoing an upright cool-down period or ≤18 beat/min in patients undergoing stress echocardiography. Frequent ventricular ectopy during recovery was defined by the presence of ≥7 premature ventricular beat/min, ventricular couplets or triplets, ventricular bigeminy or trigeminy, ventricular tachycardia, ventricular flutter, torsade de pointes, or ventricular fibrillation.[20]

The study was approved by Cleveland Clinic's institutional review board, and because all data were collected and recorded as part of routine clinical care, requirement for informed consent was waived. This project was funded by National Heart, Lung, and Blood Institute CAN# 8324207. The authors are solely responsible for the design and conduct of this study, all study analyses, the drafting and editing of the manuscript and its final contents.

Outcome

All-cause mortality, a clinically relevant and unbiased end point,[21] was ascertained by the Social Security Death Index.[22,23] Patients were followed until a common closing date of January 1, 2008. Previously, we have shown that among our patients, this method has a sensitivity of 97% for accurately detecting death.[19]

Statistical analyses

We constructed Kaplan-Meier plots[24] for each ECG variable and tested mortality differences by using the log-rank χ^2 statistic.

Modeling. We previously published a nonparsimonious prognostic model composed of 12 demographic, clinical, and exercise test variables that performed well in predicting all-cause mortality in patients with a qualitatively normal resting ECG referred for treadmill exercise testing.[13] In an analogous manner, we constructed a nonparsimonious Cox proportional mortality hazards model including all 11 quantitative ECG variables and the 12 traditional risk factor variables listed in Table I (including gender). Nonlinear associations were tested with restricted cubic splines.[25] We tested for interactions. The proportional hazards assumption was tested by scaled Schoenfeld residuals and inspection of hazard ratio (HR) plots.[26]

Constructing a composite ECG score. From this analysis, a composite ECG score was generated: ECG score = (β_1 × ECG measure 1) + (β_2 × ECG measure 2) + … (β_{11} × ECG measure 11), where β_n denoted the confounder-adjusted parameter coefficients. We constructed a second Cox proportional hazards model that included the ECG score along with the 12 traditional risk factors to determine the HR of this score.

Evaluation of prognostic importance. To assess the prognostic contribution of the ECG variables individually and as a cluster (ie, the ECG score), we used the following approaches: calculation of a bootstrapped concordance index (c-index), calculation of prediction error using out-of-bagging,[10] evaluation of reclassification,[27,28] and calculation of integrated discrimination improvement (IDI).[29,30]

Model discrimination was assessed by calculating out-of-bag (OOB) estimates of the c-index for time to event outcomes.[26]

Table I. Baseline, exercise, and electrocardiographic characteristics of 18,964 patients undergoing treadmill exercise testing according to gender

	Women (n = 6486)	Men (n = 12,478)	Entire cohort (n = 18,964)
Demographic and clinical			
Age, y	53 (46, 62)	50 (43, 58)	51 (44, 60)
Current or recent smoking, n (%)	1142 (18)	2109 (17)	3251 (17)
Non–insulin-treated diabetes mellitus, n (%)	372 (6)	549 (4)	921 (5)
Insulin-treated diabetes mellitus, n (%)	153 (2)	194 (2)	347 (2)
Hypertension, n (%)	3000 (46)	5491 (44)	8491 (45)
Typical angina, n (%)	85 (1)	72 (1)	157 (1)
Exercise			
Predicted METS achieved*, %	103 (88, 120)	103 (90, 116)	103 (90, 117)
ST-segment depression (10th, 90th percentiles), mm	0 (0, 1)	0 (0, 1)	0 (0, 1)
Test angina, n (%)	131 (2)	125 (1)	256 (1)
Abnormal heart rate recovery, n (%)	1264 (19)	1930 (15)	3194 (17)
Frequent ventricular ectopy in recovery, n (%)	219 (3)	468 (4)	687 (4)
Electrocardiographic			
Ventricular rate, beat/min	69 (61, 77)	64 (57, 73)	66 (58, 74)
P wave duration, ms	104 (96, 112)	112 (104, 116)	108 (100, 116)
PR interval, ms	156 (140, 168)	160 (148, 176)	160 (144, 176)
QRS duration, ms	84 (80, 92)	96 (88, 100)	92 (84, 100)
Sokolow-Lyon voltage, mV	2.0 (1.6, 2.4)	2.2 (1.8, 2.7)	2.1 (1.8, 2.6)
Cornell voltage, mV	1.1 (0.8, 1.5)	1.4 (1.1, 1.8)	1.3 (1.0, 1.7)
ST-segment deviation in lead V5, mV	0.29 (0.04, 0.53)	0.92 (0.48, 1.41)	0.63 (0.24, 1.17)
ST-segment slope in lead V5, mV	0.29 (0.10, 0.44)	0.78 (0.44, 1.17)	0.54 (0.29, 0.93)
QT interval, ms	396 (372, 416)	400 (376, 420)	396 (376, 420)
QRS axis, °	31 (8, 55)	27 (3, 52)	29 (4, 53)
T-wave amplitude in lead V5, mV	0.22 (0.13, 0.30)	0.31 (0.20, 0.43)	0.28 (0.17, 0.40)

Continuous variables are medians (25th to 75th percentiles), unless otherwise indicated.
*Predicted METS achieved was calculated with previously validated models[17]: for men—observed METS / (18 – [0.15 × age]) × 100; and for women—observed METS / (14.7 – [0.13 × age]) × 100.

This method is similar to calculating the area under an ROC curve, but by bootstrapping there is adjustment for overoptimism.

The OOB method involves obtaining bootstrap samples from the original cohort and using each sample to compute a prediction model that incorporated the ECG variables and potential confounders. Each bootstrap sample left out, on average, approximately 37% of the patients, which was the OOB sample. The prediction model including the ECG variables was applied to the OOB sample to calculate the OOB c-index. The value of this technique is in its ability to effectively perform valid external validations of the model without an unrelated external data set.[10,18]

To determine the change in prediction error (equivalent to change in area under an ROC curve) attributed to each variable (ie, the 'variable importance'), we recalculated the prediction error after random permutation of that variable in the OOB sample (effectively converting the variable's value into completely uninformative noise); an important variable would be expected to yield a greater degradation in the OOB c-index. The process was repeated 100 times for each variable.

To assess calibration, we performed 100 bootstrap resamplings in which patients were divided into quintiles of predicted risk. Within each quintile, actual versus predicted survival rates were calculated, and the differences were averaged to calculate a weighted calibration error. Using this technique, we found the calibration of both Cox models to be excellent.

To address clinical utility, we constructed a risk reclassification plot[27,28,31] comparing the 10-year risk of all-cause mortality predicted by the model with the ECG score and established risk factors against risk predicted by model containing only the established risk factors. We used an objective way of quantifying improvement in classification of risk: IDI.[29,30]

Presentation. Continuous variables are summarized by medians, and categorical variables by frequency and percentage. Kaplan-Meier and risk adjusted plots are presented with error bars or bands indicating 95% confidence intervals (CIs). Life expectancy data from the Human Mortality Project (//www.mortality.org) were used to produce a United States age- and sex-matched reference population mortality curve that was superimposed on a plot of ECG score quartiles.

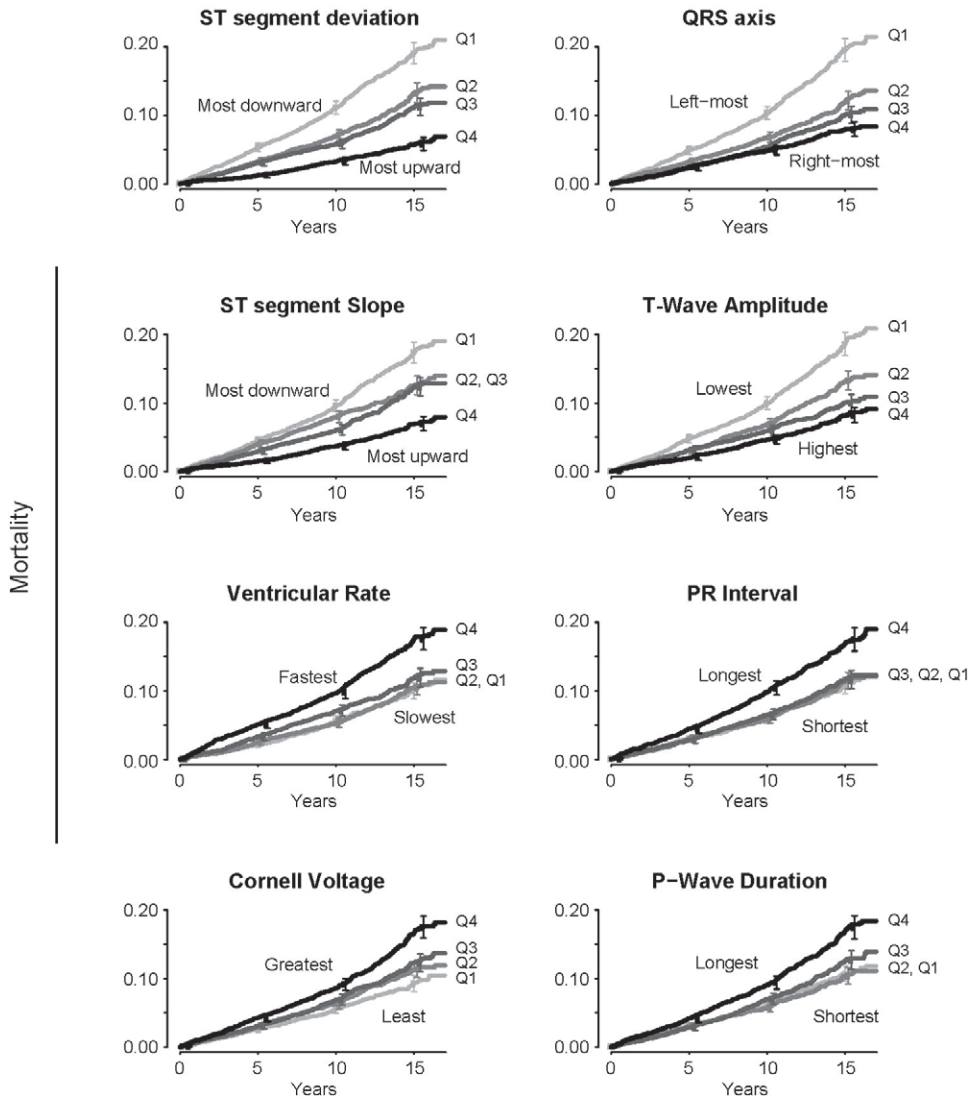
Computational methods. Data assembly was performed with SAS version 9.1.3 (SAS Institute Inc, Cary, NC). Analyses were performed using R version 2.6.2 (www.r-project.org). We used Harrell's Design and Hmisc libraries[26] for model construction and assessment, an R macro written by one of us (H.I.) for the OOB and change in prediction error analyses,[10] another macro written by Kattan[13] for constructing the reclassification table, and the *Relative Survival* library for constructing the population reference plot.[32,33] We used an SAS macro written by Pencina et al[29] for calculating improvement in classification of risk.

Results

Patient characteristics

There were 18,964 patients with both a qualitatively normal ECG and a digital 12-lead resting ECG file available for quantitative analysis; their characteristics are shown in Table I.

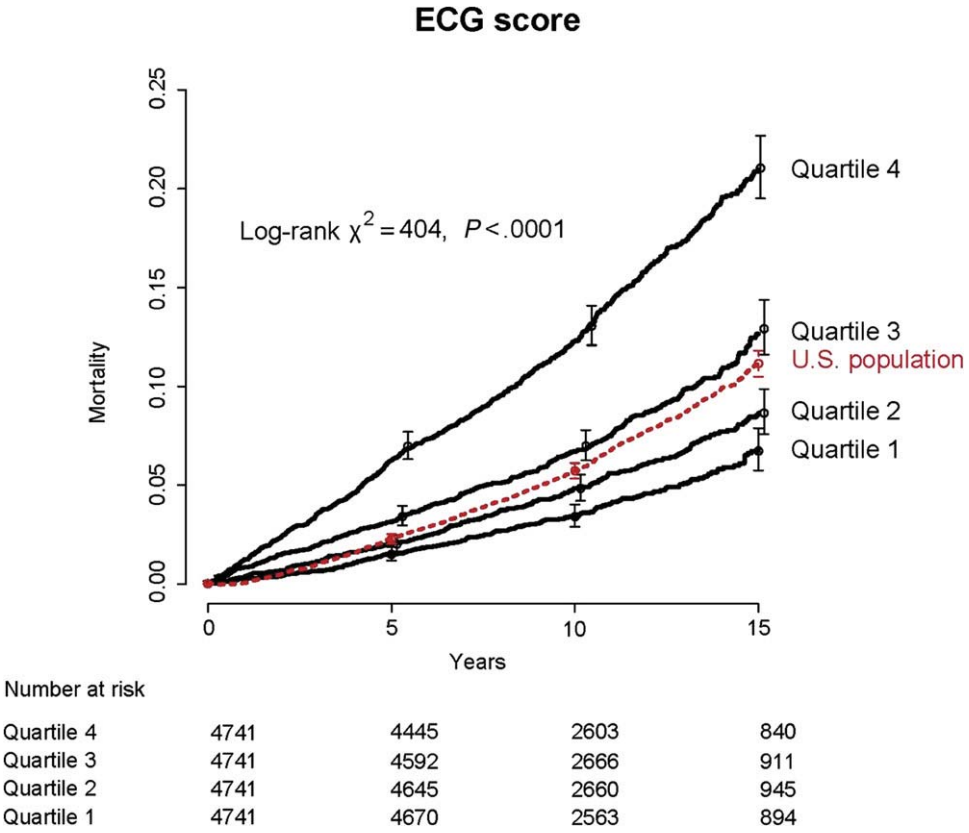
Figure 3



Mortality risk according to quartiles of quantitative ECG measures. Only shown are measures with χ^2 >50. Log-rank *P* value for all <.001. Quartile values: ST-segment deviation (Q1: <0.29 mV, Q2: 0.29-0.68 mV, Q3: 0.68-1.2 mV, Q4: ≥1.2 mV), QRS axis (Q1: <5°, Q2: 5°-30°, Q3: 30°-54°, Q4: ≥54°), ST-segment slope (Q1: <30 mV, Q2: 30-58 mV, Q3: 58-97 mV, Q4: ≥97 mV), T-wave amplitude (Q1: <0.18 mV, Q2: 0.18-0.28 mV, Q3: 0.28-0.40 mV, Q4: ≥0.40 mV), ventricular rate (Q1: <59 beat/min, Q2: 59-67 beat/min, Q3: 67-75 beat/min, Q4: ≥75 beat/min), PR interval (Q1: <148 milliseconds, Q2: 148-164 milliseconds, Q3: 164-180 milliseconds, Q4: ≥180 milliseconds), Cornell voltage (Q1: <0.98 mV, Q2: 0.98-1.31 mV, Q3: 1.31-1.68 mV, Q4: ≥1.68 mV), and P-wave duration (Q1: <104 milliseconds, Q2: 104-112 milliseconds, Q3: 112-120 milliseconds, Q4: ≥120 milliseconds). Error bars indicate 95% CIs.

Among these 18,964 patients, 3,441 (18%) were on aspirin, 1,178 (6%) were on statins, 217 (1%) were on other antihypelipidemic medications, 1,529 (8%) were on thiazide diuretics, 1,968 (10%) were on β-blockers, 1,635 (9%) were on calcium-channel blockers, and 1,447 (8%) were on angiotensin-converting enzyme inhibitors. During exercise testing, 1,917 (10%) developed between 1 and 2 mm of horizontal or downsloping ST-segment depression, and an additional 325 (2%) developed ST-segment depression >2 mm. Digital ECG variables that were correlated with abnormal ST-segment depression during exercise included ventricular rate, P-wave duration and PR interval, Sokolow-Lyon voltage, ST-segment deviation in lead V5, QT interval, and QRS axis (all $P < .0001$).

Figure 4



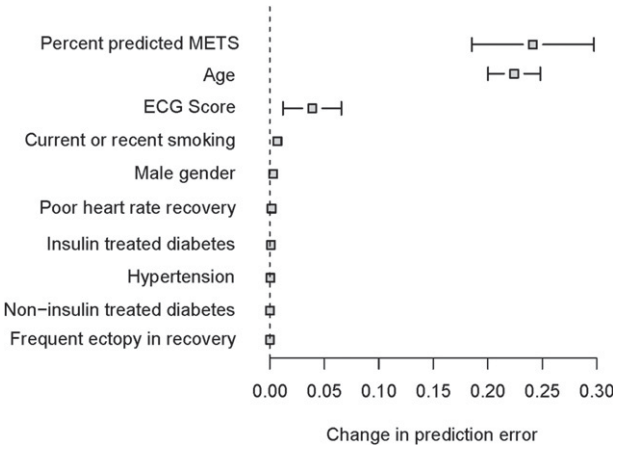
Mortality risk according to quartiles of ECG score (Q1: <22.4, Q2: 22.4-23.4, Q3: 23.4-24.5, Q4: ≥24.5). See Appendix B for ECG score equation. Error bars indicate 95% CIs. Dashed line is predicted mortality based on US vital statistics and the age and sex distribution of our cohort.

ECG findings and mortality

During a median follow-up of 10.7 years (range for survivors 5-17 years), 1,585 patients (8%) died. The following quantitative ECG findings were most strongly associated (log-rank χ^2 >100) with increased mortality in an unadjusted analysis: more downward ST-segment deviation, more leftward QRS axis, more downward ST slope, lower T-wave amplitude, faster ventricular rate, longer PR interval, greater Cornell voltage, and longer P-wave duration were all associated with higher mortality (Figure 3).

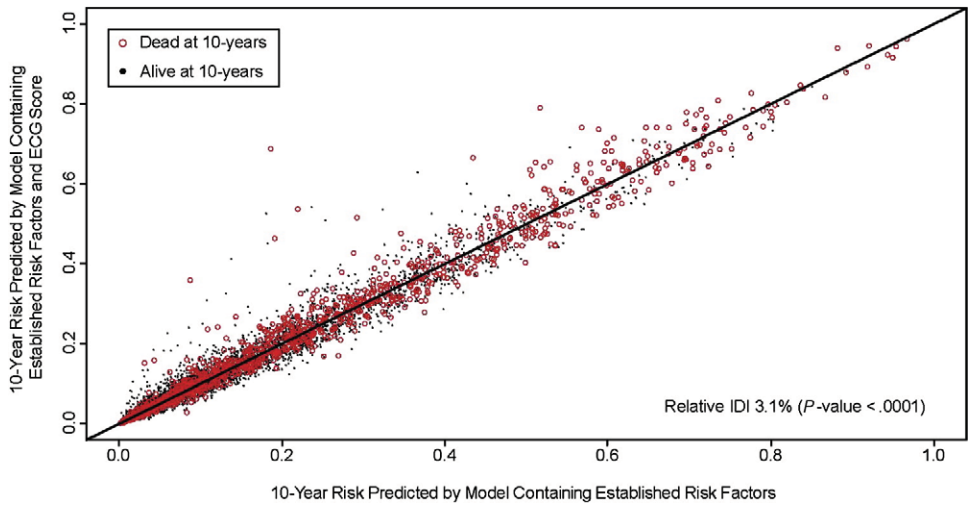
The OOB c-index of the Cox model containing the 12 established risk factors[13] and all 11 quantitative ECG measures was 0.84. Among the ECG variables, those that contributed most to improved prediction were the following: higher ventricular rate (decrease in OOB c-index = 0.004 [0.5% decrease]), more leftward QRS axis (0.003 [0.4%]), more downward ST-segment deviation (0.003 [0.4%]), longer QT interval (0.002 [0.2%]), more downward ST-segment slope (0.0007 [0.08%]), lower T-wave amplitude (0.0006 [0.07%]), greater Cornell voltage (0.0003 [0.04%]), longer PR interval (0.0003 [0.04%]),

Figure 5



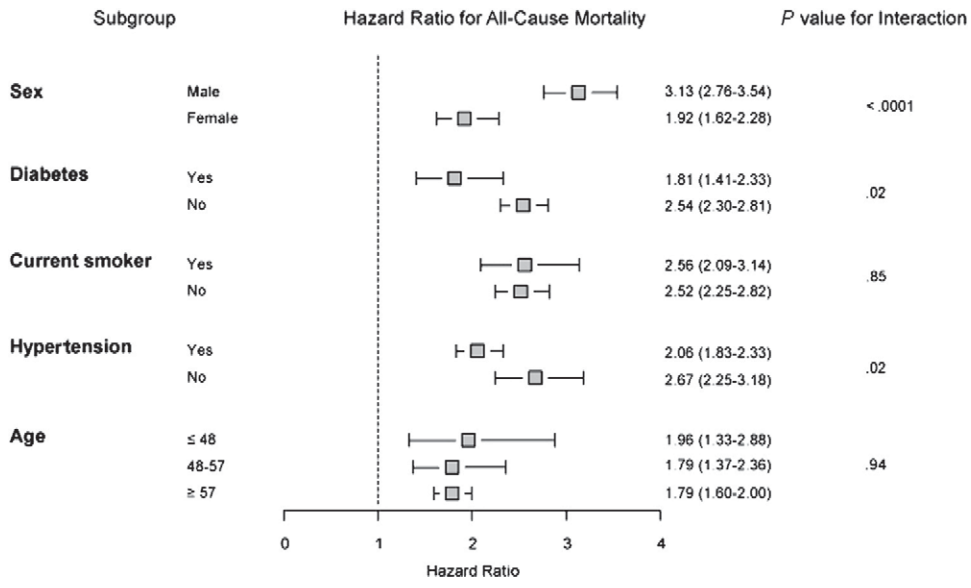
Change in prediction error by out-of-bagging. Only variables that contribute to predictive discrimination are shown. Results are based on 100 bootstrap samples. Error bars indicate 95% CIs.

Figure 6



Ten-year risk of all-cause mortality predicted by the model containing the ECG score and established risk factors against risk predicted by model containing only the established risk factors. If a variable adds no predictive value to the model, all points fall on the line of identity. Spread around the line indicates modulation of predicted risk; if the variable correctly modulates predicted risk, there should be a greater preponderance of events (red circles) above the line of identity. Improvement of classification is expressed as relative IDI, demonstrating as a 3% improvement in model performance with the ECG score.

Figure 7



Subset analysis comparing highest ECG score quartile to bottom quartiles.

longer P-wave duration (0.0001 [0.01%]), and longer QRS duration (0.0001 [0.01%]).

ECG score and mortality

The adjusted ECG score (Appendix B), a nonparsimonious cluster of the ECG variables, was predictive of death (Figure 4) even after multivariable adjustment (75th vs 25th percentile, adjusted HR 1.36, 95% CI 1.25-1.49, $P <$ .0001). There was a significant interplay between age and the adjusted ECG score (P for interaction = .005), whereby the younger the patient, the higher the predicted utility of the adjusted ECG score.

The OOB c-index of the model with the adjusted ECG score was 0.84. Adjusted ECG score was the third most important contributor to accuracy of predicting mortality (decrease in OOB c-index = 0.04 [4.8% decrease]) after percent predicted metabolic equivalents (METS; 0.24 [28.6%]) and age (0.22 [26.2%]) (Figure 5).

To assess whether the association of adjusted ECG score with mortality differed over time, we performed a supplementary analysis in 2 subsets of the cohort. In 6,797 patients who had the exercise test between 1990 and 1995 and censored at 5 years, adjusted ECG score was at least as predictive of death (75th vs 25th percentile, adjusted HR 1.38, 95% CI 1.22-1.57, $P < .0001$), as in 7,558 patients who had the test between 1995 and 2000 and censored at 5 years (75th vs 25th percentile, adjusted HR 1.58, 95% CI 1.32-1.88, $P < .0001$).

Figure 6 shows 10-year risk of mortality predicted by the model containing the ECG score and established risk factors against risk predicted by the model containing only the established risk factors. Each point represents an individual patient. The IDI, a measure of improvement in model performance, was 0.006 (95% CI 0.003-0.008, $P < .0001$). This corresponds to a relative IDI of 3%, meaning there was a 3% improvement in model performance due to the addition of ECG score. Despite this value being statistically significant, its absolute value is small and suggests only limited utility of adding ECG score to the prognostic model. The modest effect of ECG score is also reflected by the minimal spread observed around the 45 degree line of identity.

Subset analysis

To assess how ECG score performed in different subsets of the cohort, we compared mortality risk in the highest ECG score quartile versus the lower 3 quartiles in subsets of sex, diabetes, smoking, history of hypertension, and age. The ECG scores within the highest quartile portended increased risk in all subsets of patients (Figure 7), with a significantly higher utility in males and those without history of diabetes or hypertension.

Discussion

In this study of 18,964 patients without known CVD who were referred for exercise testing and who had a clinically normal resting ECG, we found that a number of digitally measured ECG findings reflective of heart rate, conduction, left ventricular mass, and repolarization were predictive of long-term mortality individually and as a composite. However, the increased risk associated with a composite ECG score led to a small change in the OOB c-index (4.8%), which suggested that adding information from quantitative ECG measures to established risk factors only modestly improved discrimination for prediction of 10-year mortality. Further, a prediction model with ECG score yielded only modest reclassification of risk (3% increase in model performance) beyond the standard prediction model, which was based on easily obtained demographic, clinical, and exercise test measures.

Our findings are consistent with prior studies demonstrating the predictive utility of subtle ECG measures for long-term mortality.[4,7,11,34-40] Denes et al[11] recently demonstrated that "minor" ECG abnormalities (defined as first- or second-degree atrioventricular block, borderline prolonged ventricular excitation, prolonged ventricular repolarization, isolated minor Q and ST-T abnormalities, LVH without ST-T abnormalities, left atrial enlargement, frequent atrial or ventricular premature beats, or fascicular blocks) added incremental prognostic information to established risk factors in asymptomatic postmenopausal women. Consistent with this, we found that in a relatively healthy stress testing population with clinically normal ECGs, quantitative ECG measures added incremental prognostic value beyond established risk factors, but the increment of improvement was modest. This may be because unlike previous investigations, we accounted for systematic measures of exercise capacity, which have consistently been shown to be one of the most powerful predictors of mortality.

Based on our analysis, health care providers utilizing exercise testing for assessment of prognosis will gain little from incorporating ECG measures into their current risk assessment algorithm. It remains unknown though if ECG measures can have greater clinical utility in non–stress-testing populations, where the powerful prognostic value of functional capacity is not assessed objectively.

Limitations

Our study population was derived from Cleveland Clinic, a single tertiary referral center, and as such, our results can only be generalized to exercise test patients in similar settings. We did not have an external data set on which to validate our findings. Because we are not aware of any other cohorts who have had these types of detailed clinical, stress testing, and quantitative ECG information collected, we opted instead to use out-of-bagging for validation.[10] With this internal validation approach, our prediction estimates were not based on one data set alone but on 100 bootstrapped segments of the larger cohort.

Potential for selection bias exists. We only studied patients who had resting ECGs collected with standard lead placement, in addition to pretest resting ECGs with modified torso leads. Some physicians did not refer patients for standard 12-lead ECGs if they were available from prior visits or from outside institutions, or if they felt that the stress laboratory–modified lead ECGs were adequate. Other reasons for missing digital ECGs included those patients who had a resting ECG done at a satellite of the hospital where central storage is not available and a small fraction whose digital file was corrupted. To assess for potential selection bias, we examined the excluded

patient's baseline characteristics and their outcomes. There were few major differences between included and excluded patients, particularly for the 2 major clinical predictors of risk, namely, age and predicted METS achieved (Appendix A). Further, the mortality rate of the excluded subset closely mirrored that of the studied subset (Figure 2).

The ECG measurements were performed digitally from surface ECGs. Intracardiac measurements done in the electrophysiology laboratory might provide even more precise measures.

Summary and conclusions

We found that the resting ECG done before exercise testing may provide incremental prognostic value beyond determining validity of ST-segment interpretation. Subtle ECG findings relating to heart rate, conduction, left ventricular mass, and repolarization portend a worse long-term prognosis but only modestly improve discrimination and clinical risk stratification for all-cause mortality.

References

1. Klocke FJ, Baird MG, Lorell BH, et al. ACC/AHA/ASNC guidelines for the clinical use of cardiac radionuclide imaging—executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (ACC/AHA/ASNC Committee to Revise the 1995 Guidelines for the Clinical Use of Cardiac Radionuclide Imaging). Circulation 2003;108:1404-18.
2. Kannel WB, Anderson K, McGee DL, et al. Nonspecific electrocardiographic abnormality as a predictor of coronary heart disease: the Framingham Study. Am Heart J 1987;113(2 Pt 1):370-6.
3. Levy D, Salomon M, D'Agostino RB, et al. Prognostic implications of baseline electrocardiographic features and their serial changes in subjects with left ventricular hypertrophy. Circulation 1994;90: 1786-93.
4. Okin PM, Devereux RB, Kors JA, et al. Computerized ST depression analysis improves prediction of all-cause and cardiovascular mortality: the Strong Heart Study. Ann Noninvasive Electrocardiol 2001;6: 107-16.
5. Sigurdsson E, Sigfusson N, Sigvaldason H, et al. Silent ST-T changes in an epidemiologic cohort study—a marker of hypertension or coronary heart disease, or both: the Reykjavik study. J Am Coll Cardiol 1996; 27:1140-7.
6. Knutsen R, Knutsen SF, Curb JD, et al. The predictive value of resting electrocardiograms for 12-year incidence of coronary heart disease in the Honolulu Heart Program. J Clin Epidemiol 1988;41:293-302.
7. Okin PM, Devereux RB, Nieminen MS, et al. Electrocardiographic strain pattern and prediction of cardiovascular morbidity and mortality in hypertensive patients. Hypertension 2004;44:48-54.
8. Iuliano S, Fisher SG, Karasik PE, et al. QRS duration and mortality in patients with congestive heart failure. Am Heart J 2002;143:1085-91.
9. Haissaguerre M, Derval N, Sacher F, et al. Sudden cardiac arrest associated with early repolarization. N Engl J Med 2008;358: 2016-23.
10. Lauer MS, Martino D, Ishwaran H, et al. Quantitative measures of electrocardiographic left ventricular mass, conduction, and repolarization, and long-term survival after coronary artery bypass grafting. Circulation 2007;116:888-93.
11. Denes P, Larson JC, Lloyd-Jones DM, et al. Major and minor ECG abnormalities in asymptomatic women and risk of cardiovascular events and mortality. JAMA 2007;297:978-85.
12. Prineas RJ, Grandits G, Rautaharju PM, et al. Long-term prognostic significance of isolated minor electrocardiographic T-wave abnormalities in middle-aged men free of clinical cardiovascular disease (The Multiple Risk Factor Intervention Trial [MRFIT]). Am J Cardiol 2002;90: 1391-5.
13. Lauer MS, Pothier CE, Magid DJ, et al. An externally validated model for predicting long-term survival after exercise treadmill testing in patients with suspected coronary artery disease and a normal electrocardiogram. Ann Intern Med 2007;147:821-8.
14. Miranda CP, Liu J, Kadar A, et al. Usefulness of exercise-induced ST-segment depression in the inferior leads during exercise testing as a marker for coronary artery disease. Am J Cardiol 1992;69:303-7.
15. Sundstrom J, Lind L, Arnlov J, et al. Echocardiographic and electrocardiographic diagnoses of left ventricular hypertrophy predict mortality independently of each other in a population of elderly men. Circulation 2001;103:2346-51.
16. Casale PN, Devereux RB, Alonso DR, et al. Improved sex-specific criteria of left ventricular hypertrophy for clinical and computer interpretation of electrocardiograms: validation with autopsy findings. Circulation 1987;75:565-72.
17. Cole CR, Blackstone EH, Pashkow FJ, et al. Heart-rate recovery immediately after exercise as a predictor of mortality. N Engl J Med 1999;341:1351-7.
18. Kim ES, Ishwaran H, Blackstone E, et al. External prognostic validations and comparisons of age- and gender-adjusted exercise capacity predictions. J Am Coll Cardiol 2007;50:1867-75.
19. Nishime EO, Cole CR, Blackstone EH, et al. Heart rate recovery and treadmill exercise score as predictors of mortality in patients referred for exercise ECG. JAMA 2000;284:1392-8.
20. Frolkis JP, Pothier CE, Blackstone EH, et al. Frequent ventricular ectopy after exercise as a predictor of death. N Engl J Med 2003;348: 781-90.
21. Lauer MS, Blackstone EH, Young JB, et al. Cause of death in clinical research: time for a reassessment? J Am Coll Cardiol 1999;34: 618-20.
22. Newman TB, Brown AN. Use of commercial record linkage software and vital statistics to identify patient deaths. J Am Med Inform Assoc 1997;4:233-7.
23. Curb JD, Ford CE, Pressel S, et al. Ascertainment of vital status through the National Death Index and the Social Security Administration. Am J Epidemiol 1985;121:754-66.
24. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assn 1958;53:457-81.
25. Harrell Jr FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996;15:361-87.
26. Harrell Jr FE. Regression modeling strategies with applications to linear models, logistic regression, and survival analysis. New York (NY): Springer-Verlag; 2001.
27. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation 2007;115:928-35.
28. McGeechan K, Macaskill P, Irwig L, et al. Assessing new biomarkers and predictive models for use in clinical practice: a clinician's guide. Arch Intern Med 2008;168:2304-10.
29. Pencina MJ, D'Agostino Sr RB, D'Agostino Jr RB, et al. Evaluating the added predictive ability of a new marker: from area under the ROC

curve to reclassification and beyond. Stat Med 2008;27:157-72 [discussion 207-12].
30. Pencina MJ, D'Agostino Sr RB, D'Agostino Jr RB, et al. Rejoinder: comments on 'integrated discrimination and net reclassification improvements—practical advice'. Stat Med 2008;27:207-12.
31. Cook NR, Buring JE, Ridker PM. The effect of including C-reactive protein in cardiovascular risk prediction models for women. Ann Intern Med 2006;145:21-9.
32. Pohar M, Stare J. Relative survival analysis in R. Comput Methods Programs Biomed 2006;81:272-8.
33. Therneau TM, Offord J. Expected survival based on hazard rates. Feb 1999. Available at http://mayoresearch.mayo.edu/mayo/research/biostat/upload/63.pdf. Last accessed July 3, 2008.
34. Cullen K, Stenhouse NS, Wearne KL, et al. Electrocardiograms and 13 year cardiovascular mortality in Busselton study. Br Heart J 1982;47:209-12.
35. De Bacquer D, De Backer G, Kornitzer M, et al. Prognostic value of ECG findings for total, cardiovascular disease, and coronary heart disease death in men and women. Heart 1998;80:570-7.
36. Cedres BL, Liu K, Stamler J, et al. Independent contribution of electrocardiographic abnormalities to risk of death from coronary heart disease, cardiovascular diseases and all causes. Findings of three Chicago epidemiologic studies. Circulation 1982;65:146-53.
37. Greenland P, Xie X, Liu K, et al. Impact of minor electrocardiographic ST-segment and/or T-wave abnormalities on cardiovascular mortality during long-term follow-up. Am J Cardiol 2003;91:1068-74.
38. Okin PM, Devereux RB, Lee ET, et al. Electrocardiographic repolarization complexity and abnormality predict all-cause and cardiovascular mortality in diabetes: the Strong Heart Study. Diabetes 2004;53:434-40.
39. Okin PM, Malik M, Hnatkova K, et al. Repolarization abnormality for prediction of all-cause and cardiovascular mortality in American Indians: the Strong Heart Study. J Cardiovasc Electrophysiol 2005;16:945-51.
40. Oikarinen L, Nieminen MS, Viitasalo M, et al. QRS duration and QT interval predict mortality in hypertensive patients with left ventricular hypertrophy: the Losartan Intervention for Endpoint Reduction in Hypertension Study. Hypertension 2004;43:1029-34.

Appendix A

Table Comparison of patients who were ("with resting ECGs") and were not ("without resting ECGs") included in the study

Characteristic	Value			
	With resting ECGs (n = 18,964)		**Without resting ECGs (n = 14,256)**	
Demographic and clinical				
Age, y	52	(44, 60)	54	(45, 62)
Male, n (%)	12,478	(66)	8004	(56)
Current or recent smoking, n (%)	3251	(17)	2426	(17)
Non–insulin-treated diabetes mellitus, n (%)	921	(5)	899	(6)
Insulin-treated diabetes mellitus, n (%)	347	(2)	318	(2)
Hypertension, n (%)	8491	(45)	7659	(54)
Typical angina, n (%)	157	(1)	201	(1)
Exercise				
Predicted METS achieved*, %	103	(90, 117)	102	(88, 116)
ST-segment depression (10th, 90th percentiles), mm	0	(0, 1)	0	(0, 1)
Test angina, n (%)	256	(1)	230	(2)
Abnormal heart rate recovery, n (%)	3194	(17)	2441	(17)
Frequent ventricular ectopy in recovery, n (%)	687	(4)	515	(4)

Continues variables are medians (25th to 75th percentiles), unless otherwise indicated.

*Predicted METS achieved was calculated with previously validated models[17]: for men—observed METS / (18 – [0.15 × age]) × 100; and for women—observed METS / (14.7 – [0.13 × age]) × 100. See text for details.

Appendix B. Formula

The equation of the ECG score: ECG Score=0.012815493×ventricular rate−0.0015165242×P wave duration+0.0016282304×PR interval−0.0033127115×QRS duration+4.2726708e−05×Sokolow−Lyon voltage+8.121262e−05×Cornell voltage−0.0019176579×ST segment deviation +0.00077421327×ST segment slope+0.0036478234×QT interval−0.0043564716×QRS axis+8.37693e−07×pmax (QRS axis+15,0)3−1.7591553e−06×pmax(QRS axis−29,0)3+9.214623e−07×pmax(QRS axis−69,0)3 +0.00038538948×T wave amplitude

**pmax: Pick maximum of the 2 possible values. For example, pmax(QRS axis + 15,0) means that if (QRS axis + 15) is >0 that value is chosen; if (QRS axis + 15) is <0, then the value zero is chosen.