Article:
Hadanny, A, Shouval, R, Wu, J orcid.org/0000-0001-6093-599X et al. (8 more authors) (2022) Machine learning-based prediction of 1-year mortality for acute coronary syndrome. Journal of Cardiology, 79 (3). pp. 342-351. ISSN 0914-5087

https://doi.org/10.1016/j.jjcc.2021.11.006

Machine learning-based prediction of 1-year mortality for acute coronary syndrome

Amir Hadanny MD[1,4], Roni Shouval MD, PhD[2], Jianhua Wu PhD, Chris P Gale MD X, Ron Unger PhD[3], Doron Zahger MD[4], Ilan Goldenberg MD[6], Roy Beigel MD[7], Shlomi Matetzky[7], Zaza Iakobishvili MD, PhD[8]

Correspondence: Amir Hadanny, MD, Sagol Center, Shamir Medical Center, 70300, Israel. Phone (972) 89779395; Fax: (972) 39542238; E-mail: amir.had@gmail.com

Acknowledgments

Source of funding

The authors had no funding source or grant for this study.

Conflict of interests

The authors had no conflict of interests

[1] Sagol Center for Hyperbaric medicine and Research, Assaf HaRofe Medical Center, Ramle, Israel; The Mina and Everard Goodman Faculty of Life Sciences, Bar-Ilan University, Ramat-Gan, Israel. Sackler School of Medicine, Tel-Aviv University, Tel-Aviv, Israel. This author takes responsibility for all aspects of the reliability and freedom from bias of the data presented and their discussed interpretation

[2] Adult Bone Marrow Transplantation Service, Memorial Sloan Kettering Cancer Center, New York, NY. The Mina and Everard Goodman Faculty of Life Sciences, Bar-Ilan University, Ramat-Gan, Israel. Sackler School of Medicine, Tel-Aviv University, Tel-Aviv, Israel. This author takes responsibility for all aspects of the reliability and freedom from bias of the data presented and their discussed interpretation

[3] The Mina and Everard Goodman Faculty of Life Sciences, Bar-Ilan University, Ramat-Gan, Israel. This author takes responsibility for all aspects of the reliability and freedom from bias of the data presented and their discussed interpretation

[4] Department of Cardiology, Soroka University Medical Center, Faculty of Health Sciences, Ben Gurion University of the Negev. This author takes responsibility for all aspects of the reliability and freedom from bias of the data presented and their discussed interpretation

X Leeds Institute of Cardiovascualr and Metabolic Medicine, University of Leeds, Leeds UK

X Leeds Institue for Data Analytics, University of Leeds, Leeds, UK

X Department of Cadiology, Leeds Teaching Hospitals NHS Trust, Leeds, UK

Keywords: STEMI, machine learning, data mining, outcome, mortality

Abstract

Background:

Most previous studies have used Cox-regression for survival analysis in patients with acute coronary syndromes (ACS). However, such classic survival analysis carries several assumptions, of which the most notable one is the proportional hazards.

Objective:

The primary objective was to compare survival prediction performance for 1-year mortality following ACS for Cox-regression (CPH) with two machine learning-based models (random survival forest (RSF) and deep learning (DeepSurv)). The secondary objective was to externally validate the findings using a nationwide registry of ACS.

Methods:

This was a retrospective, supervised learning data mining study based on the Acute Coronary Syndrome Israeli Survey (ACSIS) registry and the Myocardial Ischemia National Audit Project (MINAP). The ACSIS data was divided to train/test in a 70/30 fashion. Models were then externally validated on the MINAP data. Harell's C-index, inverse probability of censoring weighting (IPCW) and the Brier scores were used for model performance comparisons.

Results:

RSF performed best among the three models, Harell's c-index on training and testing sets reaching 0.953 and 0.924, followed by CPH Multivariate selected model (0.805, 0.849), CPH Univariate selected model (0.828, 0.806) and DeepSurv model (0.801, 0.804), and finally the traditional CPH model (0.826, 0.738). The CPH model performance on the validation set had

Harell's C-index of 0.713-0.818, 0.689-0.790 for IPCW and 0.094-0.100 for brier score. The RSF model had the highest performance on the validation data set with 0.811 for Harell's C-index, 0.844 for IPCW and 0.093 for brier score.

Conclusions*:*

This study demonstrates that RSF survival predictions for ACS long-term mortality show improved model performance compared with the classic statistical method. This can benefit patients by stratifying risks and guiding treatment options to save more lives, as well as by avoiding ineffective/unnecessary treatments.

1. Introduction

Cardiovascular diseases are the leading global cause of death, accounting for approximately 7.5 million death annually[1,2]. Acute coronary syndrome (ACS), defined as unstable angina (UA), non-ST segment elevation myocardial infarction (NSTEMI), and ST-segment elevation myocardial infarction (STEMI), causes most cardiovascular-related deaths, which represent 1.8 million deaths per year[1].

Despite tremendous achievements in the management of ACS over recent years[3,4] , the current reported rates of 1-year mortality following STEMI range of 7%-11.5%[5,6] and 14.3-14.8% for UA and NSTEMI[4,7,8] Various prognostic models for long-term mortality prediction following ACS have been developed over the last two decades, mostly through regression-based models. When it comes to the survival prediction, binary/dichotomous outcome (dead/alive) cannot sufficiently characterize the outcome but should incorporate 'time to event' as well[9].

Most previous studies have used Cox-regression for survival analysis in patients after ACS[8,10]. However, such classic survival analysis carries several assumptions and limitations. The most notable is the proportional hazards – meaning the hazard ratio for an individual with a certain risk factor to that for an individual without the risk factor is constant over time. Another assumption of the cox-regression model is that the outcome (in our case, mortality) is a linear combination of covariates that may be too simplistic for proper prognosis prediction. Additionally, Cox-based models may suffer from poor performance and high variance when including multiple variables[11]. Thus, to account for non-linear associations, violation of the proportional hazard assumption, and to reduce bias associated with feature selection, non-parsimonious non-parametric machine learning (ML) algorithms are suggested[12].

Random survival forest (RSF) model is an extension of the random forest model for survival analysis. By automatically assessing the complex effects and interactions among all variables from an unbiased view, it has demonstrated improved performance over the traditional cox proportional hazards regression (CPH) model in some cases[13].

In previous studies, artificial neural networks with various designs performed similarly to standard CPH [14,15]. DeepSurv, by Katzman et al., is a deep learning-based survival prediction algorithm that uses a multi-layer feed-forward network, which can model high-level interaction terms beyond the linear Cox model [16]. Kwon et al. recently showed this deep learning-based model predicted in-hospital and 12-months mortality with high performance compared to the global registry of acute coronary event (GRACE) and thrombolysis in myocardial infarction (TIMI)[17].

The aim of the study was to compare predictive performance of two ML algorithms designed for time-to-event outcomes – RSF and DeepSurv to the 1-year mortality after ACS prediction by traditional CPH.

2. Methods

2.1 Study population

This was a retrospective, supervised learning (i.e., an ML task of inferring a function from labeled training data), data mining study based on the Acute Coronary Syndrome Israeli Survey (ACSIS) registry and the Myocardial Ischemia National Audit Project (MINAP).

ACSIS is a mostly biennial prospective observational national survey of all patients with acute coronary syndrome (ACS) hospitalized in Israel.[18,19] In 2000, 2002, 2004, 2006, 2008, 2010, 2013, 2016 during a two-month period, data are prospectively collected from all ACS admissions

in each of the 25 cardiology wards operating in Israel.[18-21] All patients are included (without exclusion) and followed for one month. The number of patients who refused to participate was substantially small (4-5 patients each cycle – which summed to 20 patients in 8 years). The ACSIS was approved by each medical center's institutional review board. Clinical data were recorded on pre-specified forms for all patients during the years 2000 to 2016. Case report form definitions were centrally determined. The attending physicians determined admission and discharge diagnoses based on clinical, electrocardiographic, and biochemical criteria. Patients' management was at the discretion of the attending physicians.

MINAP is a national registry of patients admitted to hospitals in England and Wales with acute coronary syndromes (ACS) [6]. Data collection began in October 2000. MINAP accrues approximately 85,000 episodes of care per year for patients with ACS admitted to all acute National Health Service (NHS) hospitals in England and Wales. The dataset comprises over 120 separate fields under the following groups: patient demographics, admission method, clinical features and investigations, medical history, drug treatment before admission, detail of primary reperfusion treatment, drug treatment in hospital, clinical complications, interventional treatments, hospital outcome, discharge diagnosis and discharge (secondary prevention) treatment.

The current analysis included all ACS (STEMI, nSTEMI, UAP) patients admitted between 2006 and 2016. Statistical guidance articles have previously stated that bias is likely in analyses with more than 10% missingness and that if more than 40% of data are missing in important variables then results should not be considered[22]. Thus, variables with more than 15% missing values were not included. In addition to mortality and time, 69 variables that were available on both ACSIS

and MINAP databases (i.e., demographics, prior medical history, prior chronic drugs, clinical presentation, basic laboratory data with admission) were evaluated (Supplementary Table-S1).

The 1-year mortality time frame in ACSIS and MINAP was defined as one year from admission day. In addition to the mortality variable, all included patients had a known time to event (within 365 days). Patients alive at 1-year had a time value set to 365.

All patients gave written informed consent for inclusion in the registries. The study follows the TRIPOD (Transparent Reporting of a Multivariable prediction model for Individual Prognosis or Diagnosis) recommendations for the reporting of studies prediction scores[23].

2.2 Study Objectives

The primary objective was to compare RSF and DeepSurv and CPH models' prediction of 1-year mortality on the ACSIS data. The secondary objective was to perform external validation of the findings using the MINAP data.

2.3 Data processing

During the data preprocessing, all outliers in numeric data were converted into a null value. The ACSIS data was divided to train/test in a 70/30 fashion. Missing values on both datasets were replaced using the mean and mode values for numeric and nominal variables, respectively. Model hyperparameters were tuned on the training dataset. The models were trained on the ACSIS training dataset and tested on the ACSIS test dataset. Next, the models were externally validated on the MINAP data.

2.4 Models

2.4.1 CPH

Cox proportional hazards models are survival models that assume all patients share a common baseline hazard function multiplied by a factor based on the values of various predictor variables for an individual. The CPH model was built with the training set. Statistically significant variables in univariate analyses were taken into multivariable analysis. CPH model was trained and evaluated using the 'Survival' R package.

2.4.2 RSF

Random survival forest is an alternative method for survival analysis. It is a machine-learning technique that builds a 'forest' of decision trees, each of which calculates patient outcomes by splitting them into groups with similar characteristics. These thousands of decision trees each have random imperfections, meaning that while individual trees are then relatively poor predictors, the averaged result from the forest is more accurate and less prone to overfitting than an individual 'perfect' decision tree.

At each node in a decision tree, starting at its root, patients are split into two branches by looking at a single covariate. The algorithm selects a split point that maximizes the difference between the survival curves of patients in the two branches defined by that split. Split points are defined to maximize the homogeneity within each branch and the inhomogeneity between them. This process is repeated until the leaves of the tree are reached. These are nodes where either there are no further criteria remains by which the patients at that leaf can be distinguished, including the possibility of only a single patient remaining, or splitting may be stopped early with a minimum node size or maximum tree depth[24].

RSF Random survival forest extends random forest methodology for survival analysis. It is an alternative method for survival analysis. Thus, each tree is grown by randomly selecting a subset

of variables at each node and then splitting the node using a survival criterion involving survival time and event status information[25].

Random survival forest hyperparameters were tuned using the training set. Random forest model was trained and evaluated using the RandomForestSRC R package[26].

2.4.3 DeepSurv

DeepSurv is a multi-layer perceptron that predicts a patient's risk of an event. As a deep neural network, it provides a non-linear method (in contrast to the linear CPH) to model high level interactions. The network's output is a single node, which estimates the risk function parameterized by the network's weights. The network is built using a deep architecture (i.e., more than one hidden layer) with modern techniques (including weight decay regularization, batch normalization, scaled exponential Linear Units (SELU), gradient descent optimization algorithms, gradient clipping and learning rate scheduling)[16]. The network propagates the inputs through a number of hidden layers with weights. The hidden layers consist of fully connected non-linear activation functions followed by dropout. The final layer is a single node that performs a linear combination of the hidden features. The network's hyper-parameters were first tuned on the training set using a Random hyper-parameter optimization search. DeepSurv was implemented as an open-source Python module (https://github.com/jaredleekatzman/DeepSurv). Following training and prediction on the test set with Python, probabilities were then processed and analyzed in R.

2.5 Performance

2.5.1 Concordance (C-Index)

Harell's C-index is a method for estimating prediction error[27]. The c-index is used mainly as a metric for survival prediction and reflects a measure of how well a model predicts the ordering of patients' death times. A c = 0.5 is the average of a random model, and c = 1 refers to a perfect match of death time ranking.

2.5.2 Time-dependent AUC

The receiver operating characteristic (ROC) curve displays the sensitivity (true positive rate) versus 1- specificity (false positive rate) for all possible cut-points that define a binary test by dichotomizing a quantitative marker. The area under the ROC curve (AUC) is often used to summarize and compare diagnostic accuracy of several markers. Uno et al. [9] and Hung and Chiang [10] proposed a nonparametric estimator of the time-dependent ROC curve using the inverse probability of censoring weighting (IPCW) approach. The rationale of the IPCW approach is to mainly use the observed cases and controls and weigh them by their probability of being observed. The IPCW method was applied using the TimeROC package in R.

2.5.3 The Brier Score

The Brier score (BS) for binary classification is a cost function (or loss function) that measures the accuracy of probabilistic predictions. The lower and closer to 0 the Brier score (0-1) is for a set of predictions, the better the predictions are calibrated. Brier score was applied using the 'pec' package in R and lifelines in Python.

2.6 Statistical analysis

Results were expressed as mean and SD for parametric variables and as frequencies/percentages for nonparametric variables. Baseline differences between the training set and testing set were assessed using the two-sided independent student t-test for continuous variables and Chi-square test, Fisher's exact test for categorical variables. Survival curves were plotted using the Kaplan-Meier method and compared using the log-rank test. To estimate the prognostic effect of the features, univariate and multiple CPH regression analyses were done as well.

The analysis was performed using R programming language version 3.6.1 (R Core Team, Vienna, Austria, 2018).

3.Results

3.1 Patient characteristics

Of the 15,212 patients in the ACSIS registry, 9270 fulfilled inclusion criteria for model development (5935 were treated prior to 2006 and 7 other had unknown ACS diagnosis) . The average age was 63.8 ±12.9 years, 78% were male and all-cause mortality at 1-year was 9.4%. The data were split into two mutually exclusive datasets, 70% into the training set and 30% into the testing set. The training set was utilized to generate the prediction model and the remaining 30% was employed to estimate the model's performance (Figure-1).

There were no statistically significant differences in the features between the training an test data sets (Table-1). A difference in survival outcome was absent between the two sets as well (log-rank p=0.76) (Figure-2).

3.1 Models performance

3.1.1 Cox proportional hazards model

The CPH model was built with the training set. Univariate Cox analysis was performed (Supplementary Figure-S1). Statistically significant variables in univariate analyses were taken into multivariate analysis (Supplementary Figure-S2).

Prediction accuracy was measured using c-index on the testing set. Using all 69 features to train the model, the Harell's C-index was 0.738, time-dependent IPCW was 0.741, and Brier's score was 0.007. All features were evaluated as single predictors in univariate analysis. Using the 25 significant univariate features increased model performance; Harell's C-index to 0.806, time-dependent IPCW to 0.812, and Brier's score was 0.007. Using only the significant multivariate features increased performance further; Harell's C-index to 0.849, time-dependent IPCW to 0.860, and Brier's score were 0.007. Application of Lasso or Elastic net models did not improve performance (not shown).

3.1.2 RSF Model

The RSF model parameters were trained using the grid search method for the following variables: number of trees (trees), number of variables randomly selected as candidates for splitting a node (mtry), forest average terminal node size (nodesize), number of random split points used to split a node (nsplit) and maximum depth to which a tree should be grown (nodedepth). Optimized parameters found: trees=1000, mtry = 9, nodesize=15, nsplit =10 , nodedepth =-1. The RSF out of bag (OOB) error rates are shown in Supplementary Figure-S4.

Prediction accuracy was measured on the testing set; The Harell's c-index and time-dependent IPCW of RSF reached 0.924 and 0.928, respectively, both highest among the models. Brier score was 0.006.

3.1.2 DeepSurv Model

Following normalization of the continuous features, the DeepSurv model hyperparameters were tuned using the ADAM method of Stochastic optimization applied on the training set[28]. Optimized parameters found: L2_reg=7.5, dropout = 0.527, number of hidden layers = 48, learning rate = 0.018, lr_decay = 0.011, momentum = 0.873. Model development shown in Supplementary Figure-S3.

Prediction accuracy was measured using c-index on the testing set. Using all 69 features to train the model, the Harell's C-index was 0.804, time-dependent IPCW was 0.813 and Brier's score was 0.10.

3.1 Predictive performance

The performance of survival models based on Random survival forest (RSF), DeepSurv, and models based on CPH regression were compared on both training and the testing sets. RSF performed best among the three models, Harell's c-index on training and testing sets reaching 0.953 and 0.924, followed by CPH Multivariate selected model (0.805/0.849), CPH Univariate selected model (0.828/0.806) and DeepSurv model (0.801/0.804), and finally CPH (0.826/0.738). RSF had the highest performance on each time point and overall IPCW compared to the other models (Table-2, Figure-3).

DeepSurv had the highest (worst) Brier score (0.1) compared to similar scores for the other models (0.0006-0.0007) (Table-2).

3.2 Variable importance

To determine the top 15 important variables for RSF prediction, we applied the permutation method (VIMP – "noising up" each variable in turn and its effect on prediction error). CPH features permutation importance was evaluated for comparison. There is a considerable agreement with permutation importance in 11 of the 15 selected variables by RSF and CPH (Figure-4).

Validation

The external validation cohort included 206,915 patients out of 720,932 who fulfilled the inclusion criteria (treated after 2006 and had less than 15% variables missingness), with a mean age of 68.4±13.9, 67.1%males. There were significant differences between the ACSIS training set and the validation cohort in most of the features, including age, systolic blood pressure, diastolic blood pressure, BMI, glucose, creatinine, cholesterol, hemoglobin, chronic medical conditions, smoking status, chronic medications, and in-hospital and discharge treatments (Table-3). The validation cohort had significantly higher 1-year mortality rates (13.7% compared to 9.5%, $p<0.001$) (Table-3).

The three CPH models' performance on the validation set had Harell's C-index of 0.713-0.818, 0.689-0.790 for IPCW, and 0.094-0.100 for brier score. The model using 25 variables had the highest performance among the CPH models. The DeepSurv model performance on the validation data set had 0.801 Harell's C-index and 0.832 for IPCW. Lastly, the RSF model performance on the validation data set demonstrated 0.811 for Harell's C-index, 0.844 for IPCW, and 0.093 for brier score (Table-4, Figure-5).

4. Discussion

We compared RSF, Cox, and DeepSurv models on data-driven approaches for 1-year survival prediction on national registry. We found random survival forest to significantly outperform all other models, including the classic statistical method for survival analysis on both test set and external validation cohort. Compared to the Cox-regression models, the deep neural network underperformed on the test set, but had higher performance on the validation cohort.

Most previous publications that evaluated ACS long-term mortality, used the binary classification approach (dead/alive at 1 year) without considering the time to event as in survival analysis. Sherazi et al. evaluated models for 1-year mortality post ACS classification[29], showing deep network, gradient boosting and random forest outperform the GRACE risk score (developed using the regression method). Our survival models use the added information of the time to event, and cannot be compared to GRACE. Kwon et al. demonstrated deep network-based classification outperforms the regression-based risk scores [17]. Barrett et al. similarly developed several classification models for 1 year mortality following MI using deep network, random forest among others [30]. Importantly, these were performed on relatively small cohorts and had no external validation[31]. A recent study using survival analysis, rather than classification on a large cohort from electronic health records, did not demonstrate an improved performance using RSF compared to the traditional Cox-regression approach. [32]. To the best of our knowledge, no previous study applied deep network for survival prediction post-ACS. In our study, RSF outperformed both deep learning and Cox-regression models. Deep learning model performed better than traditional CPH, however worse than multivariate features selection CPH on the test set, but outperformed both on the validation cohort.

There are several advantages to using RSF over the CPH model. First, RSF does not require prior knowledge about the relationship (i.e., linear, non-linear) of a variable over time or to choose the best equation to transform non-linear covariates. The RSF automatically explores complex interactions between variables, which usually must be manually specified in regression-type approaches. As in our current cohort of patients post ACS, we cannot assume the data satisfies the linear proportional hazards condition. Second, the overall discrimination of an RSF model is at least comparable to standard methodologies if not above, RSF has shown its ability to outperform classic CPH regressions [13,33,34]. However, RSF may miss predictors with low representation in the population, and this would go against personalized prediction.

Previous studies applying neural networks failed to demonstrate improvements beyond the classic linear CPH model[35]. Recently Katzman et al. showed deep neural networks could outperform standard survival analysis [16,36]. One of the advantages of a deep learning-based neural network is that it discerns relationships without prior feature selection. However, in our cohort, RSF outperformed DeepSurv.

Pocock et al. demonstrated Cox based models for long-term mortality post-ACS developed on national registries have good discrimination (0.79-0.82), similar to the Cox models performance in the current study on both cohorts, and lower than RSF model in the test cohort[37]. Importantly, Pocock et al models were validated on an external cohort as in the current study. ACS diagnosis, coronary vascularization, creatinine, previous cardiac disease, discharge Aldosterone and Killip grade were found important predictive features in both our and Pocock et al Cox models. In contrast with their findings, Glucose, Hemoglobin, ejection fraction, male sex, BMI and quality of life were not found as highly important predictors in our study.

The current study has several limitations. First, the outcome is all-cause mortality without accounting for competing events and non-cardiac mortality. Second, other commonly accepted predictors of survival, such as ejection fraction, serum B-type natriuretic peptides, angiography findings, etc., were not routinely obtained in both registries.

5. Conclusions

This study demonstrates that RSF survival predictions for ACS long-term mortality show higher performance patients compared to the classic statistical method. This can benefit patients by stratifying risks and guiding treatment options by providing patient individual risk-tailored treatements (e.g., prolonged DAPT, more aggressive LDL-C lowering, and complete revascularization). ~~ineffective/unnecessary treatments.~~

6. References

1. Cardiovascular disease in Europe 2016: an epidemiological update. *European Heart Journal.* 2016;37(42):3182-3183.
2. Sleeman KE, de Brito M, Etkind S, et al. The escalating global burden of serious health-related suffering: projections to 2060 by world regions, age groups, and health conditions. *The Lancet Global health.* 2019;7(7):e883-e892.
3. Collet J-P, Thiele H, Barbato E, et al. 2020 ESC Guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation: The Task Force for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation of the European Society of Cardiology (ESC). *European Heart Journal.* 2020;42(14):1289-1367.
4. Szummer K, Wallentin L, Lindhagen L, et al. Relations between implementation of new treatments and improved outcomes in patients with non-ST-elevation myocardial infarction during the last 20 years: experiences from SWEDEHEART registry 1995 to 2014. *European Heart Journal.* 2018;39(42):3766-3776.
5. Dondo TB, Hall M, Munyombwe T, et al. A nationwide causal mediation analysis of survival following ST-elevation myocardial infarction. *Heart.* 2020;106(10):765-771.
6. Szummer K, Wallentin L, Lindhagen L, et al. Improved outcomes in patients with ST-elevation myocardial infarction during the last 20 years are related to implementation of evidence-based treatments: experiences from the SWEDEHEART registry 1995-2014. *Eur Heart J.* 2017;38(41):3056-3065.
7. Smolina K, Wright FL, Rayner M, Goldacre MJ. Determinants of the decline in mortality from acute myocardial infarction in England between 2002 and 2010: linked national database study. *BMJ.* 2012;344:d8059.
8. Santos IS, Goulart AC, Brandao RM, et al. One-year Mortality after an Acute Coronary Event and its Clinical Predictors: The ERICO Study. *Arquivos brasileiros de cardiologia.* 2015;105(1):53-64.
9. George B, Seals S, Aban I. Survival analysis and regression models. *J Nucl Cardiol.* 2014;21(4):686-694.
10. Brieger DB, Chew DP, Redfern J, et al. Survival after an acute coronary syndrome: 18-month outcomes from the Australian and New Zealand SNAPSHOT ACS study. *The Medical journal of Australia.* 2015;203(9):368.
11. Breiman L. Heuristics of instability and stabilization in model selection. *The Annals of Statistics.* 1996;24(6):2350-2383.
12. Wang P, Li Y, Reddy CK. Machine Learning for Survival Analysis: A Survey. 2019;51(6 %J ACM Comput. Surv.):Article 110.
13. Hsich E, Gorodeski EZ, Blackstone EH, Ishwaran H, Lauer MS. Identifying important risk factors for survival in patient with systolic heart failure using random survival forests. *Circulation Cardiovascular quality and outcomes.* 2011;4(1):39-45.
14. Mariani L, Coradini D, Biganzoli E, et al. Prognostic factors for metachronous contralateral breast cancer: a comparison of the linear Cox regression model and its artificial neural network extension. *Breast cancer research and treatment.* 1997;44(2):167-178.

15. Lee S, Lim H. Review of statistical methods for survival analysis using genomic data. *Genomics & informatics.* 2019;17(4):e41.
16. Katzman JL, Shaham U, Cloninger A, Bates J, Jiang T, Kluger Y. DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC medical research methodology.* 2018;18(1):24.
17. Kwon JM, Jeon KH, Kim HM, et al. Deep-learning-based risk stratification for mortality of patients with acute myocardial infarction. *PloS one.* 2019;14(10):e0224502.
18. Arbel Y, Matetzky S, Gavrielov-Yusim N, et al. Temporal trends in all-cause mortality of smokers versus non-smokers hospitalized with ST-segment elevation myocardial infarction. *International journal of cardiology.* 2014;176(1):171-176.
19. Kornowski R. The ACSIS Registry and primary angioplasty following coronary bypass surgery. *Catheterization and cardiovascular interventions : official journal of the Society for Cardiac Angiography & Interventions.* 2011;78(4):537-539.
20. Segev A, Matetzky S, Danenberg H, et al. Contemporary use and outcome of percutaneous coronary interventions in patients with acute coronary syndromes: insights from the 2010 ACSIS and ACSIS-PCI surveys. *EuroIntervention.* 2012;8(4):465-469.
21. Gottlieb S, Behar S, Hod H, et al. Trends in management, hospital and long-term outcomes of elderly patients with acute myocardial infarction. *Am J Med.* 2007;120(1):90-97.
22. Madley-Dowd P, Hughes R, Tilling K, Heron J. The proportion of missing data should not be used to guide decisions on multiple imputation. *Journal of Clinical Epidemiology.* 2019;110:63-73.
23. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Annals of internal medicine.* 2015;162(1):55-63.
24. Fen Miao Y-PC, , Yuan-Ting Zhang, Chun-Yue Li. Is Random Survival Forest an Alternative to Cox Proportional Model on Predicting Cardiovascular Disease? 6th European Conference of the International Federation for Medical and Biological Engineering; 2015.
25. Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. *The Annals of Applied Statistics.* 2008;2(3):841-860, 820.
26. Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. *The Annals of Applied Statistics.* 2008:841-860.
27. Harrell FE, Jr., Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. *Jama.* 1982;247(18):2543-2546.
28. Adam: A method for stochastic optimization. *arXiv* 2014;arXiv:1412.6980.
29. Sherazi SWA, Jeong YJ, Jae MH, Bae JW, Lee JY. A machine learning-based 1-year mortality prediction model after hospital discharge for clinical patients with acute coronary syndrome. *Health informatics journal.* 2019:1460458219871780.
30. Barrett LA, Payrovnaziri SN, Bian J, He Z. Building Computational Models to Predict One-Year Mortality in ICU Patients with Acute Myocardial Infarction and Post Myocardial Infarction Syndrome. *AMIA Joint Summits on Translational Science proceedings AMIA Joint Summits on Translational Science.* 2019;2019:407-416.
31. Pieszko K, Hiczkiewicz J, Budzianowski P, et al. Predicting Long-Term Mortality after Acute Coronary Syndrome Using Machine Learning Techniques and Hematological Markers. *Disease markers.* 2019;2019:9056402.
32. Steele AJ, Denaxas SC, Shah AD, Hemingway H, Luscombe NM. Machine learning models in electronic health records can outperform conventional survival models for predicting patient mortality in coronary artery disease. *PloS one.* 2018;13(8):e0202344.
33. Mitchell DA, Kanatas A, Murphy C, Chengot P, Smith AB, Ong TK. Margins and survival in oral cancer. *The British journal of oral & maxillofacial surgery.* 2018;56(9):820-829.

34. Gilhodes J, Zemmour C, Ajana S, et al. Comparison of variable selection methods for high-dimensional survival data with competing events. *Computers in biology and medicine.* 2017;91:159-167.
35. Sargent DJ. Comparison of artificial neural networks with other statistical approaches: results from medical data sets. *Cancer.* 2001;91(8 Suppl):1636-1642.
36. Kim DW, Lee S, Kwon S, Nam W, Cha IH, Kim HJ. Deep learning-based survival prediction of oral cancer patients. *Scientific reports.* 2019;9(1):6994.
37. Pocock SJ, Huo Y, Van de Werf F, et al. Predicting two-year mortality from discharge after acute coronary syndrome: An internationally-based risk score. *European heart journal Acute cardiovascular care.* 2019;8(8):727-737.

Tables

Table-1: Entire data, train and test patients' characteristics.

	Total	Train	Imputed	Test	Imputed	Sig.
N	9270	6489		2781		
Age (years)	63.8±12.9	63.8±12.9	0.01%	63.8±12.9	0.00%	0.864
Males	7233 (78)%	5075 (78.2)%	0.00%	2158 (77.6)%	0.00%	0.515
Heart Rate (bmp)	80.0±19.3	80.0±19.4	2.40%	80.1 ±19.1	2.10%	0.872
SBP (mmHg)	141.7±27.9	141.7±27.9	2.00%	141.7±27.8	2.30%	0.925
DBP (mmHg)	80.9±16.2	80.9±16.2	2.20%	80.9±16.3	2.40%	0.858
BMI	27.8±4.4	27.8±4.4	9.50%	27.8±4.5	9.60%	0.849
Total Cholesterol (mg/dL)	176.7±42.0	176.4±42.1	15.70%	177.4±41.8	15.40%	0.288
Glucose (mmol/L)	151.7±79.3	150.8±76.9	1.60%	153.7±84.6	1.80%	0.102
Hemoglobin (g/dL)	13.7±1.8	13.7±1.8	1.40%	13.7±1.8	7.20%	0.888
Creatinine (mg/dL)	1.19±0.89	1.19±0.93	3.70%	1.18±0.80	4.20%	0.578
Past MI	2963 (32%)	2090 (32.2%)	0.30%	837 (31.4%)	0.30%	0.44

Past Angina Pectoris	3202 (34.5%)	2227 (34.3%)	0.70%	975 (35.1%)	0.50%	0.492
Past CABG	912 (9.8%)	652 (10.0%)	0.10%	260 (9.3%)	0.10%	0.301
Chronic heart failure	744 (8.0%)	523 (8.1%)	0.20%	221 (7.9%)	0.10%	0.854
Past Revascularization	3005 (32.4%)	2093 (32.3%)	0.30%	912 (32.8%)	0.30%	0.611
Past Stroke	748 (8.1%)	515 (7.9%)	0.30%	233 (8.4%)	0.10%	0.474
Family history of MI	2437 (26.3%)	1721 (26.5%)	10.10%	716 (25.7%)	10.20%	0.437
Chronic renal failure	1131 (12.2%)	780 (12.0%)	0.30%	351 (12.6%)	0.20%	0.418
PVD	743 (8.0%)	511 (7.9%)	0.20%	232 (8.3%)	0.10%	0.447
Hyperlipidemia	6750 (72.8%)	4751 (73.2%)	0.50%	1999 (71.9%	0.50%	0.185
Diabetes Mellitus	3484 (37.6%)	2420 (37.3%)	0.30%	1064 (38.6%)	0.30%	0.379

Hypertension	5865 (63.3%)	4127 (63.6%)	0.30%	1738 (62.5%)	0.30%	0.312
Killip score at admission			1.80%		1.70%	0.443
	I	5655 (87.1%)		2420 (87.0%)		
	II	495 (7.6%)		196 (7.0%)		
	III	245 (3.8%)		119 (4.3%)		
	IV	94 (1.4%)		46 (1.7%)		
Chronic beta blockers	3431 (37.0%)	2389 (36.8%)	3.10%	1042 (37.5%)	3.30%	0.551
Chronic calcium blockers	1857 (20.0%)	1288 (19.8%)	3.80%	569 (20.5%)	3.80%	0.5
Chronic nitrates	711 (7.7%)	510 (7.9%)	4.10%	201 (7.2%)	4.40%	0.295
Chronic Aspirin	4513 (48.7%)	3162 (48.7%)	1.70%	1351 (48.6%)	1.50%	0.895
Chronic Anticoagulation	384 (4.1%)	279 (4.3%)	3.10%	105 (3.8%)	3.30%	0.246
Chronic Statins	5143 (55.5%)	3580 (55.2%)	5.50%	1563 (56.2%)	6.40%	0.359

Chronic hypoglycemic drugs	2219 (23.9%)	1546 (23.8%)	0.40%	672 (24.2%)	0.60%	0.698
Chronic insulin	828 (8.9%)	575 (8.9%)		253 (9.1%)		0.715
Chronic ACE-I/ARBB	3789 (40.9%)	2659 (41.0%)	0.40%	1130 (40.6%)	0.20%	0.757
Chronic Smoking	3554 (38.3%)	2505 (38.6%)	0.60%	1049 (37.7%)	1.00%	0.423
Previous Smoking	2055 (22.2%)	1417 (21.8%)	0.60%	638 (22.9%)	1.00%	0.241
Normal Sinus Rythem	8102 (87.4%)	5658 (87.2%)	0.00%	2444 (87.9%)	0.00%	0.36
Diagnosis			0%		0%	0.442
	STEMI	2775 (42.8%)		1164 (41.9%)		
	nSTEMI	2743 (42.3%)		1215 (43.7%)		
	UA	971 (15.0%)		402 (14.5%)		
Anterior MI	2377 (25.6%)	1654 (25.5%)	0%	723 (26.0%)	0%	0.607

Hospitalization						
Aspirin	9003 (97.1%)	6303 (97.1%)	0%	2700 (97.1%)	0%	0.903
Theophylline inhibitors	6908 (74.5%)	4812 (74.2%)	0.40%	2096 (75.4%)	0.50%	0.22
IIb/IIIa	2284 (24.6%)	1609 (24.8%)	0.20%	675 (24.3%)	0.30%	0.592
Beta Blockers	7599 (82.0%)	5333 (82.2%)	3.10%	2266 (81.5%)	3.60%	0.419
Insulin	1528 (16.%)	1039 (16.0%)	13.90%	489 (17.6%)	13.90%	0.062
Ace Inhibitors / ARB	7195 (77.6%)	5045 (77.7%)	0.30%	2150 (77.3%)	0.20%	0.644
Heparin	6190 (66.8%)	4317 (66.5%)	0.00%	1873 (67.3%)	0%	0.441
LMWH	4544 (49.0%)	3173 (48.9%)	0.00%	1371 (49.3%)	0%	0.724
Anti-coagulation	435 (4.7%)	310 (4.8%)	1.00%	125 (4.5%)	1%	0.556
Hypoglycemic drugs	1532 (16.5%)	1056 (16.3%)	0.30%	476 (17.1%)	0.30%	0.317

Aldosterone	527 (5.7%)	370 (5.7%)	19.90%	157 (5.6%)	19.20%	0.914
Angiography	8227 (88.7%)	5764 (88.8%)	0%	2463 (88.6%)	0%	0.715
PCI	6415 (69.2%)	4482 (69.1%)	0%	1933 (69.5%)	0%	0.676
Cardiac rehabilitation	3887 (41.9%)	2732 (42.1%)	12.70%	1155 (41.5%)	13.10%	0.61
CABG	515 (5.6%)	350 (5.4%)	0%	165 (5.9%)	0%	0.299
Followup Angiography	526 (5.7%)	370 (5.7%)	12.40%	156 (5.6%)	11.80%	0.971
Followup PCI	450 (4.9%)	322 (5.0%)	12.90%	128 (4.6%)	12.30%	0.46
Followup CABG	559 (6.0%)	382 (5.9%)	12.90%	177 (6.4%)	12.10%	0.376
CPR/DCS	265 (2.9%)	177 (2.7%)	0%	88 (3.2%)	0%	0.248
Discharge drugs						
Aspirin			1.80%		1.90%	
Theophylline inhibitors			2.10%		2.10%	
Insulin	1031 (11.1%)	715 (11.0%)	15.20%	316 (11.4%)	14.90%	0.629
Ace-Inhibitors/ARB	7000 (75.5%)	4894 (75.4%)	0.90%	2106 (75.7%)	0.80%	0.752

Statin	8728 (94.2%)	6105 (94.1%)	2.60%	2623 (94.3%)	2.40%	0.657
Hypoglycemic drugs	1852 (20.0%)	1279 (19.7%)	0.90%	573 (20.6%)	0.80%	0.324
Aldosterone	520 (5.6%)	358 (5.5%)	21.10%	162 (5.8%)	20.40%	0.555
Beta Blockers	7565 (81.6%)	5264 (81.1%)		2301 (82.7%)		0.065
1 Year Mortality	874 (9.4%)	616 (9.5%)	0%	258 (9.3%)	0%)	0.745

Table-2: **Predictive performance of all models for 1-year mortality.**

Model	Harell's C-Index (Train)	Harell's C-Index (Test)	IPCW (Test)	Brier (Test)
RSF	0.953	0.924	0.928	0.006
DeepSurv	0.801	0.804	0.813	0.007
CPH	0.826	0.738	0.741	0.007
CPH-Univariate selected features	0.828	0.806	0.812	0.007
CPH-Multivariate selected features	0.805	0.849	0.860	0.1

Table-3: ACSIS training set and MINAP validation set, patients' characteristics.

	ACSIS (Train)	Imputed	MINAP (Validation)	Imputed	Sig.
N	6,489		206,915		
Age (years)	63.8±12.9	0.01%	68.4±13.9	0.02%	**<0.001**
Males	5075 (78.2)%	0.00%	138,829 (67.1)%	0.01%	**<0.001**
Heart Rate (bmp)	80.0±19.4	2.40%	80.4 ±21.6	4.32%	0.12
SBP (mmHg)	141.7±27.9	2.00%	138.4±28.0	4.27%	**<0.001**
DBP (mmHg)	80.9±16.2	2.20%	79.07±16.0	4.27%	**<0.001**
BMI	27.8±4.4	9.50%	28.1±8.4	47.0%	**<0.001**
Total Cholesterol (mg/dL)	176.4±42.1	15.70%	188.4±53.9	30.9%	**<0.001**
Glucose (mmol/L)	150.8±76.9	1.60%	148.4±84.1	12.1%	**0.026**
Hemoglobin (g/dL)	13.7±1.8	1.40%	13.4±2.0	8.1%	**<0.001**
Creatinine (mg/dL)	1.19±0.93	3.70%	1.14±0.66	1.6%	**<0.001**
Past MI	2090 (32.2%)	0.30%	43248 (21.0%)	0.25%	**<0.001**

Past Angina Pectoris	2227 (34.3%)	0.70%	51042 (24.8%)	0.39%	**<0.001**
Past CABG	652 (10.0%)	0.10%	12912 (6.2%)	0.1%	**<0.001**
Chronic heart failure	523 (8.1%)	0.20%	9437 (4.6%)	0.3%	**<0.001**
Past Revascularization	2093 (32.3%)	0.30%	19968 (9.7%)	0.21%	**<0.001**
Past Stroke	515 (7.9%)	0.30%	16906 (8.2%)	0.22%	0.48
Family history of MI	1721 (26.5%)	10.10%	61211 (33.0%)	10.25%	**<0.001**
Chronic renal failure	780 (12.0%)	0.30%	11438 (5.6%)	0.41%	**<0.001**
PVD	511 (7.9%)	0.20%	8609 (4.2%)	1.83%	**<0.001**
Hyperlipidemia	4751 (73.2%)	0.50%	71611 (35.1%)	1.40%	**<0.001**
Diabetes Mellitus	2420 (37.3%)	0.30%	41860 (20.3%)	0.15%	**<0.001**

Hypertension	4127 (63.6%)	0.30%	103319 (50.1%)	0.30%	**<0.001**
Killip score at admission		1.80%		62.2%	**<0.001**
	5655 (87.1%)		62374 (79.9%)		
	495 (7.6%)		10468 (13.4%)		
	245 (3.8%)		4081 (5.2%)		
	94 (1.4%)		1181 (1.5%)		
Chronic beta blockers	2389 (36.8%)	3.10%	56960 (27.8%)	0.92%	**<0.001**
Chronic calcium blockers	1288 (19.8%)	3.80%	32197 (15.6%)	0.17%	**<0.001**
Chronic nitrates	510 (7.9%)	4.10%	49251 (23.8%)	0.16%	**<0.001**
Chronic Aspirin	3162 (48.7%)	1.70%	49019 (24.1%)	1.84%	**<0.001**

Chronic Anticoagulation	279 (4.3%)	3.10%	10093 (4.9%)	0.1%	**0.034**
Chronic Statins	3580 (55.2%)	5.50%	84971 (41.4%)	0.88%	**<0.001**
Chronic hypoglycemic drugs	1546 (23.8%)	0.40%	24649 (12.0%)	0.45%	**<0.001**
Chronic insulin	575 (8.9%)		41226 (20.0%)	0.45%	**<0.001**
Chronic ACE-I/ARBB	2659 (41.0%)	0.40%	74410 (36.3%)	1.04%	**<0.001**
Chronic Smoking	2505 (38.6%)	0.60%	59093 (29.1%)	1.91%	**<0.001**
Previous Smoking	1417 (21.8%)	0.60%	69324 (34.2%)	1.91%	**<0.001**
Normal Sinus Rythem	5658 (87.2%)	0.00%	12992 (39.3%)	84.0%	**<0.001**
Diagnosis		0%		0%	**<0.001**
STEMI	2775 (42.8%)		81657 (39.5%)		
nSTEMI	2743 (42.3%)		120081 (58.0%)		

UA	971 (15.0%)		5177 (2.5%)		
Anterior MI	1654 (25.5%)	0%	42883 (34.8%)	40.4%	**<0.001**
Hospitalization					
Aspirin	6303 (97.1%)	0%	147946 (72.8%)	1.84%	**<0.001**
Theophylline inhibitors	4812 (74.2%)	0.40%	190326 (92.1%)	0.09%	**<0.001**
IIb/IIIa	1609 (24.8%)	0.20%	24824 (12.0%)	0.2%	**<0.001**
Beta Blockers	5333 (82.2%)	3.10%	151617 (76.8%)	4.65%	**<0.001**
Insulin	1039 (16.0%)	13.90%	21168 (10.5%)	2.70%	**<0.001**
Ace Inhibitors / ARB	5045 (77.7%)	0.30%	127582 (61.8%)	0.17%	**<0.001**
Heparin	4317 (66.5%)	0.00%	54641 (26.5%)	0.47%	**<0.001**

LMWH	3173 (48.9%)	0.00%	105729 (51.2%)	0.24%	**<0.001**
Anti-coagulation	310 (4.8%)	1.00%	10093 (4.9%)	0.1%	0.72
Hypoglycemic drugs	1056 (16.3%)	0.30%	14938 (7.4%)	2.7%	**<0.001**
Aldosterone	370 (5.7%)	19.90%	12452 (6.3%)	3.93%	0.07
Angiography	5764 (88.8%)	0%	74327 (87.5%)	58.9%	**0.002**
PCI	4482 (69.1%)	0%	61460 (72.4%)	58.9%	**<0.001**
Cardiac rehabilitation	2732 (42.1%)	12.70%	161912 (81.6%)	4.11%	**<0.001**
CABG	350 (5.4%)	0%	5197 (2.7%)	8.61%	**<0.001**
Followup Angiography	370 (5.7%)	12.40%	4658 (2.3%)	1.46%	**<0.001**
Followup PCI	322 (5.0%)	12.90%	3132 (1.7%)	8.61%	**<0.001**
Followup CABG	382 (5.9%)	12.90%	4208 (2.2%)	8.61%	**<0.001**

CPR/DCS	177 (2.7%)	0%	9468 (28.7%)	84.04%	**<0.001**
Discharge drugs					
Aspirin	6215 (95.8)	1.80%	167413 (81.4%)	0.58%	**<0.001**
Theophylline inhibitors	4134 (63.7)	2.10%	150813 (75.1%)	2.92%	**<0.001**
Insulin	715 (11.0%)	15.20%	12804 (6.4%)	2.65%	**<0.001**
Ace-Inhibitors/ARB	4894 (75.4%)	0.90%	150647 (73.5%)	0.93%	**0.001**
Statin	6105 (94.1%)	2.60%	166859 (81.2%)	0.66%	**<0.001**
Hypoglycemic drugs	1279 (19.7%)	0.90%	19806 (9.8%)	2.65%	**<0.001**
Aldosterone	358 (5.5%)	21.10%	11934 (6.0%)	3.68%	0.122
Beta Blockers	5264 (81.1%)		148313 (72.3%)	0.81%	**<0.001**
1 Year Mortality	616 (9.5%)	0%	28311 (13.7%)	0%	**<0.001**

Table-4: **Predictive performance of all models for 1-year mortality in the validation cohort (MINAP**

Model	IPCW	Harell's C-index	Brier
CPH	0.818	0.781	0.094
CPH-Univariate selected features	0.826	0.790	0.095
CPH-Multivariate selected features	0.713	0.689	0.103
RSF	0.844	0.811	0.093
Deep learning	0.832	0.801	-

Figures:

Figure-1: Patients flowchart

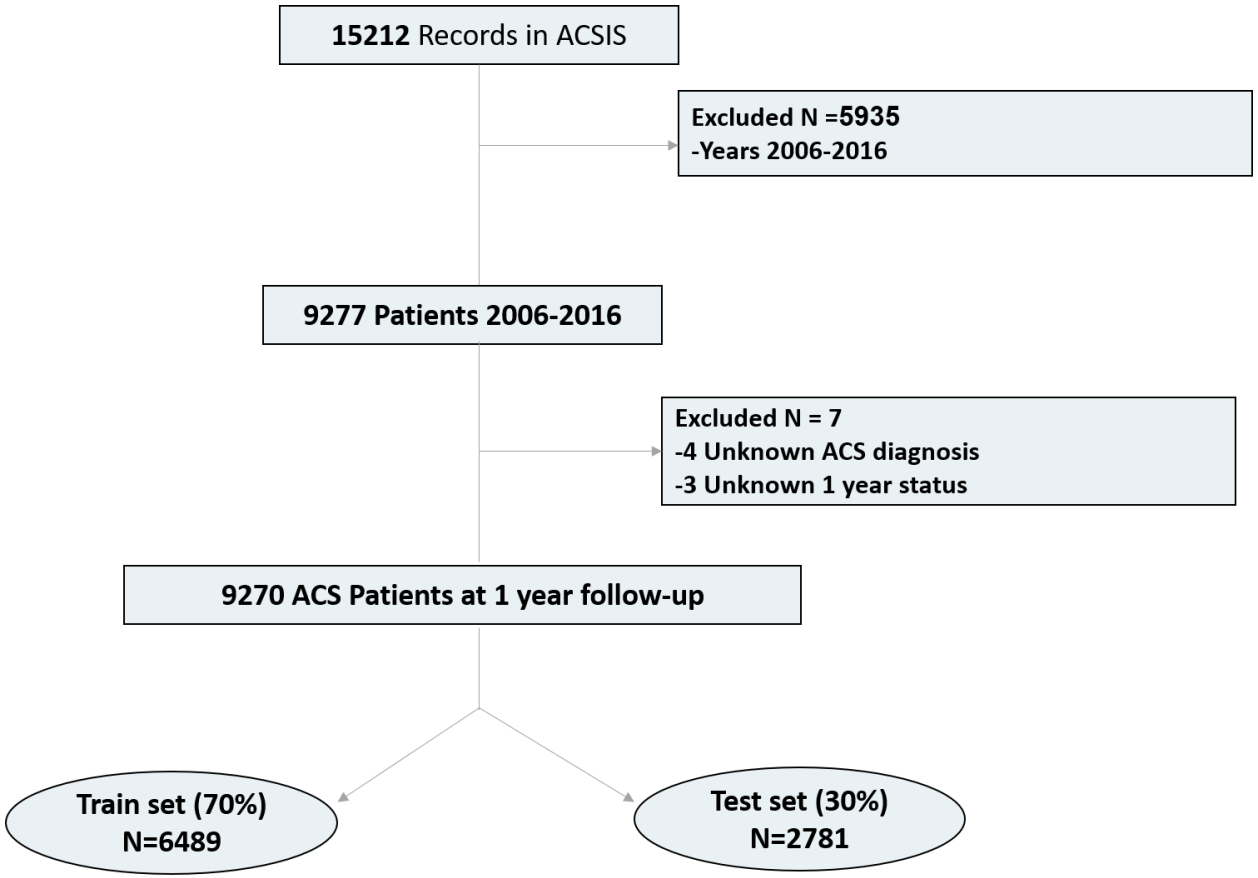


Figure-2: Kaplan-Meier curves of training and testing datasets

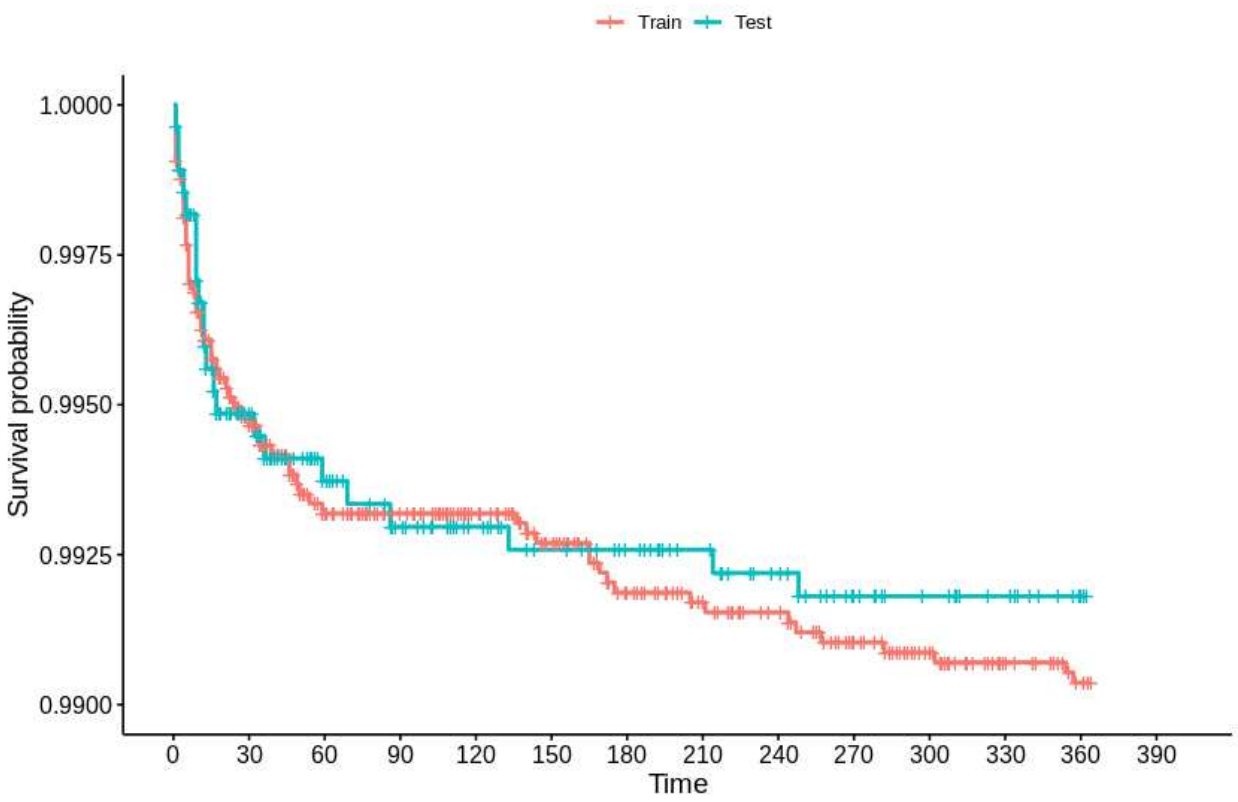
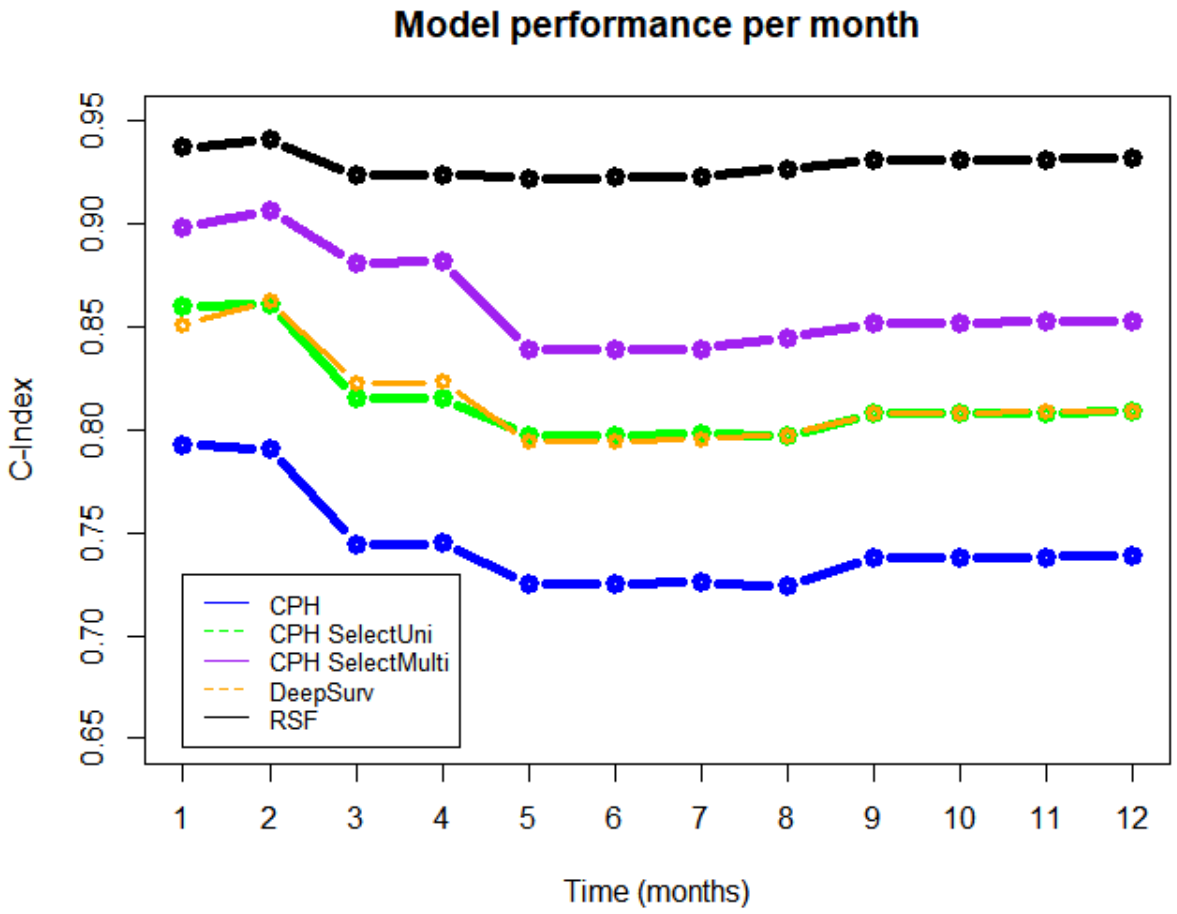


Figure-3: Testing set performance -Time dependent AUC.



RSF had the highest predictive performance on the test set compared to the other models on any time point (1-12 months post ACS).

CPH SelectUni – univariate selected features, CPH SelectMulti – multivariate selected features

Figure-4: Features importance in RSF compared to CPH

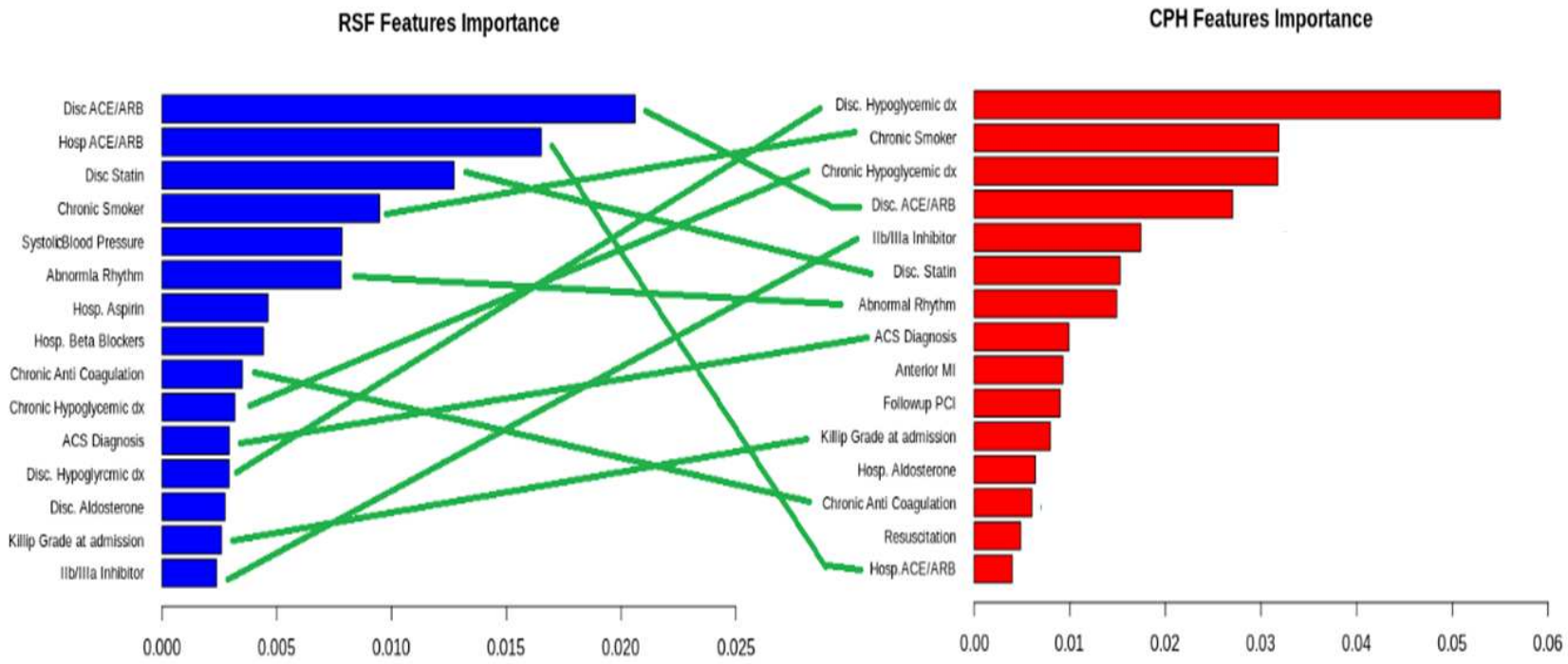
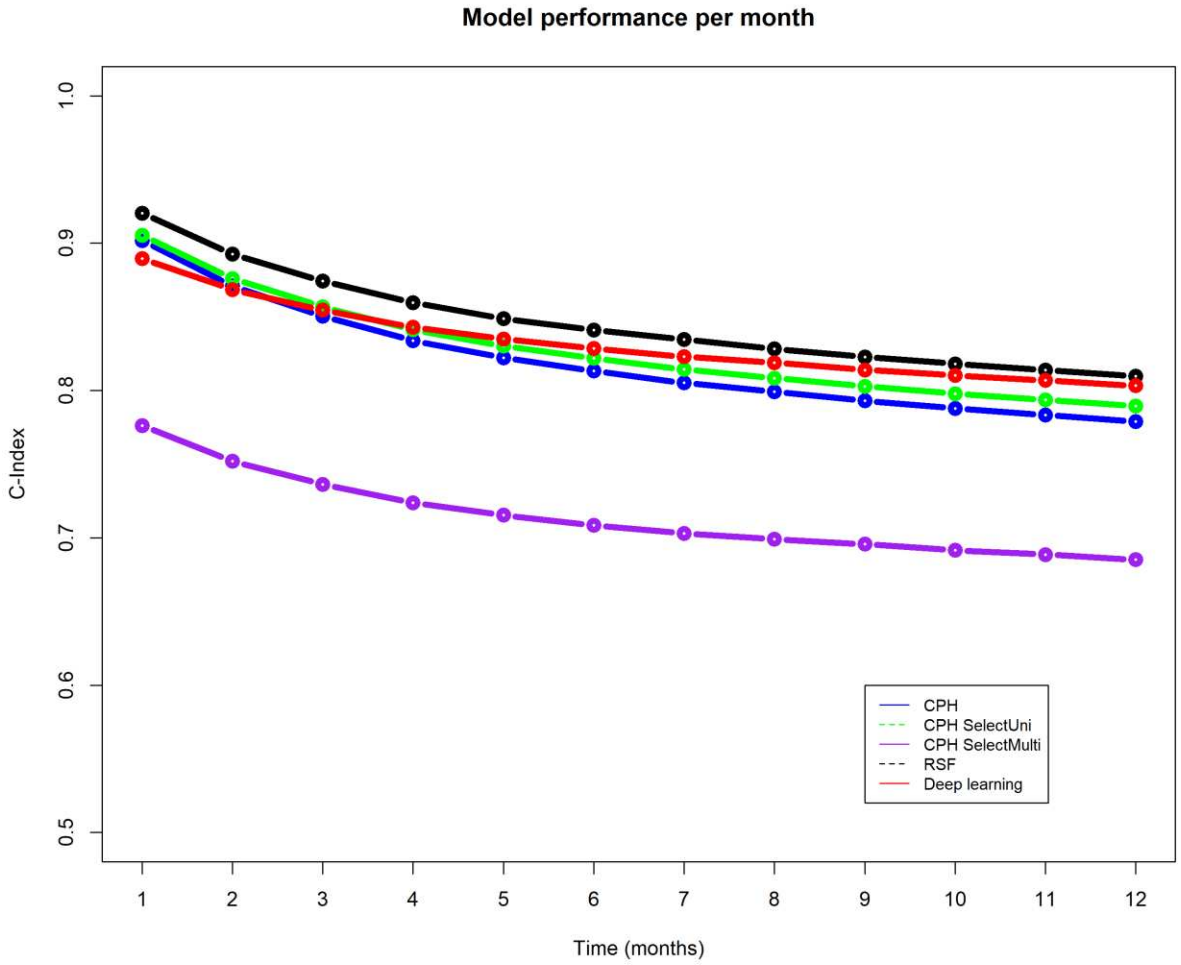


Figure-5: Validation set performance -Time-dependent AUC. RSF had the highest predictive performance on the MINAP cohort compared to the other models at any time point (1-12 months post-ACS).



CPH SelectUni – univariate selected features, CPH SelectMulti – multivariate selected features