

Risk Assessment of Sudden Cardiac Arrest using Routinely Available Clinical Variables in a Multivariate Model

Technical Report

April 16, 2012

Sanjay Mehrotra¹

Kibaek Kim¹

David Liebovitz²

Jeffrey J. Goldberger³

¹Industrial Engineering and Management Sciences, Northwestern University, Evanston, Illinois

²Division of General Internal Medicine, Feinberg School of Medicine of Northwestern University, Chicago, Illinois

³Division of Cardiology, Feinberg School of Medicine of Northwestern University, Chicago, Illinois

Abstract

Context

It is known that the sudden cardiac risk stratification achieved by using LVEF to give a dichotomous criterion is sub-optimal, leading to unacceptable false positive and false negative clinical diagnosis. Population based multivariate risk assessment models have been developed for patients at risk of developing cardiovascular disease. However, no such risk assessment model is known for patients at risk of SCA.

Objective

To develop a multivariate risk prediction model for patients at risk of SCA using easily available clinical data, and evaluate its use as a screening tool.

Data Source

Patient medical data were obtained from Northwestern Medical Enterprise Data Warehouse at Northwestern Memorial Hospital and Northwestern University for the study period from January 2006 to December 2010.

Study Selection

Patients age < 30 and patients who were pregnant or received cancer diagnosis (malignant neoplasm, sarcoma, tumors, or carcinoma) during the study period were excluded.

Results

The model was first developed and tested on a cohort of 23,041 patients with 73 SCA events. The area under the receiver operating characteristics curve (AUC) for the developed model using six basic clinical variables (Age, BMI, CHF, Diastolic BP, MI, and Ventricular Rate), and two ECG variables (QTc and P-axis) is 0.85. This was further validated using bootstrapping on a cohort of 69,670 patients with 189 SCA events given mean AUC = 0.85 (95% CI 0.82 - 0.88).

Conclusions

A population based multivariate logistic regression model using data variables that are available through routine and inexpensive clinical tests is developed for SCA risk assessment. This model has high predictive value. Such a model should be used as a baseline model before evaluating the predictive value of additional variables obtained from advanced clinical tests and methods.

Sudden cardiac arrest (SCA) is a leading cause of death (SCD) in the U.S. and worldwide. U.S. estimates of deaths due to SCA vary between 300,000 and 460,000¹. Invasive and noninvasive medical examinations have been developed over the years to identify the patients at risk of SCA¹. Left ventricular ejection fraction (LVEF) and variables derived from the ECG provide noninvasive markers to assess a patient's risk of SCA²⁻⁴. Currently, dichotomous criteria are used to identify patients with prior myocardial infarction (MI) or congestive heart failure (CHF) as high risk for SCA, which may warrant insertion of an implantable cardioverter defibrillator (ICD)^{2,5-8}. The current ACC/AHA/ESC 2006 guidelines for management of patients with ventricular arrhythmias and the prevention of SCD recommend a LVEF of less than 35 - 40% as a critical point to consider ICD implantation^{8,9}.

Studies have shown that LVEF alone is an inadequate predictor for SCA events with respect to both true positive rate (TPR)^{10,11} and false positive rate (FPR)^{2,7,12-14}. For example, TPR was 30% for patients with LVEF < 35% in the Oregon Sudden Unexpected Death Study¹⁰. It has been recognized that improved risk assessment represents an important challenge towards reducing SCA mortality and improving the cost-benefit ratio of ICD implantation^{1,15}. The etiology of SCA is considered multivariate, and the limitation of the use of LVEF as a risk criterion has been discussed^{12,15}. Factors such as age, prior MI, presence of CHF, hypertension, left ventricular hypertrophy, intraventricular conduction block, elevated serum cholesterol, glucose intolerance, decreased vital capacity, smoking, weight, and heart rate have been identified as risk factors for SCA¹⁶⁻¹⁸.

While methods such as the Framingham risk score¹⁹ are known for cardiovascular disease risk assessment, risk assessment for SCA is unknown. There is a great need for population based SCA risk assessment to facilitate screening of patients at risk^{20,21}. The goal of this study is to

develop a risk assessment model for SCA for possible applicability in screening of large populations.

STUDY DESIGN AND MEHODS

Data.

The study data were collected from the Northwestern Medical Enterprise Data Warehouse (EDW) and its feeder databases. The data in the EDW is stored from different patient health record systems currently being used at the Northwestern Memorial Hospital and Northwestern Memorial Faculty Foundation²². The in and out-of-hospital SCA events were not distinguished. The study was approved by Northwestern University Institutional Review Board. The study included all records in EDW from January 2006 to December 2010 with the following exclusion criteria: age < 30 years; pregnant women during the study period; and patients with cancer. The data further excluded patients who have QT interval greater than 1000, PR interval greater than 600, blood pressure greater than 400, or triglyceride greater than 10000. SCA was identified by International Classification of Diseases, Ninth Revision (ICD-9) codes: 427.41, Ventricular fibrillation; 427.42, Ventricular flutter; and 427.5, Cardiac arrest. Charts for patients with these ICD-9 codes were manually reviewed for correctness.

The clinical data used for this study ranged from 180 days prior to until 14 days prior to the SCA event. For patients with multiple SCA event records, the first SCA event was used. For patients with no SCA records, all data from the study duration was used. When multiple records of a clinical variable were available in the study time window, the following four derived data were used: the average, the minimum, the maximum, and the most recent. The value of these derived variables was set to the same value if only one record was found.

Data on the following four groups of variables were collected: (i) clinical variables (age, gender, race, body mass index [BMI], systolic and diastolic blood pressure [BP], history of congestive heart failure [CHF] history of myocardial infarction [MI], statin therapy); (ii) LVEF; (iii) ECG variables (atrial rate, ventricular rate, PR interval, QRS duration, QT interval, QTc, P-axis, R-axis, T-axis); and (iv) lipid variables (high density lipoprotein [HDL], low density lipoprotein [LDL], and triglycerides). MI was identified by the following ICD-9 codes: 410, 410.6, 410.7, 410.8, 410.9, 411.81, 412, 414.2, 429.7, and V12.5. CHF was identified by the following ICD-9 codes: 398.91, 402, 402.00, 404, 428, 428.0, 428.1, 428.2, 428.3, and 428.9. Categorical variables were used to indicate patient gender, race, a diagnosis of CHF, prior MI, and the use of statin therapy. All other data were represented by continuous variables.

Variable Selection, Model Calibration, and Validation.

Logistic regression analysis was used to construct the prediction function²³. The forward stepwise selection method was used on the study data to select the variables to be used for all subsequent analyses²³. Univariate logistic regression analysis first generated a pool of candidate variables. A variable was put in the pool if its likelihood ratio statistic (p-value) was less than 0.25. After generating the variable pool, the forward stepwise selection method started with an empty set, and sequentially considered a variable for addition in the logistic model being trained. If a variable was added, the method also considered a variable currently in the model for removal. The variable with the smallest p-value from the remaining pool of variables was considered for addition. A variable was added if the likelihood ratio of the model prior to adding this variable was improved with p-value < 0.15. If no such variable was found, the process terminated. The stepwise selection method next considered removing a variable from the model. A variable was removed if the likelihood ratio of the model did not reduce

significantly (p-value < 0.10) by removing this variable. If more than one derived variable among Average, Min, Max, Recent derived variables remained in the model at the end of the forward stepwise selection method, only the variable giving the largest likelihood ratio was kept in the model.

The area under the receiver operating characteristic (ROC) curve (AUC) was used to assess the power of the prediction function¹². AUC was calculated using all patient records for training and subsequent testing of the model. The 2-fold cross validation (2xV) method²⁴ was used as a bootstrapping method to generate an estimate of the mean and confidence interval for the AUC. The 2xV method used a random selection of 50% of patients to form a dataset to train the model. It then validates the predictive power of the trained model on the remaining 50% patient data. The confidence interval for AUC was generated from 50 replications of the 2xV method. All analyses, including the stepwise selection method, the calculation of AUC, the construction of the ROC curve, and the 2xV method, were performed using STATA²⁵.

RESULTS

Patient Characteristics

The patient characteristics of the cohort are given in Table 1. The group consists of 23,041 patients, of which only 73 patients (0.32%) had a SCA recorded in EDW. Older patients were more likely to have SCA (p-value < .001). In this study population the SCA rate is not significantly different across gender (p-value = .32) and race (p-value = .66). The BMI, diastolic and systolic blood pressure, atrial rate, ventricular rate, QRS duration, QT interval, QTc, R-axis, T-axis, LVEF value, LDL, and HDL of patients with SCA are significantly different from those for whom no SCA is recorded (see Table 1). Prior MI and CHF are more

prevalent in SCA patients (p-values = .09 and < .001, respectively). Patients with prior MI and/or CHF had relative risks of 6.31 and 8.45, respectively. Differences in PR interval, and triglyceride variables were not statistically significant.

Patient SCA Risk Prediction Function

The univariate analysis excluded the following variables from further consideration in the forward stepwise selection method: gender, ethnicity, atrial rate (minimum), PR-interval (average, minimum, recent), QT-interval (recent), P-axis (maximum, recent), R-axis (maximum), triglyceride (average, minimum, maximum, recent), systolic BP (maximum), and BMI (minimum, recent).

The forward stepwise selection method sequentially added variables in the following order: CHF (p-value < .001), ventricular rate (maximum) (p-value < .001), diastolic BP (recent) (p-value < .001), LVEF (average) (p-value < .001), LDL (maximum) (p-value = .002), LDL (average) (p-value = .006), MI (p-value = .01), QTc (minimum) (p-value = .02), P-axis (average) (p-value = .05), QTc (recent) (p-value = .05), T-axis (minimum) (p-value = .05), T-axis (average) (p-value = .05), BMI (average) (p-value = .05), age (p-value = .05), and diastolic BP (average) (p-value = .07). Two derived variables for LDL (maximum, average), diastolic BP (recent, average), QTc (minimum, recent), and T-axis (minimum, average) remained in the model at the termination of the forward stepwise selection method. Of these, LDL (maximum), diastolic BP (recent), QTc (minimum), and T-axis (minimum) were removed from further consideration. Table 2 presents the coefficients of the ten variables in the final multivariate logistic regression model and their corresponding odds ratios (OR).

The ROC, TPR, and TNR curves are given in Figures 1a and 1b. The AUC for this model is 0.87. At the TPR of 30% (a number in the Oregon Sudden Unexpected Death Study¹⁰), the FPR was 2.4% (equivalently, 97.6% specificity). Figure 1c presents the decile of mean predicted event rate and the observed event rate.

Model Validation using Time Sensitivity Analysis

The main analysis incorporated data from 180 days prior to 14 days prior to SCA. Two time sensitivity analysis were performed without retraining the model. In the first analysis, none of the data immediately prior to the first SCA event was excluded (180 days prior up to the time of SCA). In the second analysis data up to 30-days prior to SCA was excluded (180 days to 30 days prior to SCA). The AUCs for these cohorts are 0.85, and 0.86, respectively.

Variable Sensitivity Analysis

A further analysis was performed to evaluate the contribution of specific subsets of variables to the prediction model. These models are called sub-models below. The sub-models were retrained using the derived variables (i.e., minimum, maximum, average, recent) of the variable subsets defining the sub-model.

The results from this analysis are given in Table 3. The ECG sub-model has clinical and ECG variables. The No LVEF sub-model has all variables except LVEF variable. The No MI-CHF sub-model has all variables except prior CHF and MI variables. Two additional sub-models: one using age and LVEF (Age+LVEF) and one using LVEF only, were also formed in the variable sensitivity analysis. Finally, the Recent sub-model considers only the most recent values of the variables when training the model.

The AUCs of ECG group and No LVEF group were not significantly different from the Base cohort (p-values = .30 and .90, respectively). The lack of difference in AUCs among Base, No LVEF and ECG groups suggests that the additional predictive value of LVEF and lipid variables is already captured in other variables. The model without MI or CHF knowledge has a reduced AUC value both in the standard analysis (AUC = 0.85) and the 2xV analysis (mean AUC = 0.82, p-value < .001). Note that QT interval and QRS duration were not chosen in the Base group variable selection method. In the variable selection method the Recent sub-model identified the following seven variables: indication of prior MI, indication of prior CHF, diastolic BP, ventricular rate, LVEF, LDL, and age. The corresponding sub-model AUC value is 0.83, and the mean 2xV AUC value is 0.81 (95% CI, 0.76 – 0.86). The Recent sub-model AUC was 0.82 when LVEF was dropped from the set of seven variables in this model. The LVEF sub-model has AUC = 0.64 with mean 2xV AUC = 0.64 (95% CI, 0.57 – 0.72). The Age+LVEF sub-model has AUC = 0.73, with mean 2xV AUC = 0.73 (95% CI, 0.65 – 0.78). At 33% TPR, the FPR of different sub-models are also reported in Table 3. The FPR in the multivariate model is more than ten fold better than the FPR in the Oregon Sudden Unexpected Death Study¹⁰, where FPR=30% was reported. The FPR of the base model is significantly better than the FPR of LVEF or Age+LVEF sub-models.

Validation of the ECG Model on a Larger Cohort

A larger cohort was generated by only using the clinical variables and the ECG variables. This larger cohort had 69,670 patients with 189 SCA events. The patient characteristics of this cohort are given in Table 4. The ECG sub-model trained from the cohort in Table 2 was used for validation on the cohort in Table 4. The AUC = 0.85 and 2xV mean AUC = 0.85 (95% CI,

0.82 - 0.88) was found with no statistical difference (p-value = .156) with the 2xV AUC resulting from the original cohort (Table 1) on which the model was trained.

DISCUSSION

This study developed an effective risk stratification technique to identify particularly high risk individuals for SCA among a broad population of patients. The model was developed in a very low risk population with a prevalence of SCA of 0.32% and validated in a larger population with a SCA prevalence of SCA of 0.27%. Using only clinical (Age, BMI, CHF, Diastolic BP, MI, and Ventricular Rate) and ECG based variables (QTc and P-axis) is AUC 0.85 (95% CI 0.82 - 0.88). The model had excellent discrimination with an AUC of 0.85 and excellent calibration. These data provide a robust, semi-automatic clinical algorithm to be used in the initial screening of large populations at risk of SCA.

There are several notable features of this study and the approach taken. First, by the nature of the selection process, the endpoint – SCA – selected individuals who were survivors of SCA, as they needed to arrive for medical attention and have SCA coded in their visit. This is precisely the population that is targeted for risk stratification for prevention of sudden cardiac death – those with reversible causes. In contrast, it is well known that the adjudicated diagnosis of sudden cardiac death in clinical trials has many etiologies that are all not amenable to defibrillation, such as ruptured aortic aneurysm, pulmonary embolism, and myocardial infarction/rupture, among others^{26,27}.

Various approaches have been taken to risk stratification for the large public health problem of sudden cardiac death. As initially proposed by Myerburg²⁸, there is a dissociation between the incidence and prevalence of sudden death, so that the highest risk groups account for only a

minority of actual sudden death cases. It is important to note that risk stratification is necessary and important both in the high risk population (i.e. those with severe left ventricular dysfunction) and the low risk population. A variety of invasive and noninvasive testing approaches have been tested in the highest risk groups²⁹. It is notable that even within this group, several studies have highlighted the importance of clinical markers^{12,13,30}. As the majority of sudden death cases occur in the lower risk, but larger population, efforts to impact the sudden death rate must focus on developing risk stratification in this group. The initial step must be a quick and rapid screen. The present study provides such a simple screening tool using easily available clinical and ECG variables.

Although other broad population studies have evaluated risk factors for sudden cardiac death³¹⁻³⁵, none have evaluated the utility of a specific algorithm to perform screening of a large population. To the best of our knowledge, this is the first study using a large population data providing SCA risk assessment. The AUC = 0.85 as observed in the ECG model should be considered very good for screening purposes in a large population. By way of comparison, the AUC of the Framingham risk score alone and the Framingham risk score with additional predictors for predicting cardiovascular disease has ranged from 0.50 to 0.83 (median, 0.74) and 0.57 to 0.84 (median, 0.75), respectively¹⁹.

Variables in the Multivariate Model and their Relationship with Other Studies

Most variables remaining in our multivariate model have been identified as independent risk factors before. Age, BMI, and atrial rate were identified as independent risk factors in the Paris Population Study³² following a patient population over 23 year. Prior MI and CHF are also well known independent risk factors^{8,9,15,16,36,37}. QTc was identified as a risk factor in several studies^{8,38,39}. The inverse relationship of diastolic blood pressure with SCA risk, as identified in

our study, has also been previously identified⁴⁰. The negative correlation of lower LDL with SCA risk may be indicative of an aggressive use of LDL lowering statin therapy in the population with MI and CHF^{19,20}. However, we also found that the exclusion of LDL results in a model with no difference in AUC values.

The American College of Cardiology (ACC), American Heart Association (AHA), European Society of Cardiology (ESC), the Heart Rhythm Society and the U.S. Department of Health and Human Services use LVEF < 35% as a primary screening criterion for SCA. In our dataset only 29% of the patients with SCA had LVEF < 35%. This is consistent with the findings of the Oregon Sudden Unexpected Death Study¹⁰ and the Maastricht Study⁴¹. Exclusion of the LVEF in the model that included only clinical and ECG variables only minimally reduced the predictive value of the model.

Implications for use in Further Invasive and Non-invasive Risk Stratification

The results presented here are important for future risk stratification studies. While many ECG based risk stratification tests such as heart rate variability, heart rate turbulence, heart rate recovery after exercise, baroreceptor sensitivity, baro-reflex sensitivity, deceleration capacity of heart rate, R-wave and T-wave morphology (length, area, deflection, amplitude, T-wave alternans, etc.), QRS-complex morphology (center of mass), wave-absence (P, T), morphologic variability (e.g., RR interval, etc.), spectral energy, and frequency range changes in certain regions, etc., have been identified, and efforts to identify new risk marker continue⁴²⁻⁴⁸, these are not practical tests to administer as a screening tool in the general population. On the other hand, using the current model as a “baseline model” for screening, it is possible that some of these tests

could provide further incremental risk stratification. It will be important to evaluate the ability of this next tier of tests to appropriately reclassify individual risk.

Limitations

The present study is based on a retrospective analysis of data available from a single institution. In addition, the choice of variables in this study was limited by the data availability in EDW for patients with SCA events. Furthermore, the data used in the study does not distinguish between in and out-of-hospital cardiac arrest, and it is limited by digitized records in EDW. While variables measured at multiple time points were considered by taking average, minimum, maximum, and most recent values, the predictive model suggested in the present study may still not capture the risk resulting from temporal variations in these markers. Recent studies have presented the importance of risk assessment timing to effectively capture the temporal variations in data^{49,50}. Nevertheless, we find that the risk assessment from our model is robust in terms of availability of patient data that are easily collectable in an outpatient clinical setting.

Implications for Population Risk Stratification

Ideally, the goal of identifying individuals with high SCA risk is to be able to provide appropriate pharmacological and device therapy. While these therapies may evolve over time or new ones may be introduced, risk stratification will remain an important component of the evaluation of the risk-benefit and cost-benefit ratios for these therapies. A multivariate score of SCA using standard risk factors is particularly relevant for primary screening of patients in office-based primary care practices. The risk assessment function is easily implementable in the electronic medical record or can be provided as an online service²⁶.

REFERENCE

1. Fishman GI, Chugh SS, Dimarco JP, et al. Sudden cardiac death prediction and prevention: report from a National Heart, Lung, and Blood Institute and Heart Rhythm Society Workshop. *Circulation*. Nov 30 2010;122(22):2335-2348.
2. Moss AJ, Zareba W, Hall WJ, et al. Prophylactic Implantation of a Defibrillator in Patients with Myocardial Infarction and Reduced Ejection Fraction. *New England Journal of Medicine*. 2002;346(12):877-883.
3. Das MK, Khan B, Jacob S, Kumar A, Mahenthiran J. Significance of a Fragmented QRS Complex Versus a Q Wave in Patients With Coronary Artery Disease. *Circulation*. May 30, 2006 2006;113(21):2495-2501.
4. Monnig G, Eckardt L, Wedekind H, et al. Electrocardiographic risk stratification in families with congenital long QT syndrome. *European heart journal*. Sep 2006;27(17):2074-2080.
5. Buxton AE, Lee KL, Fisher JD, Josephson ME, Prystowsky EN, Hafley G. A Randomized Study of the Prevention of Sudden Death in Patients with Coronary Artery Disease. *New England Journal of Medicine*. 1999;341(25):1882-1890.
6. Moss AJ, Hall WJ, Cannom DS, et al. Improved Survival with an Implanted Defibrillator in Patients with Coronary Disease at High Risk for Ventricular Arrhythmia. *New England Journal of Medicine*. 1996;335(26):1933-1940.
7. Bardy GH, Lee KL, Mark DB, et al. Amiodarone or an Implantable Cardioverter-Defibrillator for Congestive Heart Failure. *New England Journal of Medicine*. 2005;352(3):225-237.

8. Zipes DP, Camm AJ, Borggrefe M, et al. ACC/AHA/ESC 2006 Guidelines for Management of Patients With Ventricular Arrhythmias and the Prevention of Sudden Cardiac Death. *Circulation*. September 5, 2006 2006;114(10):e385-e484.
9. Members WC, Antman EM, Anbe DT, et al. ACC/AHA Guidelines for the Management of Patients With ST-Elevation Myocardial Infarction, Executive Summary. *Circulation*. August 3, 2004 2004;110(5):588-636.
10. Stecker EC, Vickers C, Waltz J, et al. Population-Based Analysis of Sudden Cardiac Death With and Without Left Ventricular Systolic Dysfunction: Two-Year Findings from the Oregon Sudden Unexpected Death Study. *Journal of the American College of Cardiology*. 3 2006;47(6):1161--1166.
11. Goldberger JJ. Evidence-based analysis of risk factors for sudden cardiac death. *Heart Rhythm*. 2009;6(3 Suppl):S2-7.
12. Buxton AE, Lee KL, Hafley GE, et al. Limitations of Ejection Fraction for Prediction of Sudden Death Risk in Patients With Coronary Artery Disease: Lessons From the MUSTT Study. *Journal of the American College of Cardiology*. 9 2007;50(12):1150--1157.
13. Goldenberg I, Vyas AK, Hall WJ, et al. Risk Stratification for Primary Implantation of a Cardioverter-Defibrillator in Patients With Ischemic Left Ventricular Dysfunction. *Journal of the American College of Cardiology*. January 22, 2008 2008;51(3):288-296.
14. Buxton AE, Moss AJ. Should everyone with an ejection fraction less than or equal to 30% receive an implantable cardioverter-defibrillator? *Circulation*. 2005;111(19):2537.
15. Goldberger JJ, Buxton AE, Cain M, et al. Risk Stratification for Arrhythmic Sudden Cardiac Death. *Circulation*. May 31, 2011 2011;123(21):2423-2430.

16. Zipes DP, Wellens HJJ. Sudden Cardiac Death. *Circulation*. November 24, 1998
1998;98(21):2334-2351.
17. Parmley WW, Chatterjee K. Congestive heart failure and arrhythmias: an overview. *The American Journal of Cardiology*. 1986;57(3):B34-B37.
18. Kannel WB, Schatzkin A. Sudden death: lessons from subsets in population studies. *Journal of the American College of Cardiology*. 1985;5(6):141B-149B.
19. Tzoulaki I, Liberopoulos G, Ioannidis JPA. Assessment of claims of improved prediction beyond the Framingham risk score. *JAMA: the journal of the American Medical Association*. 2009;302(21):2345-2352.
20. D'Agostino RB, Vasan RS, Pencina MJ, et al. General Cardiovascular Risk Profile for Use in Primary Care. *Circulation*. February 12, 2008 2008;117(6):743-753.
21. Lloyd-Jones DM, Liu K, Tian L, Greenland P. Narrative Review: Assessment of C-Reactive Protein in Risk Prediction for Cardiovascular Disease. *Annals of internal medicine*. July 4, 2006 2006;145(1):35-42.
22. Northwestern University. Northwestern Medical Enterprise Data Warehouse. 2011;
<http://informatics.northwestern.edu/>.
23. Hosmer DW, Lemeshow S. *Applied logistic regression*. 2nd ed. New York: Wiley; 2000.
24. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. *BeiJing: Publishing House of Electronics Industry*. 2004.
25. *Stata Statistical Software: Release 12* [computer program]: College Station, TX: StataCorp LP; 2011.

26. Pouleur AC, Barkoudah E, Uno H, et al. Pathogenesis of Sudden Unexpected Death in a Clinical Trial of Patients With Myocardial Infarction and Left Ventricular Dysfunction, Heart Failure, or Both. *CLINICAL PERSPECTIVE. Circulation*. 2010;122(6):597-602.
27. Pratt LA, Ford DE, Crum RM, Armenian HK, Gallo JJ, Eaton WW. Depression, psychotropic medication, and risk of myocardial infarction: prospective data from the Baltimore ECA follow-up. *Circulation*. 1996;94(12):3123-3129.
28. Myerburg M, Robert J, Interian A, et al. Frequency of sudden cardiac death and profiles of risk. *The American Journal of Cardiology*. 1997;80(5):10F-19F.
29. Goldberger J, Cain M, Hohnloser S, et al. AHA/ACC/HRS scientific statement on non-invasive risk stratification techniques for identifying patients at risk for sudden cardiac death: A scientific statement from the American Heart Association Council on Clinical Cardiology Committee on Electrocardiography and Arrhythmias and Council on Epidemiology and Prevention. *Circulation*. 2008;118:1497-1518.
30. Goldenberg I, Moss AJ, McNitt S, et al. Time dependence of defibrillator benefit after coronary revascularization in the Multicenter Automatic Defibrillator Implantation Trial (MADIT)-II. *Journal of the American College of Cardiology*. 2006;47(9):1811-1817.
31. Albert CM, Chae CU, Grodstein F, et al. Prospective study of sudden cardiac death among women in the United States. *Circulation*. 2003;107(16):2096-2101.
32. Jouven X, Desnos M, Guerot C, Ducimetière P. Predicting Sudden Death in the Population : The Paris Prospective Study I. *Circulation*. April 20, 1999
1999;99(15):1978-1983.
33. Seccareccia F, Pannoizzo F, Dima F, et al. Heart rate as a predictor of mortality: the MATISS project. *American journal of public health*. 2001;91(8):1258.

34. Gorgels APM, Gijsbers C, de Vreede-Swagemakers J, Lousberg A, Wellens HJJ. Out-of-hospital cardiac arrest-the relevance of heart failure. The Maastricht Circulatory Arrest Registry. *European heart journal*. 2003;24(13):1204-1209.
35. Jouven X, Empana J-P, Schwartz PJ, Desnos M, Courbon D, Ducimetière P. Heart-Rate Profile during Exercise as a Predictor of Sudden Death. *New England Journal of Medicine*. 2005;352(19):1951-1958.
36. Piccini JP, Zhang M, Pieper K, et al. Predictors of sudden cardiac death change with time after myocardial infarction: results from the VALIANT trial. *European heart journal*. January 1, 2010 2010;31(2):211-221.
37. Members WC, Epstein AE, DiMarco JP, et al. ACC/AHA/HRS 2008 Guidelines for Device-Based Therapy of Cardiac Rhythm Abnormalities. *Circulation*. May 27, 2008 2008;117(21):e350-e408.
38. Straus SMJM, Kors JA, De Bruin ML, et al. Prolonged QTc Interval and Risk of Sudden Cardiac Death in a Population of Older Adults. *Journal of the American College of Cardiology*. December 22, 2005 2005;j.jacc.2005.2008.2067.
39. Algra A, Tijssen JG, Roelandt JR, Pool J, Lubsen J. QTc prolongation measured by standard 12-lead electrocardiography is an independent risk factor for sudden death due to cardiac arrest. *Circulation*. Jun 1991;83(6):1888-1894.
40. Goldenberg I, Moss AJ, McNitt S, et al. Inverse Relationship of Blood Pressure Levels to Sudden Cardiac Mortality and Benefit of the Implantable Cardioverter-Defibrillator in Patients With Ischemic Left Ventricular Dysfunction. *Journal of the American College of Cardiology*. April 3, 2007 2007;49(13):1427-1433.

41. de Vreede-Swagemakers JJM, Gorgels APM, Dubois-Arbouw WI, et al. Out-of-Hospital Cardiac Arrest in the 1990s: A Population-Based Study in the Maastricht Area on Incidence, Characteristics and Survival. *Journal of the American College of Cardiology*. 1997;30(6):1500-1505.
42. Barron HV, Lesh MD. Autonomic nervous system and sudden cardiac death. *Journal of the American College of Cardiology*. 1996;27(5):1053-1060.
43. Bauer A, Kantelhardt JW, Barthel P, et al. Deceleration capacity of heart rate as a predictor of mortality after myocardial infarction: cohort study. *The Lancet*. 2006;367(9523):1674-1681.
44. Rashed U, Mirza MJ. Identification of sudden cardiac death using spectral domain analysis of Electrocardiogram (ECG)2008.
45. Schwartz P, La Rovere M, Vanoli E. Autonomic nervous system and sudden cardiac death. Experimental basis and clinical observations for post-myocardial infarction risk stratification. *Circulation*. 1992;85(1 Suppl):I77.
46. Steinberg JS, Regan A, Sciacca RR, Bigger JT, Fleiss JL. Predicting arrhythmic events after acute myocardial infarction using the signal-averaged electrocardiogram. *The American journal of cardiology*. 1992;69(1):13-21.
47. Syed Z, Scirica B, Stultz CM, Gutttag JV. Electrocardiographic prediction of arrhythmias2009.
48. Zabel M, Acar B, Klingenhoben T, Franz MR, Hohnloser SH, Malik M. Analysis of 12-lead T-wave morphology for risk stratification after myocardial infarction. *Circulation*. 2000;102(11):1252-1257.

49. Exner DV, Kavanagh KM, Slawnych MP, et al. Noninvasive Risk Assessment Early After a Myocardial Infarction: The REFINE Study. *Journal of the American College of Cardiology*. 2007;50(24):2275-2284.
50. Huikuri HV, Raatikainen MJP, Moerch-Joergensen R, et al. Prediction of fatal or near-fatal cardiac arrhythmia events in patients with depressed left ventricular function after an acute myocardial infarction, Å†. *European heart journal*. March 1, 2009 2009;30(6):689-698.

Table 1. Characteristics of the study population from Northwestern Medical Enterprise Data Warehouse*

Characteristics	SCA Present (n = 73)	SCA Absent (n = 22,968)	P Value
Clinical Variables			
Age, mean (SD), y	67.74 (12.55)	60.20 (14.29)	< .001
Gender			
Female, No. (%)	34 (46.58)	12,031 (52.38)	
Male, No. (%)	39 (53.42)	10,937 (47.62)	.38
Race			
Asian, No. (%)	2 (2.74)	566 (2.46)	
Black or African American, No. (%)	25 (34.25)	7,110 (30.96)	
Other, No. (%)	1 (1.37)	942 (4.10)	.66
White, No. (%)	45 (61.64)	14,351 (62.48)	
Body Mass Index, mean (SD), lb/in ²	30.42 (8.61)	29.98 (8.72)	.001
Systolic BP, mean (SD), mmHg	115.83 (25.32)	126.17 (24.89)	< .001
Diastolic BP, mean (SD), mmHg	65.72 (14.49)	69.25 (14.81)	< .001
Congestive Heart Failure, No. (%)	48 (65.75)	4,436 (19.31)	< .001
Myocardial Infarction, No. (%)	15 (20.55)	1,543 (6.72)	< .001
On-Statin, No. (%)	53 (72.60)	12,567 (54.72)	.001
ECG Variables			
Atrial Rate, mean (SD), bpm	99.48 (59.51)	85.1 (46.95)	< .001
Ventricular Rate, mean (SD), bpm	87.62 (26.41)	77.76 (20.15)	< .001
PR interval, mean (SD), ms	164.33 (39.39)	165.29 (34.01)	.42
QRS duration, mean (SD), ms	114.54 (33.91)	96.88 (24.55)	< .001
QT interval, mean (SD), ms	422.6 (67.72)	413.34 (52.06)	< .001
QTc, mean (SD), ms	474.01 (49.81)	448.5 (41.1)	< .001
P-axis, mean (SD), degree	48.31 (36.78)	48.77 (25.04)	.68
R-axis, mean (SD), degree	12.61 (62.5)	18.64 (47.94)	< .001
T-axis, mean (SD), degree	47.44 (85.84)	43.32 (56.98)	.07
LVEF, mean (SD), %	33.63 (16.95)	48.47 (17.19)	< .001
Lipid Variables			
LDL, mean (SD), mg/dL	81.09 (36.23)	101 (39.59)	< .001
HDL, mean (SD), mg/dL	37.5 (14.84)	44.82 (16.52)	< .001
Triglyceride, mean (SD), mg/dL	141.29 (123.61)	132.78 (249.72)	.32

* Multiple records of a variable for a patient were treated as separate patient records when calculating the mean and standard deviation given in this table.

Table 2. A Population based multivariate logistic model for sudden cardiac arrest risk assessment

Variable	Coefficient (95% CI)	OR* (95% CI)	P value	Variable Type
Base Model				
CHF	1.08 (0.53 – 1.62)	2.93 (1.70 – 5.05)	< .001	Categorical
Ventricular Rate (Max.)	0.02 (0.02 – 0.03)	1.36 (1.24 – 1.50)	< .001	Continuous
MI	0.68 (0.09 – 1.28)	1.98 (1.09 – 3.59)	.02	Categorical
LDL (Max.)	-0.01 (-0.02 – -0.00)	0.84 (0.75 – 0.95)	.005	Continuous
Diastolic BP (Recent)	-0.03 (-0.05 – -0.01)	0.66 (0.51 – 0.84)	.001	Continuous
QTc (Min.)	0.01 (0.00 – 0.01)	1.21 (1.03 – 1.41)	.02	Continuous
P-axis (Avg.)	-0.01 (-0.02 – 0.00)	0.86 (0.73 – 1.00)	.06	Continuous
BMI (Avg.)	0.03 (0.01 – 0.06)	1.15 (1.02 – 1.28)	.02	Continuous
LVEF (Max.)	-0.02 (-0.04 – -0.01)	0.82 (0.71 – 0.94)	.004	Continuous
Age	0.02 (0.00 – 0.04)	1.13 (1.01 – 1.26)	.04	Continuous
Constant	-8.83 (-12.5 – -4.82)	NA	< .001	NA
ECG Model				
CHF	1.33 (0.81 – 1.85)	3.78 (2.25 – 6.35)	< .001	Categorical
Ventricular Rate (Max.)	0.02 (0.02 – 0.03)	1.37 (1.25 – 1.51)	< .001	Continuous
MI	0.74 (0.15 – 1.33)	2.10 (1.16 – 3.79)	.01	Categorical
Diastolic BP (Recent)	-0.04 (-0.06 – -0.02)	0.63 (0.49 – 0.80)	< .001	Continuous
QTc (Min.)	0.01 (0.00 – 0.02)	1.30 (1.12 – 1.52)	.001	Continuous
P-axis (Avg.)	-0.01 (-0.02 – 0.00)	0.85 (0.72 – 1.00)	.04	Continuous
BMI (Avg.)	0.03 (0.00 – 0.05)	1.12 (1.00 – 1.26)	.06	Continuous
Age	0.02 (0.00 – 0.04)	1.13 (1.01 – 1.26)	.03	Continuous
Constant	-12.2 (-15.9 – -8.50)	NA	< .001	NA

* Odds ratio. The odds are calculated per 0.5 standard deviation of value for the continuous variables.

Table 3. Predictability of the population based multivariate logistic regression model and variable sensitivity results

Variable set (No. of variables)	Total patients, No.	SCA, No. (%)	AUC*	FPR*	Mean AUC (95% CI) [†]	P-value
Base (10)	23,041	73 (0.32)	0.87	2.4%	0.86 (0.81 – 0.90)	NA
ECG (8)	23,041	73 (0.32)	0.87	4.1%	0.85 (0.80 – 0.90)	.30
No LVEF (9)	23,041	73 (0.32)	0.88	2.7%	0.86 (0.82 – 0.90)	.90
No MI-CHF (8)	23,041	73 (0.32)	0.85	3.1%	0.82 (0.77 – 0.88)	< .001
Age+LVEF (2)	23,041	73 (0.32)	0.73	7.9%	0.73 (0.65 – 0.78)	< .001
LVEF (1)	23,041	73 (0.32)	0.64	8.5%	0.64 (0.57 – 0.72)	< .001
Recent (7)	23,041	73 (0.32)	0.83	2.7%	0.81 (0.76 – 0.86)	< .001
ECG (Larger) Validation	69,481	189 (0.27)	0.85	3.0%	0.85 (0.82 – 0.88)	.156

* This was calculated using 100% of patients for training and subsequent testing the model.

† This was calculated by 2-fold cross validation.

Table 4. Patient characteristics on a larger cohort using the clinical variables and the ECG variables*

Characteristics	SCA Present (n = 189)	SCA Absent (n = 69,481)	P-value
Clinical Variables			
Age, mean (SD), y	66.57 (13.70)	57.26 (14.83)	.07
Gender			
Female, No. (%)	91 (48.15)	38,246 (55.05)	
Male, No. (%)	98 (51.85)	31,235 (44.95)	.07
Race			
Asian, No. (%)	7 (3.70)	1,578 (2.27)	
Black or African American, No. (%)	58 (30.69)	17,455 (25.12)	
Other, No. (%)	2 (1.06)	3,895 (5.61)	.01
White, No. (%)	122 (64.55)	46,553 (67.00)	
Body Mass Index, mean (SD), lb/in²			
Body Mass Index, mean (SD), lb/in ²	29.42 (7.93)	29.46 (8.33)	.70
Diastolic BP, mean (SD), mmHg			
Diastolic BP, mean (SD), mmHg	65.89 (14.16)	69.68 (14.27)	< .001
Systolic BP, mean (SD), mmHg			
Systolic BP, mean (SD), mmHg	118.15 (24.72)	125.41 (23.64)	< .001
Congestive Heart Failure, No. (%)			
Congestive Heart Failure, No. (%)	114 (60.32)	6,469 (9.31)	< .001
Myocardial Infarction, No. (%)			
Myocardial Infarction, No. (%)	28 (14.81)	2,159 (3.11)	< .001
On-Statin, No. (%)			
On-Statin, No. (%)	125 (66.14)	22,307 (32.11)	.001
ECG Variables			
Atrial Rate, mean (SD), bpm			
Atrial Rate, mean (SD), bpm	99.85 (61.10)	82.92 (42.53)	< .001
Ventricular Rate, mean (SD), bpm			
Ventricular Rate, mean (SD), bpm	87.16 (25.63)	77.11 (19.38)	< .001
PR interval, mean (SD), ms			
PR interval, mean (SD), ms	165.47 (38.38)	164.24 (32.18)	.19
QRS duration, mean (SD), ms			
QRS duration, mean (SD), ms	110.02 (31.83)	94.45 (22.32)	< .001
QT interval, mean (SD), ms			
QT interval, mean (SD), ms	417.79 (66.44)	409.18 (49.35)	< .001
QTc, mean (SD), ms			
QTc, mean (SD), ms	467.52 (47.89)	443.1 (38.61)	< .001
P-axis, mean (SD), degree			
P-axis, mean (SD), degree	48.86 (35.17)	45.15 (24.07)	< .001
R-axis, mean (SD), degree			
R-axis, mean (SD), degree	14.36 (59.57)	22.76 (45.41)	< .001
T-axis, mean (SD), degree			
T-axis, mean (SD), degree	44.76 (83.23)	42.02 (50.13)	.12

* Multiple records of a variable for a patient were treated as separate patient records when calculating the mean and standard deviation given in this table.

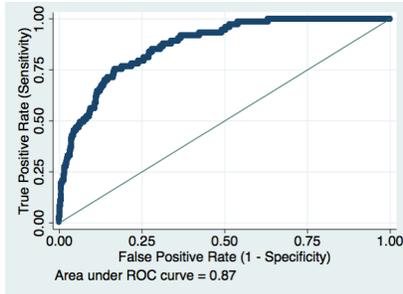


Figure 1a. Receiver operating characteristic curve of the multivariate logistic regression model for sudden cardiac arrest

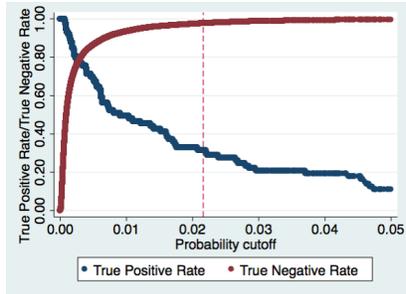


Figure 1b. True positive rate (sensitivity) and true negative rate (specificity) curves of the multivariate logistic regression model for sudden cardiac arrest

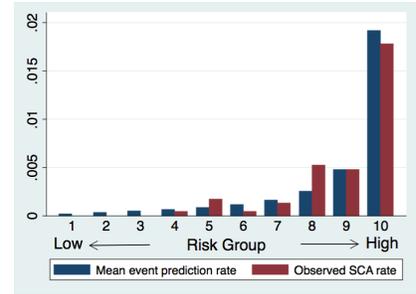


Figure 1c. Risk group deciles by mean event prediction rate and the observed SCA event rates in the study population.