Title
Analyzing 30-Day Readmission Rate for Heart Failure Using Different Predictive Models.

Authors
Mahajan, Satish
Burman, Prabir
Hogarth, Michael

Publication Date
2016

DOI
10.3233/978-1-61499-658-3-143

Peer reviewed

Analyzing 30-day Readmission Rate for Heart Failure Using Different Predicitve Models

Satish MAHAJAN, PhD, MStat, MEng, RN[a,b,c1], Prabir BURMAN, PhD[d], Michael HOGARTH, MD[b]

[a]*Betty Irene Moore School of Nursing, University of California, Davis, CA, USA*

[b]*School of Medicine, University of California, Davis, CA, USA*

[c]*Nursing Informatics Service, Veterans Affairs Palo Alto Health Care System, Palo Alto, CA, USA*

[d]*Department of Statistics, University of California, Davis, CA, USA*

Abstract. The Center for Medicare and Medical Services in the United States compares hospital's readmission performance to the facilities across the nation using a 30-day window from the hospital discharge. Heart Failure (HF) is one of the conditions included in the comparison, as it is the most frequent and the most expensive diagnosis for hospitalization. If risk stratification for readmission of HF patients could be carried out at the time of discharge from the index hospitalization, corresponding appropriate post-discharge interventions could be arranged. We, therefore, sought to compare two different risk prediction models using 48 clinical predictors from electronic health records data of 1037 HF patients from one hospital. We used logistic regression and random forest as methods of analyses and found that logistic regression with bagging approach produced better predictive results (C-Statistics: 0.65) when compared to random forest (C-Statistics: 0.61).

Keywords. Predictive models, readmission, heart failure, Electronic Health Records, Logistic Regression, Random Forest

1. Introduction

A vast amount of literature over the past 20 years as well as projections into the future indicate that heart failure incidence and prevalence will continue to rise in the United States [1][2]. This situation has quality and cost implications for both patients and health care organizations. Hospitals concerned with disease burden, cost containment, and worries about penalty for high readmission rates are critically looking at 30-day readmissions after initial hospitalization for heart failure (HF). We intend to examine predictivity of the risk factors for such readmissions.

---

[1] Corresponding Author: Satish Mahajan. University of California, Davis, Betty Irene Moore School of Nursing, Education Building, 4610 X Street, Sacramento, CA 95817, USA.

A systematic review of the previous models indicates that researchers have suggested various administrative, clinical, and psychosocial predictors of HF readmissions [3][4]. We intend to apply the same set of clinical predictors to two different predictive models—Logistic Regression with Bagging and Random Forest—to understand the effort involved in constructing and running them and to see which model is better at predicting readmission risk in our dataset. Clinical predictors are measures of physiologic change and typically consist of vital signs such as heart rate, blood pressure, some HF specific laboratory values, and comorbidities. We also intend to understand the role of tuning parameters, if any, for the models and review predictor importance suggested by the models.

2. Methods

We employed observational retrospective cohort study design using health care data from Electronic Health Records (EHR) at Veterans Affairs Palo Alto Health Care System, Palo Alto, CA, USA to derive and validate the models. This system represents one of the oldest electronic data sources of patient information in the United States. Institutional Review Board of the hospital approved this project and protocols to protect patient-specific identifiable information were followed.

2.1. Definitions and Data Set

The patient cohort consisted of outcome variable with two classes of patients: Class 1 representing the patients who were readmitted for any cause within 30 days of their last hospitalization and the other, Class 0, representing the patients who were either not readmitted within 30 days of their last hospitalization or readmitted after the 30-day window. The 30-day window was chosen based on the prior empirical studies that indicated that the probability of readmission of patients with heart failure is highest during the first 30-day period from the earlier hospitalization [5] and also on the Center for Medicare Services (CMS) guidance based on the Affordable Care Act of 2010 [6].

We extracted six years worth of clinical data using International Classification of Diseases version 9 – Clinical Modification (ICD-9-CM) codes for heart failure. If patients had multiple episodes of 30-day readmissions, only the last readmission was considered. In addition, if the patient was readmitted more than once during the 30-day period from his/her last discharge, the last episode within the 30-day range was considered. We also made sure that the same patients were not repeated in the non-30-day set if, in case, they had had other episodes of admission that were not 30-day readmissions. These rules made sure that (i) we did not repeat any patient within the 30-day readmission set, and (ii) we did not repeat any patient across 30-day and non-30-day readmission sets. We, thus, had a statistically independent and mutually exclusive sample of patients across the two classes of the cohort.

The raw dataset extractions were carried out from database system running on Linux servers to Windows server running R language environment. In all, about 25 million records and 10 GB of data were manipulated to arrive at the dataset representing 48 predictors with 1037 patient readmissions; 180 of which were within 30 days and 857 represented the non-30-day class. This indicated 180/1037 = 17.36% of 30-day readmission rate. If the repeated readmissions for the same patient were counted, we found the readmission rate of 260/1037 = 23.96%. This rate coincided with the industry reported rate of 24% by the other hospitals including United States government's Medicare website. Table 1 represents the predictor set for the classification models.

Table 1. Predictor set used for classification models.

Predictor Subset	Predictor Name	Data Type
Vitals	Heart Rate	Numeric
	Respiratory Rate	Numeric
	Systolic Blood Pressure	Numeric
Laboratories	Glucose	Numeric
	Urea Nitrogen	Numeric
	Sodium	Numeric
	Potassium	Numeric
	Albumin	Numeric
	Hemoglobin	Numeric
	Hematocrit	Numeric
	B-Natriuretic Peptide (BNP)	Numeric
Comorbidities (Only Disease Groups reported here)	Diabetes Mellitus	Categorical
	Coronary Artery Disease	Categorical
	Ischemic Heart Disease	Categorical
	MI	Categorical
	Valvular Heart Disease	Categorical
	Vascular/Circulatory Disease	Categorical
	Aortic Stenosis	Categorical
	Arrhythmias	Categorical
	Idiopathic Cardiomyopathy	Categorical
	Prior Cardiac Surgery	Categorical
	ICD placement during hospitalization	Categorical
	History of percutaneous coronary intervention	Categorical
	Mechanical ventilation during admission	Categorical
	Renal Disease or ESRD or Dialysis	Categorical
	Chronic Lung Disease/COPD/Asthma	Categorical
	Cerebrovascular Accident /TIA	Categorical
	Metastatic Cancer/Acute Leukemia/severe hematological disorder	Categorical
	Liver Disease	Categorical

2.2. Logistic Regression

Logistic regression is a method of modeling a binary categorical response variable. It uses a generalized linear model with a logit link function, estimating its beta parameters via Weighted Least Squares regression. Given observations of the predictors, Logistic Regression estimates the probability that the response falls into a particular category. When used in classification as a binomial response, a probability greater than 0.5 indicates classification in one category, and that less than or equal to 0.5 indicates classification in the other category (Readmission or No-readmission).

The Logistic Regression implemented via **glm()** and **step()** methods in stats package in R was wrapped with bagging layer in this project [7]. We examined final model selection using both Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) for the dataset.

2.3. Random Forest

Random Forest method stems from general decision tree methods of Classification and Regression Tree (CART) family. These methods involve segmentation of the predictor space into a number of simple regions with recursive binary splitting and then use of the mean or the mode of the observations in the region for making a prediction. The basic Classification Tree method results are sensitive to predictor measurements and thus introduce high variance in the prediction decision and hence bootstrap aggregation or bagging procedure was suggested [8]. Further improvements of random selection of certain number of predictors were suggested to reduce the influence of highly correlated predictors to create the resulting method of Random Forest [9].

There are 3 main parameters that control the behavior of the Random Forest algorithm: (i) *ntree* parameter decides the number of trees to grow in a bootstrapped sample. (ii) *mtry* parameter determines the number of predictors to select randomly at each tree split. There are two methods (**tuneRF()** and **rfcv()**) to estimate this parameter. We tried both methods to select optimal value of *mtry* parameter. (iii) *nodesize* parameter decides the depth of each tree that is grown during bagging. It indicates the minimum number of terminal nodes in the tree. The resultant Random Forest model was finally run with all the tuned parameters.

3. Results

We ran Logistic Regression with bagging for 100 iterations for both AIC and BIC based stepwise selections. The prediction error and Class 0 error performance with the AIC based predictor selection was slightly better than that of the BIC selection. However, the Class 1 error for the AIC criterion was 18.7% lower than that of the BIC criterion. The AIC criterion, by its definition, selected many more predictors (in the range of 20-24 out of 48) as compared to the BIC case (8-12 out of 48). The optimal model using this method provided C-Statistics of 0.65 and 0.62 for the AIC and BIC criteria respectively.

For the Random Forest method, the Out Of Bag (OOB) errors in each class varied significantly with the tree size. The Class 1 produced bigger errors, as its class size was considerably smaller than Class 0 and *ntree* = 300 parameter was used for the final run. Our simulations for *mtry* parameter showed that *mtry* = 24 setting had the lowest Cross Validation (CV) error whereas 3, 6, 12 predictors produced about the same CV error. Our simulations with various values for *nodesize* parameter showed *nodesize* = 10 provided minimum class errors. The optimal model using this method provided C-Statistics of 0.61.

4. Discussion

We compared parametric model approach using Logistic Regression with bagging to a non-parametric model of Random Forest as a statistical technique applied to the same dataset. With Logistic Regression approach, the predictors related to comorbidities had positive coefficients and indicated increased probability of 30-day readmission with their presence. The other significant predictors with negative coefficients were blood pressure and hemoglobin indicating that as their values go down, patient's chance of 30-day readmission went up. These findings coincided with the empirical observations in our patient cohort. For the Random Forest approach, Gini Index based method seemed to favor continuous variables whereas the classification error based prediction favored categorical variables representing comorbidities. One lab test (BNP) and vital signs appeared to be important predictors of 30-day readmission according to Random Forest method.

The direct comparison of the methods indicated that Stepwise AIC Logistic Regression model provided the highest predictivity as measured by C-Statistics. Tuning of Random Forrest algorithm became incrementally time consuming and resource intensive as the number of simulations that had to be run to find optimal values for all the parameters increased exponentially with the grid of trial values for each parameter.

This work used dataset derived from the EHR system of one hospital and hence it requires additional validation studies. On the other hand, it demonstrated the feasibility of using incrementally better methods of risk prediction in the HF patient cohort. It also provided a baseline for exploring use of additional predictor domains and more advanced algorithmic techniques.

5. Acknowledgments

This project was carried out under the IRB jointly approved by VA Palo Alto Health Care System (VAPAHCS) and Stanford University (eProtocol number: 12015). Robert King, MS, Health Science Specialist helped in the data extraction. We also want to thank Dr. Paul Heidenreich, Cardiology Service, VAPAHCS and Professor of Medicine, Stanford University for his comments on the initial manuscript.

Reference

[1] M.W. Rich, Quality improvement in heart failure: A simple solution to the beta-blocker problem, *Journal of Cardiac Failure* **10(3)** (2004), 225-227.
[2] N. Fida & I.L. Pina, Trends in heart failure hospitalizations, *Curr Heart Fail Rep* **9(4)** (2012), 346-353.
[3] J.S. Ross, G.K. Mulvey, B. Stauffer, V. Patlolla, S.M. Bernheim, P.S. Keenan, & H.M. Krumholz, Statistical models and patient predictors of readmission for heart failure: A systematic review, *Arch Intern Med.* **168(13)** (2008), 1371-1386.
[4] A.M. Hersh, F.A. Masoudi, & L.A. Allen, Postdischarge environment following heart failure hospitalization: expanding the view of hospital readmission, *J Am Heart Assoc* **2(2)** (2013), e000116.
[5] H.M. Krumholz, L. Horwitz, C. Partovian, Z. Lin, J. Herrin, J. Grady, . . . E. Drye, *Hospital-Wide (All-Condition) 30-Day Risk-Standardized Readmission Measure*, New Haven, CT: Yale New Haven Health Services Corporation/Center, 2008.
[6] Center for Medicare and Medicaid Services. (2013). *Readmissions Reduction Program*. Bethesda, Maryland. Retrieved from http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/Readmissions-Reduction-Program.html.
[7] R Development Core Team. (2006). *R: A language and environment for statistical computing [Computer software manual]*. Vienna, Austria. Retrieved from http://www.R-project.org.
[8] L. Breiman, Bagging predictors, *Machine Learning*, **24** (1996), 123–140.
[9] L. Breiman, Random forests, *Machine Learning*, **45** (2001), 5–32.