A Prediction Model for Disease-Specific 30-Day Readmission Following Hospital Discharge

By

Dara L. E. Mize

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

in

Biomedical Informatics

May 11, 2018

Nashville, Tennessee

Approved:

Mia Levy, M.D., Ph.D.

Shubhada Jagasia, M.D., M.M.H.C.

Colin Walsh, M.D., M.A.

To Mom and Dad, Mark, Catherine, and Elizabeth. Thank you for your love and support.

ACKNOWLEDGEMENTS

I would like to thank my committee for their mentorship through this process. Colin Walsh was a truly dedicated mentor and I could not have been successful without his guidance. Mia Levy has been a thoughtful mentor to me through this process, both personally and professionally and I look forward to continuing to work with her. Shubhada Jagasia has been an inspirational leader and mentor both personally and professionally for many years. I am grateful for her ongoing support. Kevin Johnson inspired me to consider a career change in a single lecture several years ago. For his inspiration that day and many days since then, I am thankful. I am grateful to Al Powers for allowing me the opportunity to pursue an Endocrinology fellowship at Vanderbilt many years ago and for all of his support and sound advice since then. I would like to thank Cindy Gadd, Gretchen Jackson, Rischelle Jenkins and other DBMI faculty for allowing me to have this educational opportunity. I am grateful to my all of my DBMI colleagues and fellow students for creating such an engaging and collegial atmosphere. I would like to acknowledge Matt Lenert and Johnathan Wanderer for allowing me to use their LACE data in my work. I want to express my gratitude to Neal Patel and my other colleagues in HealthIT who continue to teach me how to create informatics solutions which can translate into clinical practice. Howard Baum has been a mentor and sounding board in many aspects of my life both personally and professionally, including during the early stages of this work. I am thankful for him and his sense of humor. I am indebted to my patients who inspire me every day to try to improve healthcare for all of us.

My parents, Daniel and Deloris Eckerle, have spent their lives sacrificing so that I can follow my dreams. That has never been more true than the last 4 years. I could spend my life

trying to repay them but will fall short. I promise to do the same for my daughters. My husband, Mark Mize, makes all of this possible. He is a true partner and I am so thankful for his love, support and friendship. My two daughters, Catherine and Elizabeth, are my joy. They have felt the sacrifice of this process but have also demonstrated grace, patience and understanding beyond their years. I love you all and hope to always make you proud.

This work would not have been possible without funding from NIH T32 DK007061 and NLM 5T15LM007450.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Value-Based Healthcare

For decades, policy makers have been developing legislation in attempt to bend the healthcare cost curve to improve value while maintaining quality healthcare delivery. The National Health Expenditure has grown steadily since the 1960s reaching 17.9% of the Gross Domestic Product (GDP) in 2016[1]. The Health Maintenance Act of 1973[2] promoted the formation of Health Maintenance Organizations (HMOs) and other managed care organizations which aimed to reduce healthcare costs through various techniques including capitation agreements for physicians, providing a set of services for fixed payment, restricting access outside of a preferred network and cost-sharing features such as copayments and coinsurance. The number of managed care organizations continued to grow in the 1980s and 1990s[3].

However, the quality of healthcare delivery in the United States began receiving intense scrutiny following the release of *To err is human: Building a safe health system* by the Institute of Medicine (IOM) in 1999[4]. In this report, the authors outline a healthcare system marred by preventable medical errors and patient safety concerns. In their follow-up series, *Crossing the Quality Chasm*, the IOM raised awareness of the inherent disincentive of current healthcare payment policies to improve the value of healthcare delivery by reducing cost[5]. In light of these two publications, policies supporting value based payment systems or "pay for performance" initiatives began to gain traction as a mechanism to support high quality care while also reducing cost[6,7].

The passage of the Affordable Care Act (ACA) in 2010 resulted in healthcare reform with the goal of reducing healthcare costs while enhancing the quality of care[8]. Underlying this directive is a major transformation in the way healthcare is administered. There is an ongoing transition from the traditional fee-for-service model of healthcare delivery in the United States to that of bundled payments and Accountable Care Organizations (ACOs). With this, the reimbursement structure is changing such that the burden of providing high-quality, cost-effective care falls on healthcare systems. ACOs which meet established quality metrics become eligible for "incentive payments.[8]" Additionally, penalties in the form of reduction in payment will be imposed on systems which fail to meet accepted quality standards. One such example, the Hospital Readmissions Reduction Program (HRRP), permits Centers for Medicare and Medicaid Services (CMS) to reduce reimbursement to hospitals with excess 30-day hospital unplanned readmissions[8].

The HRRP, established as a provision of the ACA, reflects the 2008 recommendations of the Medicare Payment Advisory Commission (MedPAC) in their report to Congress[9]. In this report, MedPAC recognized that the current Medicare fee-for-service (FFS) payment model rewarded health care systems for increased volume. They recommended a novel payment structure which would hold providers accountable for the delivery of high-quality care and incentivize providers to work together. The proposed changes included public reporting of hospital readmission rates and a readmission reduction program, whereby Medicare would reduce payments to hospitals with excess readmissions. With a 30-day readmission rate of approximately 20% (17.6%[9], 19.6%[10]) among Medicare beneficiaries[9,10], these changes were proposed to save Medicare as much as $12 billion on preventable readmissions while improving quality of care delivered[9].

Since implementation of HRRP, penalties imposed on hospitals with excess readmissions increased from $290 million in 2013 to $528 million in 2017[11]. As additional medical conditions are added to the evaluation of the readmission penalty, the average hospital penalty and percent of hospitals receiving the maximum penalty continues to rise (table 1). While these regulations resulted in significant savings for Medicare, the loss of revenue for hospitals has prompted them to focus attention on ways to reduce hospital readmissions and improve safety during transitions of care.

Table 1. Financial Summary from First Five Years of Hospital Readmission Reduction Program (Reproduced with permission[11])

Year penalties apply	**FY 2013**	**FY 2014**	**FY 2015**	**FY 2016**	**FY 2017**
Performance (measurement) period	June 2008-July 2011	June 2009-July 2012	June 2010-July 2013	June 2011-July 2014	June 2012-July 2015
Diagnoses of initial hospitalization	Heart attack Heart failure Pneumonia	Heart attack Heart failure Pneumonia	Heart attack Heart failure Pneumonia COPD Hip or knee replacement	Heart attack Heart failure Pneumonia COPD Hip or knee replacement	Heart attack Heart failure Pneumonia* (expanded) COPD Hip or knee replacement CABG
Penalties: Percentage reduction in base payments on all Medicare inpatient admissions					
Maximum rate of penalty	1%	2%	3%	3%	3%
Average hospital payment adjustment (among all hospitals)	-0.27%	-0.25%	-0.49%	-0.48%	-0.58%
Average hospital penalty (among penalized hospitals only)	-0.42%	-0.38%	-0.63%	-0.61%	-0.74%
Percent of hospitals penalized	64%	66%	78%	78%	79%
Percent of hospitals at max penalty	8%	0.6%	1.2%	1.1%	1.8%
CMS estimate of total penalties	$290 million	$227 million	$428 million	$420 million	$528 million

Diabetes in Hospitalized Patients

The number of diabetes-related emergency department visits and hospitalizations has increased with the increasing prevalence of diabetes in the U.S[12]. While only 9.4% of the U.S. population had diabetes in 2015[12], 25-30% of hospitalized patients have diabetes.[13] Patients with diabetes incur higher hospital costs and longer lengths of stay compared with their non-diabetic counterparts[14]. They are also more likely to require hospitalization through the emergency department, another high-cost resource[14].

Several studies[15–28] identify diabetes as an independent risk factor for hospital readmission. These studies include a broad range of patient populations including those admitted for renal transplant[15], vascular surgery[29], coronary artery bypass graft (CABG) or other cardiac surgery[16,17,21], congestive heart failure(CHF)[22,23], acute myocardial infarction (AMI)[24], any cardiovascular disease admission (AMI, CHF, ischemic heart disease, stroke)[25], stroke[26,27], liver disease[28] and general medical patients[19,20,30]. Compared to the general population whose 30-day readmission rate is 5-14%[31–33], patients with diabetes have a 30-day readmission rate of 14.4-22.7%[31]. Anti-diabetic agents are some of the highest risk medications for causing emergency hospitalization for adverse drug events placing this patient population at even greater risk after a hospital discharge[34].

Inpatient diabetes education[35], case management transition resources[36] and inpatient medication adjustment to improve glycemic control[37–39] may be effective interventions to reduce hospital readmission among patients with diabetes. Understanding causes and trends of readmission in patients with type 2 diabetes (T2DM) has the potential to improve transition-of-care strategies for this at-risk population. Enhancing the quality of care delivery to patients with

diabetes during transitions of care has may reduce healthcare costs by avoiding preventable hospital readmissions.

Predictive Analytics in Hospital Readmissions

As a result of regulations established through the HRRP, the literature abounds with strategies for reducing excess 30-day hospital readmission[40,41]. In a review of the existing literature, Hanson et al established a terminology for classification of the types of interventions implemented based on the timing and setting of the interventions. Under their model, interventions are classified as predischarge interventions, postdischarge interventions or bridging interventions. Yet, given the heterogeneity of the patient populations, interventions and outcomes reviewed, they were unable to determine which individual components of an intervention were responsible for the desired effect[41]. In the most comprehensive systematic review and meta-analysis of interventions to reduce 30-day readmissions to date, Leppin et al use the same classification framework and demonstrated that effective interventions are more complex, more comprehensive and increase a patient's capacity for self-care[40]. Higher complexity required more resources to implement the intervention and required a higher number of patient interactions. Allocating these resource-intensive interventions to the highest-risk patients increases the effectiveness of readmission reduction programs[42].

Given the inability of health care providers to accurately anticipate patients with highest risk for readmission[43], predictive analytics have been employed to aid in identification of highest risk patients. Two systematic reviews[44,45] describe 99 unique, published hospital readmission prediction models. Using prediction models to target resources to high-risk patients has shown

benefit in heart failure patients where this approach significantly reduced readmissions from 26.2% to 21.2%[42].

Before application in clinical practice, the quality of these predictions models must be assessed. An evaluation of prediction models includes an assessment of various aspects of model performance, most commonly discrimination and calibration[46]. Discrimination, how well the model separates those with the outcome from those without, is typically assessed using the concordance-statistic (c-statistic). For a binary outcome, such as hospital readmission, this correlates to the area under the receiver operating characteristic (ROC) curve. C-statistic values may range from 0.5 to 1.0, where 0.5 implies that the model's ability to accurately discriminate is equivalent to chance and 1.0 implies perfect discrimination. One proposed framework classifies a C-statistic of >0.5 to <0.7 as poor discrimination, ≥0.7 to <0.8 as acceptable discrimination, ≥0.8 to <0.9 as excellent discrimination and ≥0.9 as outstanding discrimination[47]. For calibration, the measure of agreement between observed and predicted outcomes, the evaluation method is more variable and not consistently reported in the biomedical literature[44,45].

Most of the studies describing readmission risk prediction models demonstrate only acceptable or poor discriminatory power. In the systematic review of hospital readmission risk prediction models by Kansagara et al, 26 unique models were identified with c-statistic values ranging from 0.55 to 0.83. However, only 6 models demonstrated a c-statistic greater than 0.70[44]. In a more recent systematic review by Zhou et al, 60 studies of 73 unique models reported a c-statistic range from 0.21 to 0.88. Two studies reported a c-statistic of >0.8 (excellent), 11 reported a c-statistic of ≥0.7 to <0.8 (acceptable) and all other studies demonstrated poor discrimination performance with a c-statistic <0.7[45].

In addition to their discriminatory performance, predictive models can be characterized by other features such as the development cohort, readmission outcome, variables included in the predictive model and the statistical algorithm(s) used to develop the model.

Most commonly, the cohort is selected based on the diagnosis at index hospitalization, typically one or more of those penalized under the HRRP (i.e., CHF, pneumonia, COPD). Diagnosis-specific models are increasing in prevalence[45] and demonstrate improved accuracy compared to models developed in more heterogeneous cohorts[48]. Institution-specific risk readmission models are also common. They may offer improved discrimination[49] when compared with publically available models and those endorsed by CMS[23,50] but suffer in the ability to generalize to other populations.

Historically, there has been an overreliance on administrative billing data as features in readmission predictions models. One significant limitation to that approach is the lack of the availability of that information at the time of hospital discharge limiting their use in real-time clinical decision-making. The widespread implementation of electronic health records has increased electronic access to clinical data. More recent literature has shown that clinical data and utilization history make the greatest contributions to predictive accuracy[45,51] and improve model performance[52].

The majority of existing models tend to focus on risk of all-cause readmission for a pre-selected cohort or for all patients hospitalized at the institution. However, the discriminatory power of a predictive model can vary by more than 20% when the readmission diagnosis is changed[51]. Models focusing on all-cause readmission are limited in their clinical utility to focus transition and post-discharge resources to patients who need them the most. The ability to

predict the reason for readmission may guide more specific interventions by transition coordinators and population health managers.

Despite the emergence of newer machine learning techniques, logistic regression remains the most common method for the development of readmission risk prediction models. All models from the updated[45] systematic review of hospital readmissions and all except one model[49] from the original systematic review[44] continue to use logistic regression for model development. The most widely used models for readmission risk prediction in clinical practice are logistic regression models[53], LACE[54] and LACE+[55]. LACE, developed in medical and surgical patients from 11 community hospitals in Ontario from 2004 to 2008, considers length of stay (L), acuity of the hospital admission (A), comorbidities (C) and emergency department visits in the previous 6 months (E) to make a prediction. During model evaluation, it demonstrated poor discrimination with a c-statistic of 0.684. The acuity was determined by whether the admission was classified as emergent or elective. Comorbidities were measured as the Charlson comorbidity index. Charlson comorbidity index is a widely-used method for predicting mortality using weighted scoring of conditions[56]. It includes weights for 17 conditions and has been validated across several clinical domains[57–62]. It has also been adapted for use with ICD-9 codes[63]. The LACE+ model includes the predictors from the original LACE index and adds patient age and sex, teaching status of the discharge hospital, number of urgent admissions in previous year, number of elective admissions in previous year, case-mix group score and number of days on alternative level of care status.

Newer machine learning algorithms have shown promise of improved performance compared with traditional logistic regression models, but have rarely been used in the evaluation of hospital readmissions. Least absolute shrinkage and selection operator (LASSO), Random

Forest (RF) and Support Vector Machines (SVM) are novel machine learning techniques which are readily used in the biomedical literature across a wide array of clinical domains. LASSO has been used to predict colon cancer diagnosis[64], pancreatic cancer prognosis[65], response to therapy in Schizophrenia patients[66], mortality after violent crime[67], hip fracture surgery[68], sepsis[69] cardiac procedures[70], neurologic outcomes in pediatric intensive care unit patients[71], pneumonia admissions in the general population[72], infection after burn[73] and hospital acquired pneumonia in stroke patient[74] among others. RF has seen application in non-small cell lung cancer response to chemotherapy[75], infectious complications in combat casualties[76], mortality in cholangitis[77], sepsis[78] and AMI[79], Clostridium difficile recurrence[80], severe Hand, Foot and Mouth disease[81], cardiovascular event prediction[82], extrauterine disease in patients with endometrial cancer [83] and relapse in childhood acute lymphoblastic leukemia (ALL)[84]. Applications of SVM include prediction of post-operative sepsis and acute kidney injury (AKI)[85], lung cancer[86], mortality after trauma[87], sepsis[88] and after cystectomy for bladder cancer[89] and breast cancer survival[90].

Despite the promising results seen in other biomedical domains, there are few examples in the biomedical literature applying newer machine learning techniques to the prediction of unplanned hospital readmission. Yu et al used an SVM framework to develop and evaluate institution-specific and diagnosis-specific readmission risk prediction models. These were compared to a widely used logistic regression model, LACE[91], and consistently demonstrated improved discrimination performance[49]. Futoma et al directly compared the discrimination performance of logistic regression to several other commonly used statistical techniques including logistic regression with multi-step variable selection (LRVS), penalized logistic regression (PLR), RF and SVM. The evaluation was performed across 280 cohorts as determined by the visit DRG and the same set of variable predictors was used in each. RF and

PLR consistently outperformed all other techniques[48]. Jamei et al analyzed several different methods including logistic regression, RF and artificial neural networks (ANN)[92]. ANN demonstrated significantly greater performance than all other methods, including RF, the model with the second best performance. Frizzell et al reported no difference in discrimination performance between naïve Bayesian network, RF, gradient-boosted logistic regression and LASSO models when applied to predict 30-day readmission in 56,477 Medicare patients with heart failure[93]. In a study of short (30-day) and long-term (180-day) hospital readmission in patients with heart failure, Mortazavi et al demonstrated that RF and boosting improved 30-day all-cause and 30-day heart failure readmission, respectively, when compared to logistic regression[94].

Hospital Readmissions in Patients with Diabetes

A review of existing literature reveals several studies describing independent risk factors for readmission in patients with diabetes. Most of these studies use logistic regression to identify independent risk factors for readmission but do not assess the validity of the model or serve as a standalone tool. Within a population of patients with diabetes, age[95], race/ethnicity[95–97], payer[95,35], socioeconomic status[95,97], source of admission[98], comorbidities[98–102], length of stay[97,35], number of prescribers in previous year[99], hospitalizations in the previous 6 months[99], smoking status[100], polypharmacy[101], living in an urban setting[103], presence of secondary hypoglycemia during admission[104], and failure to record a diabetes diagnosis at discharge in patients with previous diabetes diagnosis[105] are risk factors for 30-day hospital readmission. Three studies summarized in Table 2 present readmission risk prediction models developed in diabetes cohorts, evaluated in a separate validation sample and presented as standalone risk

prediction tools. Their cohorts of study vary slightly but the features used to develop the algorithm are similar and they all evaluated the risk of all-cause readmission. On internal validation, they each demonstrated acceptable or excellent discrimination but none of these used newer machine learning techniques which have shown promise to improve model performance in other readmission risk prediction models and other areas of study in biomedical literature**.**

There are currently no published studies examining readmission risk in diabetes patients which use newer machine learning techniques or a diagnosis-specific readmission outcome to try to improve model performance and clinical utility.

Table 2. Summary of Current, Internally Validated Readmission Risk Prediction Models in Patients with Diabetes

Study	Cohort	Features	Outcome	Algorithm	Internal Validation
Rubin[106] 2015 DERRI	Hospitalized patients with diabetes	• Demographic • Laboratory • Medications • Microvascular complications • Utilization history	30-day All-Cause	Logistic regression	C-statistic 0.69
Rubin[107] 2017 DERRI-CVD	Patients with diabetes hospitalized for cardiovascular disease (CVD)	• Comorbidities • Demographic • Laboratory • Medication • Microvascular complications • Visit utilization	30-day All-Cause	Logistic regression	C-statistic 0.71
Collins[108] 2017	Hospitalized Medicare patients with diabetes	• Clinical conditions • Demographics • Utilization metrics *from claims data	30-day All-Cause	Logistic regression	C-statistic 0.82

We hypothesize that a risk prediction model using novel machine learning techniques (LASSO, RF, SVM) to identify hospitalized patients with type 2 diabetes at highest risk of diagnosis-specific 30-day hospital readmission (DM and CHF) will outperform all-cause readmission, logistic regression-based prediction models. First, we used LASSO, RF and SVM to develop and evaluate the validity of prediction models of hospitalized patients with type 2 diabetes at risk for diagnosis-specific (DM, CHF, All-Cause) 30-day readmission. Next, we compared these model performance metrics with the published and validated LACE all-cause readmission prediction tool.

CHAPTER 2

METHODS

Dataset

We identified a retrospective cohort of inpatient admissions at Vanderbilt University Medical Center (VUMC) in Nashville, Tennessee, between October 1, 2010, and September 15, 2015. This time frame was selected to begin after the start of the initial HRRP benchmarking period to capture encounters over a relatively stable readmission reduction strategy. The study population included adults aged 18 and older with type 2 diabetes (T2DM). A diagnosis of T2DM was based on the presence of the PheWAS parent code 250.2 prior to the index encounter. PheWAS is a research method which uses custom combinations of International Classification of Disease 9th edition (ICD-9) codes to describe phenotypes in electronic health record (EHR) data[109,110]. Encounters for patients classified as "observation" status were excluded as only inpatient admissions are penalized under HRRP.

Diagnosis Related Groups (DRG) and All Patient Refined Diagnosis Related Group (APR-DRG) were extracted for each inpatient encounter. DRG is a framework for associating the types of conditions treated during a hospitalization with the costs associated with treating them. APR-DRG is a proprietary classification scheme developed to enhance the traditional Diagnosis Related Group (DRG) classification of inpatient admissions by incorporating measures of severity of illness and risk of mortality[111]. We used APR-DRG version 31[112] to classify the reason for hospitalization and to define the diagnosis-specific 30-day readmission outcome. All APR-DRGs except 693, chemotherapy, were included in the all-cause readmission

outcome. Readmissions for chemotherapy were excluded from the outcome evaluation as these represent planned inpatient encounters which are not penalized under HRRP.

While diabetes is a relatively less common reason for readmission among subjects in our cohort, it is a common chronic medical condition for which there are known strategies to reduce readmission. Additionally, it is likely underrepresented as the primary reason for readmission due to its relatively lower service intensity weight. Service intensity weight is a measure of cost or resources needed to treat an associated APR-DRG. Given that the discriminatory power of a predictive model can vary significantly when the readmission diagnosis is changed[51], we selected three readmission outcomes for which distinct models were developed. In addition to predicting diabetes-specific readmissions, our original goal, we chose all-cause and heart failure APR-DRGs as readmission outcomes of interest. All-cause excluding chemotherapy was selected as this is the closest representation to the current HRRP implementation which penalizes all-cause unplanned readmissions for certain index admission diagnoses. Heart failure is the most common reason for 30-day readmission in our population and is also one of the index hospitalizations penalized under HRRP.

Structured Query Language (SQL) was used to extract data from the [113], a database of clinical and related data derived from VUMC's clinical systems and restructured for research. Data were preprocessed in Python[114] before being imported into R[115], an open-source software environment for statistical computing. R was used for model development and internal validation.

Feature Selection and Pre-Processing

Features were selected based on domain expert opinion of those clinical variables relevant to the readmission risk of hospitalized patients with diabetes. In order to support real-

time application in a clinical setting, only those features that are available prior to discharge were included. Features selected include demographic information, utilization history and laboratory results. Table 3 describes features from each category used for training of all models.

Demographic features include age, gender, race, insurance payer and area deprivation index. Area of deprivation index is a geographically-based measure of socioeconomic deprivation associated with a 9-digit zip code[116]. Because the full 9-digit zip code was not available for our cohort, we truncated area deprivation index 9-digit zip codes to include the first 5 digits and assigned the median area deprivation index for each grouping. Median, as opposed to mean, area deprivation index was chosen to reduce sensitivity to outliers.

Utilization history included the length of stay (LOS) of the current admission, number of VUMC emergency department visits in the 6 months preceding admission, number of VUMC outpatient clinic visits in the 1 year preceding admission. Six months was chosen as the lookback time for emergency department visits as this metric previously demonstrated validity in predicting hospital readmissions[54]. In contrast, the cadence of many outpatient specialty appointments is less frequent so a longer lookback time of 1 year was used to capture outpatient utilization history. In addition, we chose to include active use of the locally-developed VUMC patient portal, My Health at Vanderbilt (MHAV), prior to the current admission as one of our utilization measures. Patient portals are secure, internet-based platforms where patients may access their personal health information and communicate with health care providers. Patient portal use has been associated with both no impact on[117] and increased risk of[118] 30-day hospital readmission. Utilization data for other hospitals were not available for inclusion.

Laboratory tests included pre-admission and admission values. Laboratory data ranges were reviewed and discarded where not physiologically possible. We noted this systematic error

in point of care (POC) A1C values. These likely represent transcription errors where laboratory equipment was not integrated with the EHR. A1C values <2% and >25% were discarded as spurious, representing 0.03% of A1C readings.

Table 3. Demographic, Utilization History, and Laboratory Results Used as Features in Model Training

Feature Class	Features
Demographic	○ Age ○ Gender ○ Race ○ Insurance payer ○ Area deprivation index[116]
Utilization History	○ Length of Stay ○ # VUMC emergency department visits in 6 months ○ # VUMC outpatient clinic visits in 1 year ○ Active use of VUMC patient portal
Laboratory Results	○ Admission glucose ○ Admission bicarbonate ○ Maximum A1C in last 1 year ○ Boolean value representing if blood glucose checked on day of admission ○ Absolute value of difference between maximum and minimum creatinine during admission ○ Absolute value of difference between maximum and minimum sodium during admission ○ Absolute value of difference between maximum and minimum blood glucose during last 24 hours of the admission ○ Median number of blood glucose readings per day during the admission

We evaluated candidate features for missingness. For blood glucose, missing values may indicate patient or provider-specific characteristics related to the likelihood of hospital readmission. For example, forgetting to check a blood glucose in a patient with diabetes could indicate substandard care delivery increasing the risk for readmission. Alternatively, lack of blood glucose data on the day of admission may indicate that the patient has well-controlled

diabetes and admitted for routine hospital services with low risk of complication and readmission. To address informative missingness, a Boolean variable was created to indicate whether or not blood glucose was measured on the day of admission. All other missing data were assumed to be missing at random.

Where data were missing at random, multiple imputation was used to replace missing values. Multiple imputation uses bootstrapping, sampling with replacement from original non-missing data, to replace missing values. Then, using all cases from the imputed dataset, nonparametric regression is used to generate variable coefficients. Using this model, predicted values are generated for all cases of the variable, missing and non-missing. Last, predictive mean matching is used to fill in the original missing values. With predictive mean matching, a missing value is filled from among the original non-missing values of that variable. Variables are randomly selected from cases where the regression-predicted values of the missing variables are closest to the regression-predicted value for the non-missing variables based on the simulated regression model. We used the "Hmisc" package in R to generate 5 complete datasets[119].

Statistical Modeling

A wide range of machine learning methods have been studied in the biomedical literature. Logistic regression remains a commonly used technique. As a parametric method, it assumes the form of the unknown target function which offers both advantages and disadvantages. Because the form and complexity of the target function are assumed, parametric methods require less computational time and less data to generate predictions. However, they may suffer in accuracy if the assumptions do not match the underlying data. Alternatively, nonparametric methods do not constrain the form of the target function. As a result, the target function will change in shape

and complexity to best fit the underlying data. This flexibility may offer improved accuracy compared to parametric methods but also increases the risk of overfitting. Additionally, they may suffer from high computational time as they have more parameters to train. We studied the behaviors of different machine learning techniques in current biomedical literature as discussed in Introduction. Modern predictive studies should include both techniques in order to determine the optimal approach for the given problem. In designing a data-driven model, it's challenging to know which approach will best fit the problem a priori. Based on methods used in prior readmission work, we selected one parametric and 2 nonparametric methods to study.

LASSO is a form of penalized logistic regression where regularization parameters are used to reduce the magnitude of regression coefficients to avoid overfitting. LASSO tends to select one predictor out of multiple correlated predictors and discards the others resulting in feature selection[120]. An additional tuning parameter, λ, controls the overall strength of the penalty. 10-fold cross validation was used with each imputed dataset to select the shrinkage parameter λ. LASSO was performed using the "glmnet" package in R.

Unlike LASSO which assumes a constrained form of the mapping function, SVM with a radial kernel and RF are non-parametric methods. RF uses bagging, selection of a random subset of observations and a random subset of features, to develop an ensemble of decision trees before polling the trees to create a ranking of classifiers[121]. The use of bagging and random selection of features allow RF to overcome limitations such as sparse and missing data. Our random forest models used 500 trees and were developed using the "ranger" package in R[122]. We used 4 variables available for splitting at each node, the default setting of the "ranger" package, determined as the rounded down square root of the number of predictor variables.

SVM attempts to find a hyperplane which separates observations of different classes[123]. Support vectors are the observations from each class closest to the hyperplane. The best SVM model leaves the largest margin between support vectors of different classes and, thus, reduces overfitting. Model margin of error and complexity are modified with two parameters, cost and γ. Lower cost values increase the margin to allow for incorrect classification. The value of γ determines the complexity of the curve that best separates observations of different classes. When the value of γ is too large, overfitting will result. For our model, we used $\gamma = 1$ and cost = 1.25. We used a radial kernel to develop a nonlinear classifier of the input data. The "e1071" package in R was used to develop an SVM model for diabetes-specific readmission outcome[124].

For each of three readmission APR-DRGs, we developed prediction models using LASSO and RF. SVM was also used for the development of a third model for diabetes-specific readmission (APR-DRG 420). Due to high computational time and inferior performance as discussed in Results, we did not develop an SVM model for HF and all-cause readmission outcomes.

In order to compare our models to the widely used LACE algorithm, we developed a univariate regression model with LACE score as the feature. We evaluated the performance of this model across all three readmission outcomes studied.

Internal Validation

Internal validation involves using available data to estimate how well a given model will perform in a new dataset[125]. Several strategies exist for performing internal validation. In split-sample validation, the observations are randomly divided into two sets, the training set and the testing set. The training set is used to develop the model which is then evaluated in the testing

set. While it benefits from computational simplicity, split-sample validation has several disadvantages. First, the trained model may vary significantly depending on the split of the data, particularly if the predictors or outcome are skewed. Similarly, because the evaluation is performed on a relatively small subset of the larger population, it may not provide a reliable estimate of how the model will perform in practice.

Cross-validation is a variation on split-sample validation which can be used to yield a more consistent model. To perform cross-validation, the data set is divided into some number of equally sized subsets. One of the subsets is held out to serve as the test set while the remaining subsets serve as the training set for model development. This approach of holding out one subset and training an all of the others is continued until each subset serves as the testing set one time. The average error across each of the testing sets is calculated. This approach is less sensitive to variation based on the splitting of the data since all subsets serve as the test set once. Because the final model has learned from all of the available data, the result demonstrates improved performance compared to split-sample validation. However, it can be computationally intensive due to the need to repeatedly train the model[125].

Another approach is bootstrap validation[126]. With this approach, the data are sampled with replacement to create subsets that are equal in size to the original data. Prediction models are developed in each bootstrap subset and on the original data. An evaluation of the difference in performance between them gives an estimate of accuracy. Compared with split-sample and cross-validation, the estimate of model performance demonstrates less variability because the sizes of the subsets are equal to the size of the original. Like cross-validation, it allows the model to train on all available data yielding a more accurate estimate. Because of the need to repeatedly train models on bootstrapped subsets that are as large as the original data, it is more

computationally complex than split-sample and cross-validation. Additionally, bootstrap validation has been demonstrated to work well in high-dimensional datasets where the number of predictors is much larger than the number of cases.

For each model, we used Harrell's algorithm[127] for estimating optimism to calculate the optimism-adjusted performance of the model. This method relies upon bootstrapping to quantify model optimism. For each of 5 imputed complete datasets, models are developed and performance metrics calculated. Each of the 5 datasets is then sampled with replacement 100 times to create 100 new bootstrapped datasets for each. Models developed and evaluated on bootstrapped samples are then evaluated on the imputed dataset from which they were derived. The difference in performance between the original dataset and bootstrapped datasets determines the degree of overfitting. We followed Rubin's rules for pooling results for combining results into an overall multiple imputation estimate[128]. However, our results were not normally distributed based on an evaluation using the Kolmogorov-Smirnoy Goodness-of-fit test[129] so we reported the median value of the performance metric.

Model Performance Evaluation

For each model, we report discrimination, calibration and a precision-recall curve. Discrimination, how well the model separates those with the outcome from those without, was assessed using the c-statistic. For our binary outcome of hospital readmission within 30 days, this correlates to the area under the receiver operating characteristic (ROC) curve. An ROC curve plots true positive rate (sensitivity) by false positive rate (1-specificity) over the range of possible cutoffs for classifying observations as positive or negative. The c-statistic then represents the probability that a random observation with the outcome was given a higher score

than a random observation without the outcome. C-statistic values may range from 0.5 to 1.0, where 0.5 implies that the model's ability to accurately discriminate is equivalent to random chance and 1.0 implies perfect discrimination. The "ROCR" package in R was used to generate ROC curves[130]. Standard error was used to generate confidence intervals around the c-statistic for each test.

For calibration, we report a calibration plot with its slope and intercept values. The calibration plot is a graphical representation of predicted probability compared to observed probability. For binary outcomes, where the observed probability is either 0 or 1, this plot was generated by binning observations into equal size groups based on an ordered list of predicted probabilities and plotting the proportion of outcomes per bin. The "rms" package in R was used to create a calibration plot for one imputed dataset for each model[126]. The calibration intercept measures the extent to which predictions are consistently too low or high using a comparison of the mean of all predicted risks to the mean observed risk[46]. Calibration slope represents the degree of overfitting or underfitting by the regression coefficients where a slope less than 1 suggests overfitting.[131] A perfectly calibrated model is represented by a diagonal with slope = 1 and an intercept = 0. Calibration plots were made and slope and intercept values calculated using the val.prob function from the "rms" package in R[126].

When evaluating a dataset with heavily imbalanced classes, additional measures are needed to present an accurate view of the model's performance. In this setting, ROC curves tend to present overly optimistic results[132,133] as a high number of false positive have only a minimal effect on false positive rate. Precision-recall curves can present a more accurate representation of model performance by accounting for the inappropriate labeling of false positive observations. This penalty for mislabeling negative outcomes is primarily accounted for in precision which

represents the fraction true positive cases out of all cases with a positive label (true positives plus false positives). We used the "ROCR" package in R to create precision-recall curves for each of our models in order to present a more informative representation of our models' performance in the setting of class imbalance.

Feature Importance

Feature selection is inherent to the LASSO algorithm but there may be inconsistency when evaluating models over several bootstraps. We used the Bolasso algorithm to pool these results over 500 bootstraps (100 bootstraps for each of the 5 imputed datasets) [134]. Bolasso is a variable selection algorithm which finds the intersection of all features with non-zero weights in all bootstraps. To calculate odds ratio and confidence intervals, we performed unregularized logistic regression using the Bolasso-selected features for each model and outcome.

For each branch in a decision tree, RF uses Gini impurity to select the variable that provides the best split of the remaining observations[122]. This method seeks the variable that accounts for the greatest variance in the data at each step. For example, a feature which is present for 90% of the observations and not for 10%, has a higher Gini impurity than one present in 50% of observations. We obtained the Gini impurity for all variables for each model and calculated the median Gini impurity across all imputed datasets to report these results.

Clinical Application

In addition to considering performance evaluation metrics such as discrimination, calibration and precision, we want to know how best to apply the model in clinical practice. We need to determine the optimal threshold for predicted probability. Cases with a predicted

probability above that threshold are classified as having the outcome and cases with a predicted probability below that threshold are classified as not having the outcome. We selected RF, the model with the best performance metrics, and created a confusion matrix at various outcome thresholds based on 30-day readmission for diabetes. We randomly selected 112 encounters from our population. None of those encounters were associated with a 30-day readmission for diabetes. We also selected 3 observations with the outcome for a total of 115 patients. This is equivalent to the average number of unique patients seen by the diabetes consult services in one week at VUMC. Using this approach, the outcome prevalence for this subset of our population was 2% compared to the true outcome prevalence in our total population of 0.3%. While we acknowledge this difference in outcome prevalence will impact our results, this exercise has utility in the demonstration of how to implement this model in clinical practice. We also include a discussion of the limitations to this approach.

We calculated the sensitivity, specificity and precision at various cutoffs. Because hospital readmission for diabetes is a rare but serious event, we prioritized sensitivity over specificity and precision when selecting the ideal threshold. Whereas sensitivity and specificity indicate a test's ability to properly detect or reject cases, respectively, precision indicates the likelihood of the outcome given a positive test rest. Unlike sensitivity and specificity, precision is affected by the prevalence of the outcome in the population. With a lower outcome prevalence, we expect to see lower precision. As such, we use this method to demonstrate how to apply this model in clinical practice but would need to evaluate prospectively to verify the validity.

CHAPTER 3

RESULTS

We identified 56,258 inpatient encounters for patients with type 2 diabetes admitted between October 1, 2010, and September 15, 2015. The dataset included 29,013 unique patient identifiers of which 10,660 had more than one inpatient encounter during the study period. Table 4 presents the 10 most common 30-day readmission APR-DRGs in this cohort. Although our population demonstrated a 17% rate of all-cause readmission within 30-days, our diagnosis-specific readmissions had a low prevalence. Heart failure was the single most common readmission diagnosis in our population and only accounted for 1% of the 30-day readmissions. While diabetes may have been a factor in many of the hospital readmissions, for reasons previously discussed, only 0.3% of the readmissions were coded with diabetes as the readmission diagnosis.

Table 4. Summary of the 10 Most Common Readmission APR-DRGs for a Cohort of 56,258 Inpatient Encounters of Adults with Type 2 Diabetes Over the Study Period Excluding APR-DRG 693, Chemotherapy

Readmission Reason (APR-DRG)	**Number of 30-day Readmissions** (% of encounters)	**Number of unique patient identifiers**
All-Cause	9,762 (17.4%)	5,293
Heart Failure (194)	531 (0.94%)	394
Septicemia and Disseminated Infections (720)	440 (0.78%)	388
Renal Failure (460)	348 (0.62%)	311
Post-operative, Post-traumatic or other device infections (721)	344 (0.61%)	310
Malfunction, Reaction & Complications of Genitourinary Device Or Procedure (466)	226 (0.40%)	182
Diabetes (420)	191 (0.34%)	133
Cardiac Arrhythmia & Conduction Disorder (201)	186 (0.33%)	162
Other Pneumonia (139)	171 (0.30%)	157
Kidney & urinary tract infection (463)	144 (0.26%)	131
Percutaneous cardiovascular procedures without AMI (175)	121 (0.22%)	116

The absence of a blood glucose check on day of admission may represent data which are missing not at random according to domain expert opinion. To manage informative missingness, we added a Boolean variable to indicate whether or not the test was performed. For all other features, we performed a missingness analysis with results as summarized in Table 5.

Table 5. Summary of Missing Data

Feature	Number of missing values	Proportion of encounters missing data
Age	0	0
Sex	0	0
Race	0	0
Payer	0	0
ED visit count	0	0
Outpatient Visit Count	0	0
Glucose checked day of admission (yes/no)	0	0
MHAV Use (yes/no)	0	0
LOS	5	0.001%
Median BG readings per day	1839	3.3%
Admission Glucose (1st day)	3464	6.2%
Change in blood glucose last 24 hours	5917	10.5%
Change in creatinine during admission	6083	10.8%
Change in sodium during admission	6448	11.5%
Area deprivation index	21511	38.2%
A1C max (in last year)	21753	38.6%
Admission bicarbonate	28217	50.2%

Note. BG = blood glucose. ED = emergency department, MHAV = My Health at Vanderbilt patient portal, LOS = length of stay, A1C max = maximum A1C in the last year.

Predictive performance varied across readmission outcomes and statistical models. Optimism-adjusted discrimination results are presented in Table 6. Across all readmission outcomes, RF demonstrated significantly better discriminatory performance than LASSO or SVM. For diabetes-specific readmission, SVM had the lowest discriminatory performance and highest computational time and, therefore, was not used to develop models for heart failure and all-cause readmission. LASSO performed significantly better for both diagnosis-specific readmission outcomes than for all-cause readmission. While there was less variation in the discriminatory performance of RF across readmission outcomes, RF did demonstrate a

statistically significant improvement in performance for diabetes-specific readmission when compared to all-cause or heart failure readmissions.

LACE demonstrated inferior discriminatory performance in predicting all-cause readmission in our population than it did at model development[54]. While developed to predict all-cause readmission, LACE demonstrated its best discriminatory performance when predicting heart failure readmission in our external validation. It was no better than chance at predicting readmission for diabetes. Across all readmission outcomes, LACE demonstrated inferior discriminatory performance in this external validation when compared to our models.

Table 6. Discriminatory Performance of Each Statistical Model Across all Readmission Outcomes

Statistical Model	**Readmission Outcome** ROC (95% CI)		
	Diabetes	**Heart Failure**	**All-Cause**
LASSO	0.85 (0.849-0.850)	0.71 (0.709-0.710)	0.64 (0.648- 0.648)
RF	0.95 (0.949-0.951)	0.93 (0.929-0.931)	0.94 (0.939-0.940)
SVM	0.84 (0.838-0.842)	---	---
LACE	0.477	0.670	0.594

To assess calibration, we measured calibration slope and intercept and examined a calibration plot for each model. Optimism-adjusted calibration performance metrics for slope and intercept are summarized in Table 7. LASSO demonstrated the most consistent and well-calibrated models with slope near 1 and intercept near 0 for all models. SVM demonstrated poor calibration with a slope indicating the model underfit the data and an intercept indicating predictions are systematically too high. RF tended to underfit data, more for all-cause

readmission than for diagnosis-specific readmission. RF also demonstrated a slight tendency for high predictions, particularly in heart failure and all-cause readmission outcomes, when compared with LASSO.

Table 7. Calibration Performance of Each Statistical Model Across all Readmission Outcomes

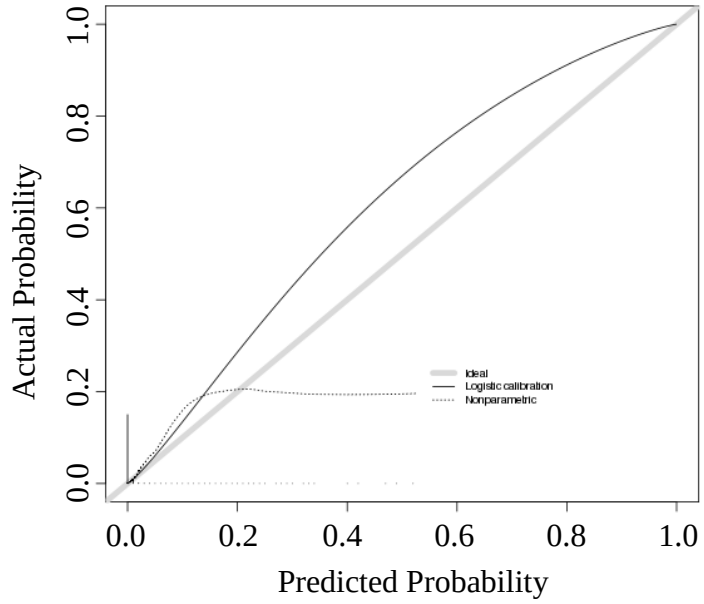
Statistical Model	**Readmission Outcome**					
	Diabetes		**Heart Failure**		**All-Cause**	
	Slope	Intercept	Slope	Intercept	Slope	Intercept
LASSO	0.9924598	-0.0423310	0.97364	-0.1168155	1.000334	-0.0015909
RF	3.273016	0.5016695	2.679752	2.71451577	7.267741	2.2598
SVM	14.5954	70.4701	---	---	---	---

Whereas slope and intercept rely on a single value to describe model fit and systemic tendency for high or low predictions, calibration plots show cases where the model calibration may vary across the distribution of observations. A single, representative calibration plot for each model is shown below in Figure 1. For each plot, a gray, shaded line along the diagonal represents ideal calibration. A darker, solid, gray line plots the logistic calibration curve. The logistic calibration curve represents the proportion of outcomes per bin when observations are grouped into equal size bins based on an ordered listed of predicted probabilities. This is useful when plotting calibration for a binary outcome. LASSO demonstrates good calibration across all readmission outcomes. Compared to LASSO, RF demonstrated inferior calibration across all models, particularly for all-cause readmission. However, the calibration plots show that the RF models for diabetes and heart failure readmission outcomes demonstrated reasonable calibration at low outcome probabilities. SVM demonstrated poor calibration at all outcome probabilities

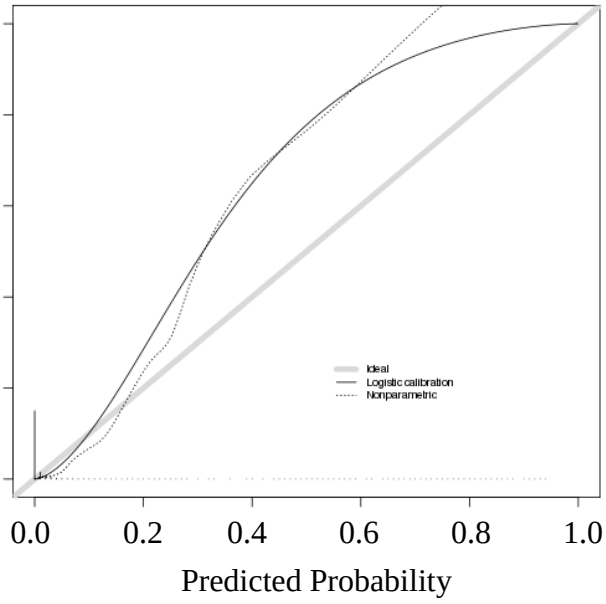
when evaluated for diabetes readmission. Given these results and SVM's inferior discriminatory performance, we did not use SVM to develop CHF or all-cause readmission models.

Figure 1. LASSO, RF and SVM Calibration Plots for Diabetes, Heart Failure and All-Cause Readmission Outcomes

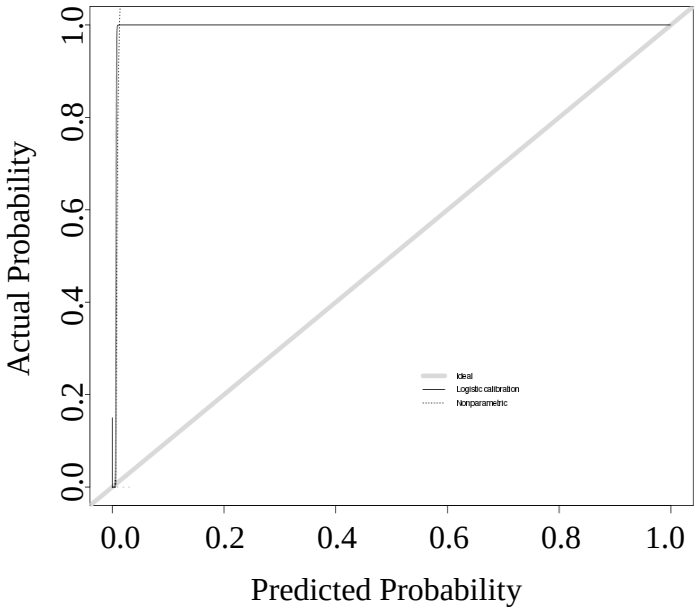
A. Diabetes—LASSO

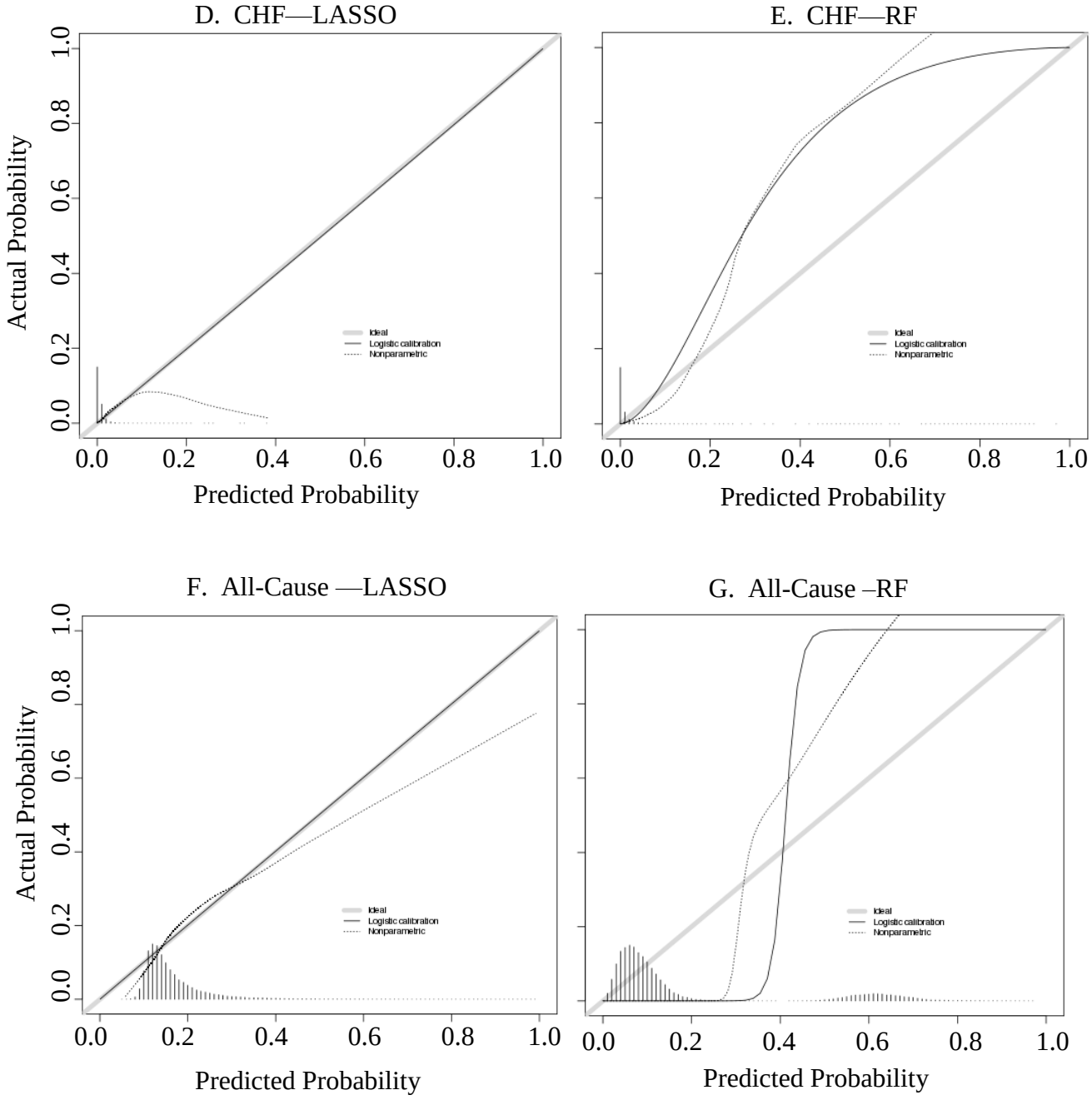


B. Diabetes—RF



C. Diabetes—SVM



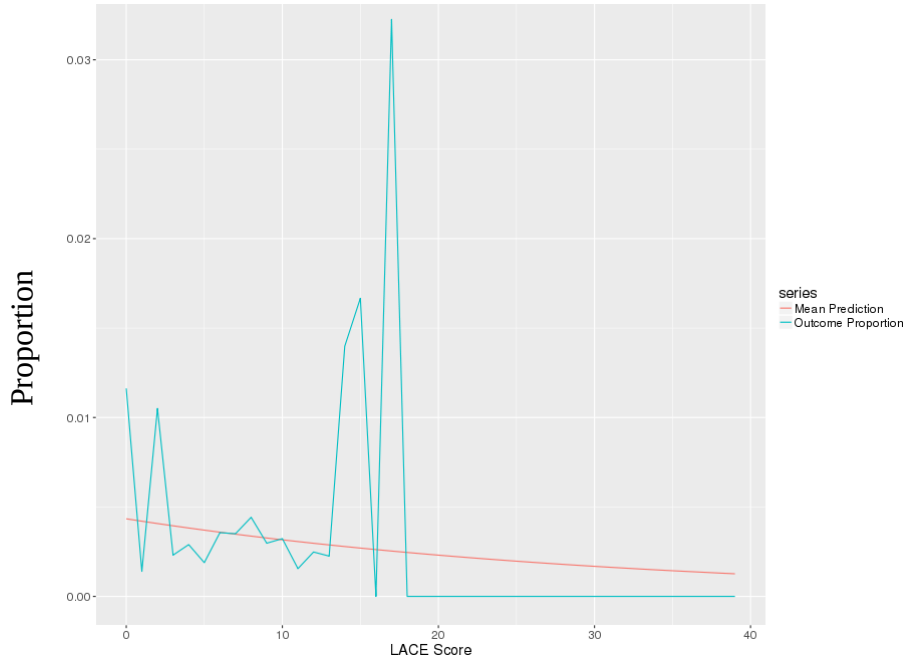


Note. 1A, 1B and 1C demonstrate a representative calibration curve from a single bootstrap for LASSO, RF and SVM, respectively, when evaluated for diabetes-specific readmission. 1D and 1E show representative plots for LASSO and RF, respectively, for heart failure readmission. 1F and 1G display a representative calibration plot for LASSO and RF when evaluated for all-cause readmission. For each plot, the light gray, shaded line along the diagonal represents ideal calibration. The darker, solid gray line plots the logistic calibration curve in which we are most interested.

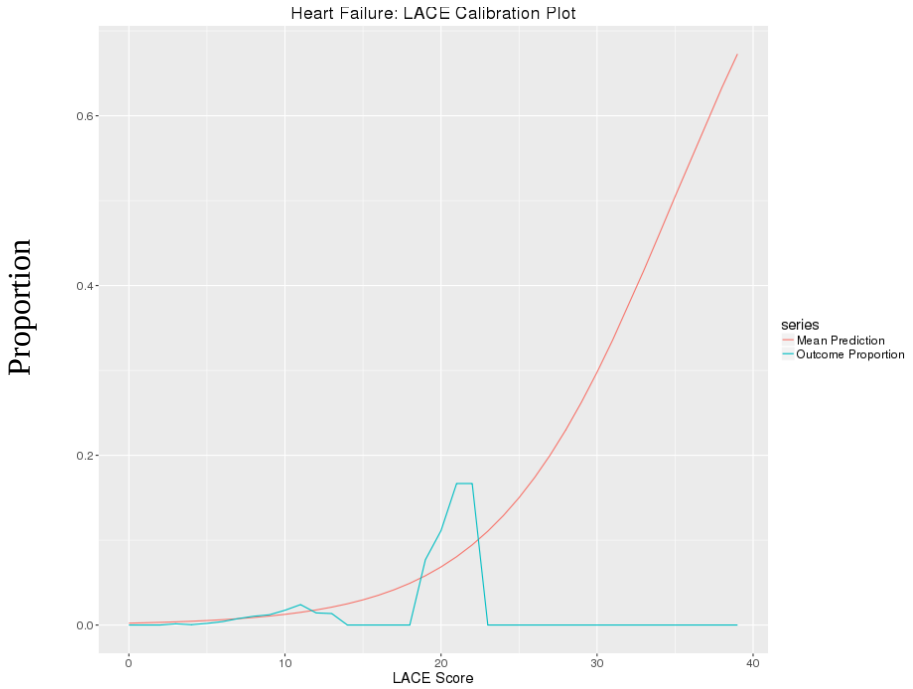
To evaluate the calibration of LACE, we used the univariate logistic regression model with LACE score as the variable to plot the mean predicted probability at each LACE score against the outcome proportion at each LACE score. These plots are shown in Figure 2 below. The maximum outcome proportion for diagnosis-specific readmission outcomes is much lower than for all-cause readmission as reflected by variation in the y-axis scale on these figures. This reflects the overall low prevalence of diabetes and heart failure-specific readmission outcomes in our population. For all-cause readmission, this model demonstrated good calibration, particularly at lower LACE scores where there were more observations. For the diagnosis-specific readmission outcomes, the calibration is not as good but is difficult to evaluate given the low prevalence of the outcome at some LACE scores.

Figure 2. Calibration Plots for LACE Score Univariate Regression Model for Each Readmission Outcome

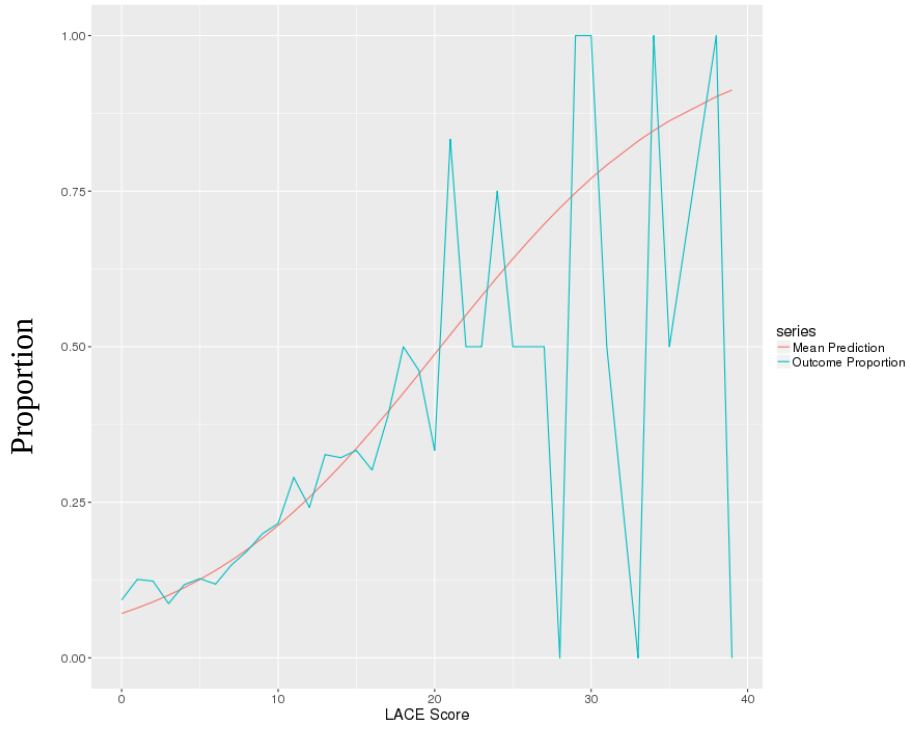
LACE Calibration for Diabetes Readmission Outcome



LACE Calibration for Heart Failure Readmission Outcome

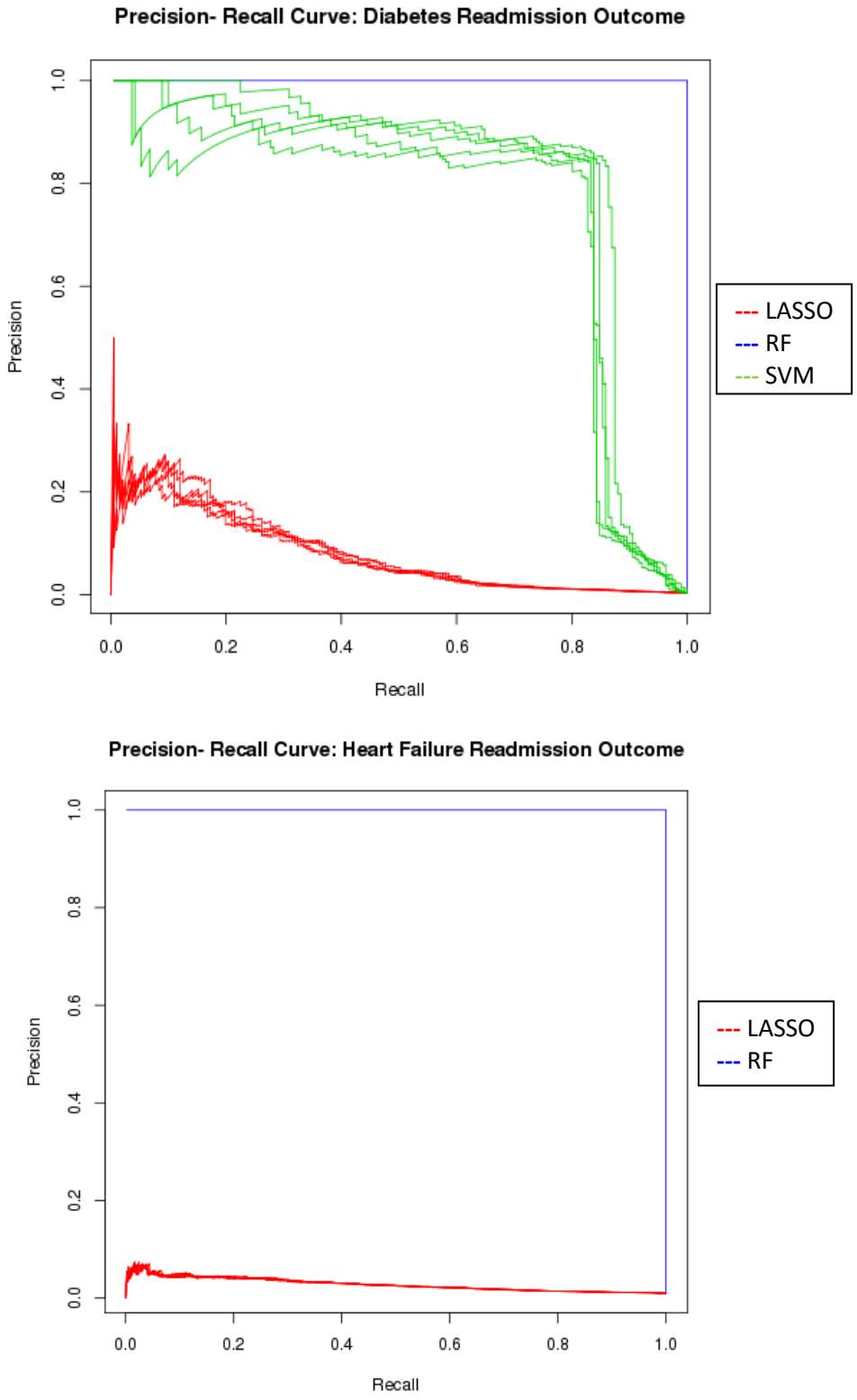


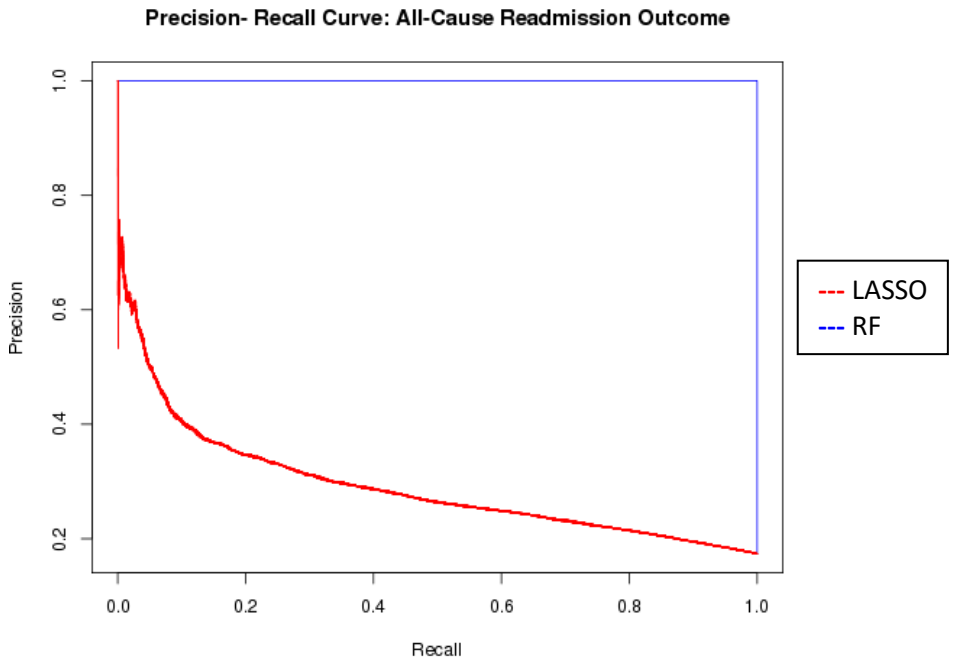
LACE Calibration for All-Cause Readmission Outcome



In a skewed dataset where the number of outcomes is rare, precision can give a more accurate representation of model performance by accounting for the number of false positives. With a rare outcome, a model may demonstrate excellent discrimination simply with a base-rate classifier which has a high true positive rate but also a high number of false positive results. Figure 3 illustrates precision-recall curves for each readmission outcome. LASSO demonstrates poor precision across all models, particularly for diagnosis-specific readmission outcomes. Despite demonstrating poor discrimination and calibration, compared with LASSO, SVM more accurately labeled negative outcomes resulting in better precision. RF demonstrated excellent precision and recall for all models.

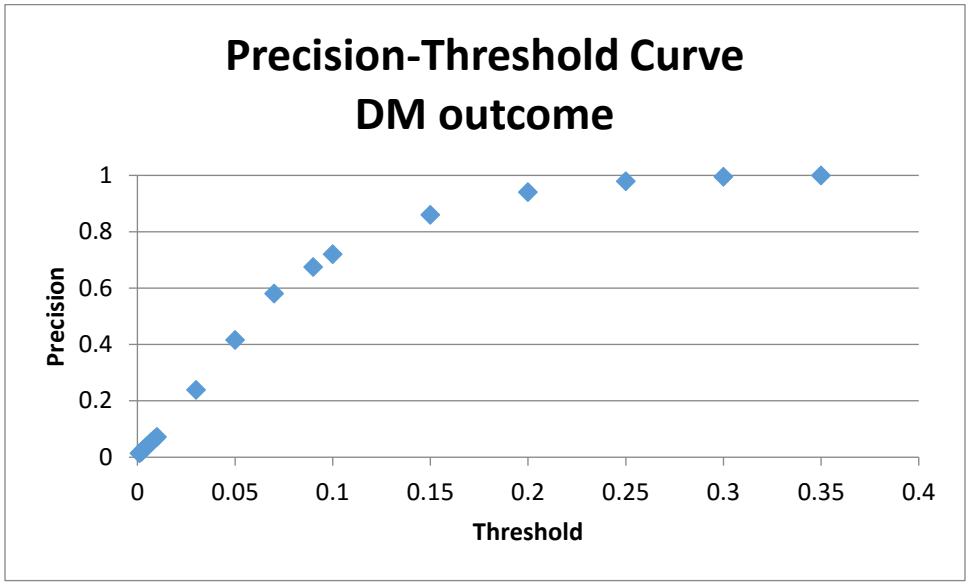
Figure 3. Precision-Recall Curves for LASSO, RF and SVM Models for Each Readmission Outcome





Because the precision of RF appears to be spuriously high, we examined it more closely at low thresholds near the outcome prevalence for diabetes. These results can be seen in Figure 4. At low thresholds, the precision for RF is very low but it rises quickly with increasing threshold. Choosing the threshold based on outcome prevalence alone will result in a model with low precision due to a high number of false positives. This could lead to misallocation of resources to patients who have low risk of readmission. As a result, we will want to select the predicted risk threshold at which we would implement the model in clinical practice based on optimizing sensitivity, specificity and precision.

Figure 4. Plots of Precision Versus Threshold for RF Model on DM Outcome



Note. Precision is poor at low thresholds, including the outcome prevalence of 0.3%, but rises quickly with increasing threshold.

The features selected by the Bolasso algorithm for each readmission outcome are shown in table 8 below. All selected features were statistically significant with p-value <0.001 except race in the heart failure readmission model. Of the 17 features included in model development, 12 were selected after Bolasso evaluation for at least one readmission outcome. Payer, length of stay, area deprivation index, change in creatinine during admission and use of the patient portal did not appear in any model. Only two features (age and number of emergency department visits in previous 6 months) were selected by the Bolasso for all 3 readmission outcomes.

In all 3 models, more emergency department visits were associated with an increased risk of readmission. Compared with the heart failure readmission model where higher age increased the risk of readmission, increasing age was associated with a slightly reduced risk for readmission in diabetes-specific and all-cause readmission models. While this counterintuitive to clinical intuition, the odds ratio in both of these models was very close to 1. One other

consideration to explain this result is death as a competing risk with increasing age. While we considered the impact of death on readmission risk, it was not directly modeled.

For diabetes-specific readmission, 2 other features were associated with a reduced risk of readmission. These included blood glucose being checked on day of admission and increasing admission bicarbonate. Both are clinically plausible. Monitoring blood glucose upon admission indicates and awareness of and attention to diabetes management by the clinician. Low bicarbonate values indicate acidosis, a serious condition associated with some diabetes-related conditions. Elevated A1C indicates poorly controlled diabetes which increases risk of cardiovascular disease and stroke as well as numerous other diabetes-related conditions[135] so it is not unexpected to see that as one of the strongest predictors of hospital readmission in patients with diabetes.

Some of the same features present in the diabetes-specific model were also present in the heart failure model, but a few new features also emerged. A larger change in sodium during the hospital admission was associated risk of readmission. Sodium alterations can exist in a number of complex medical conditions including diabetes, renal failure, liver failure, nutritional deficiency and heart failure among others. Heart failure is the only Bolasso model selecting race as a feature. While none reach statistical significance in the unregularized regression model, black, other, and unknown race were associated with higher risk of readmission compared to white race. The effect of race on heart failure prognosis is mixed in reported literature[136–139].

The Bolasso all-cause model is the only one to select sex as a feature with male sex indicating an 11% increased risk of readmission compared to female. Two measures of utilization history, emergency department visits in 6 months and outpatient visits in 1 year, appear in the all-cause readmission model. Several of the laboratory values present in either the

diabetes or heart failure models appear in the all-cause model including change in blood glucose in last 24 hours of admission, change in sodium during admission and admission glucose. All are likely markers of disease severity and complexity.

Table 8. Features Selected from LASSO Models Using Bolasso Algorithm With OR and P-value

Features	**Readmission Diagnosis**		
	DM	**HF**	**All-Cause**
A1c max (in last year)	1.2752145*	1.096570*	
Age	0.9729791*	1.039310*	0.9924642*
Median BG readings per day	1.1663328*		
Admission bicarbonate	0.8061060*		
Change in BG (in last 24 hours)	1.0041854*		1.0008262*
Change in sodium during admission		1.054966*	1.0390031*
ED visit count (in 6 months)	1.1070867*	1.132465*	1.1618994*
Admission Glucose (1st day)			0.9991449*
Glucose checked day of admission (yes/no)	0.1758684*		1.4191464*
Outpatient Visit Count (in last year)		1.015681*	1.0143048*
Race		Black 3.508632 (p= 0.0789) Other 3.610592 (p=0.1626) Unknown 3.828665 (p=0.3380) White 2.165657 (p=0.2772)	
Sex			Male 1.1130738*

Note. All features selected demonstrate statistical significance except race in the LASSO heart failure model. *p-value <0.001. BG = blood glucose, ED = emergency department, A1C max = maximum A1C in the last year.

Random forest evaluates the amount of variance explained by a variable with each split of the decision tree. This is reported as importance. Table 9 summarizes the feature importance results from all three random forest models. It is notable that age is reported as the most important feature in all 3 RF models and admission glucose is the 2nd. Age was one of only two features selected in all 3 Bolasso evaluations.

Whereas no markers of socioeconomic status appeared in any of the final Bolasso models, area deprivation index appeared in the top 5 importance for all 3 RF models. This supports prior work[140,141] which found that low socioeconomic status was associated with an increased risk for readmission likely due related to low self-efficacy, low health literacy and limited access to healthcare resources[142].

Laboratory values including admission glucose, change in creatinine during admission, change in blood glucose in last 24 hours of admission and max A1C in last year make up the remainder of the top 5 across all RF models. Variation in blood glucose and creatinine during the admission reflect the severity and lability of the underlying disease. Alterations in kidney function, as reflected by changes in serum creatinine, directly impact glycemic control due to the role kidneys play in the metabolism of insulin. Labile renal function can cause a broad range of glycemic excursions which may include hypoglycemia with impaired renal function due to reduced insulin degredation or hyperglycemia if renal function improves and insulin metabolism is increased. Additionally, if glycemic control is highly labile in the hospital where dietary choices and physical activity are often more consistent than what patients experience when not hospitalized, it is clinically plausible that diabetes will also be difficult to control after discharge, placing the patient at increased risk for readmission. Admission glucose and maximum A1C in the last year are more likely to reflect a patient's capacity for self-care. The goal of many care

transition programs is to increase the patient's capacity for self-care, however, these are time and resource-intensive efforts which are often difficult to sustain.

The 5 least important variables were the same across all 3 RF models. These are payer, race, use of the patient portal MHAV and whether or not blood glucose was checked on admission. With the exception of race, which did appear in the Bolasso model for heart failure readmission but did not reach statistical significance on unregularized regression, these are the same features that did not appear in the final Bolasso model for any outcome. The inclusion and exclusion of many of the same features between LASSO and RF supports the validity of their findings.

Table 9. Random Forest Feature Importance for Diabetes, Heart failure, and All-Cause Readmission Models

Importance	Diabetes	Heart Failure	All-Cause
1	Age	Age	Age
2	Admission Glucose	Admission Glucose	Admission Glucose
3	A1C Max	Change in Blood Glucose	Area deprivation index
4	Change in Blood Glucose	Change in creatinine	A1C Max
5	Area deprivation index	Area deprivation index	Change in creatinine
6	Change in creatinine	A1C Max	Outpatient visit count
7	Median BG readings per day	Outpatient visit count	Change in Blood Glucose
8	Outpatient visit count	LOS	LOS
9	Admission bicarbonate	Admission bicarbonate	Admission bicarbonate
10	ED visit count	Change in sodium	Change in sodium
11	Change in sodium	ED visit count	ED visit count
12	LOS	Median BG readings per day	Median BG readings per day
13	Payer	Payer	Payer
14	Race	Race	Race
15	MHAV Use	Sex	Sex
16	Sex	MHAV Use	MHAV Use
17	Glucose checked day of admission	Glucose checked day of admission	Glucose checked day of admission

Note. Features with higher importance measures explain greater variance in the data. BG = blood glucose.

Clinical Application

In order to simulate a population of patients seen by the Endocrinology consultation service at VUMC in 1 week, we identified a subset of our population containing 115 observations. Of those, 112 were randomly chosen and included no cases. Three cases were randomly selected for inclusion in order to enable us to perform the evaluation. This outcome prevalence of 2% is much greater than 0.3%, the true outcome prevalence in our population.

While this difference will impact our results, we will discuss these limitations and recognize the value of this exercise in demonstrating how we could implement our model in clinical practice.

We evaluated sensitivity, specificity and precision of the classification at various predicted probability thresholds. Results of the evaluation are summarized in table 10. This model suffers from poor precision at our true outcome threshold of 0.3% but improves with increasing cutoffs for the threshold. For this rare but serious outcome, we prioritize sensitivity to ensure all at risk patients receive the intervention. Based on these results, we can achieve ideal sensitivity, specificity and precision at a threshold of 0.075 (or 7.5% risk of readmission) allowing us to match our resources with the highest-risk patients. Using this example, we would recommend the intervention for a patient with a predicted probability of readmission 7.5% or greater.

Given that the outcome prevalence of 2% in our example is much greater than that of the underlying population, this evaluation likely overestimates the true precision and needs prospective, external validation. As demonstrated in Figure 4 above, precision across the entire population from a 5 year period is poor at low threshold values, including the outcome prevalence of 0.3%, but becomes 1 at a threshold of 35% where sensitivity and specificity are also calculated to be 1.

Table 10. Sensitivity, Specificity, and Precision Evaluation of 115 Observations, Including 3 Cases of Diabetes-Specific Readmission

Threshold	**FN**	**FP**	**TN**	**TP**	**Precision** (PPV)	**Sensitivity** (Recall)	**Specificity**
0.002	0	16	96	3	0.16	1.00	0.86
0.003	0	9	103	3	0.25	1.00	0.92
0.005	0	6	106	3	0.33	1.00	0.95
0.007	0	4	108	3	0.43	1.00	0.96
0.008	0	4	108	3	0.43	1.00	0.96
0.009	0	4	108	3	0.43	1.00	0.96
0.01	0	4	108	3	0.43	1.00	0.96
0.015	0	4	108	3	0.43	1.00	0.96
0.02	0	3	109	3	0.50	1.00	0.97
0.03	0	3	109	3	0.50	1.00	0.97
0.05	0	1	111	3	0.75	1.00	0.99
0.075	0	0	112	3	1.00	1.00	1.00
0.1	0	0	112	3	1.00	1.00	1.00

CHAPTER 4

CONCLUSIONS

This study presents the performance of three machine learning methods to predict three different 30-day readmission outcomes. We used LASSO, RF and SVM to predict unplanned 30-day readmission for diabetes. Due to the inferior performance and high computational time associated with SVM, we used only LASSO and RF to predict unplanned readmission for heart failure and all-cause readmission.

RF offered the best discriminatory performance among all models across all three readmission outcomes. While it was not as well-calibrated as LASSO, next steps would include the implementation of techniques such as binning, Platt scaling or isotonic regression to improve calibration[143]. Additionally, many of our performance metrics for RF seem high. While there is literature supporting similar behavior of these methods in other settings, given our great class imbalance, we are concerned that our results may represent overfitting. Future work should also include other methods such as oversampling and undersampling in attempt to avoid overfitting in a highly imbalanced dataset.

While LASSO's discriminatory performance was not as high as RF, it demonstrated excellent discrimination for diabetes-specific readmission and acceptable discrimination for heart failure-specific readmission. Its discriminatory performance is also superior to the readmission prediction models most commonly used in clinical practice[91]. LASSO benefits from low computational time. Additionally, LASSO is well-calibrated for all outcomes so it would not require recalibration methods which may impact discriminatory performance. Lastly, the

features selection aspect of LASSO improves the interpretability of the results. These advantages of LASSO make it an important model to consider, possibly in combination with other methods, for future application.

The informatics contribution of this work is the application of newer machine learning techniques to a novel population and evaluating the performance while varying the outcome of interest. It also demonstrated the value of using domain knowledge in the development pipeline and not solely relying upon available structured data. This builds on the large body of readmission prediction model literature which consists mostly of logistic regression models with an overemphasis on administrative billing data to predict all-cause readmission.

Clinically, there has been little work predicting readmission in patients with diabetes. Of the three published studies, none use novel machine learning approaches and none are considering outcomes other than all-cause readmission which limits their utility and impacts their model performance. We expand that body of knowledge by adding new methods to the approach of hospital readmission risk prediction for patients with diabetes. Additionally, by predicting diagnosis-specific readmission, our results are directly actionable by a diabetes, or other disease-specific, service line.

There are several strengths to our approach. One is the use of both parametric and nonparametric methods applied to our population of interest. It's difficult to know a priori which method will best suit the problem and this approach enables comparison across methods while maintaining consistency of the underlying data. Another is the use of domain expertise to select and transform data for inclusion as features in our study. This, in combination with presentation of selected features and feature importance from LASSO and RF, respectively, leads to models which are clinically meaningful and more likely to be accepted by end users. It may also explain

why our models performed so well with relatively few features. Because we are predicting the specific reason that the patient is returning to the hospital, a disease-specific service line can target disease-specific interventions to highest-risk patients in attempt to prevent the readmission. Unlike some all-cause readmission prediction models, the output is focused and actionable related to a specific condition.

One of the limitations of our study is an academic medical center as the single source of data which may impact the generalizability of the results. Additionally, we did not have data regarding utilization history and readmissions to other facilities which could skew our results and underestimate the outcome prevalence. Another important limitation to our study is the use of APR-DRG to define readmission diagnosis, particularly for the diabetes outcome. In the current reimbursement structure, APR-DRG is used to determine payment for a hospitalization based on service intensity weight. Medical coders will look for criteria to assign the APR-DRG with the highest possible service intensity weight to the admission in order to maximize reimbursement. Because the service intensity weight for diabetes is low, medical coders will look for almost any other APR-DRG to define the hospitalization even if diabetes is or is not directly related to the true reason for admission. Lastly, while we consider it a strength to use domain expertise in the pre-selection of model features, one could consider an argument for allowing the models to use as much available clinical data as possible.

Future work must include a prospective evaluation of performance validity in order to address the above limitations and understand the degree of overfitting of the results. Future research should also include enhancement of the methods used to assign readmission for diabetes. This could include natural language processing to evaluate documentation during the admission as well as other tools to develop a phenotype based on any available EHR data.

REFERENCES

1. NHE-Fact-Sheet [Internet]. 2017 [cited 2018 Jan 2];Available from: https://www.cms.gov/research-statistics-data-and-systems/statistics-trends-and-reports/nationalhealthexpenddata/nhe-fact-sheet.html

2. Smith Mueller M. Health Maintenance Organization Act of 1973 [Internet]. 1974 [cited 2018 Jan 2]. Available from: https://www.ssa.gov/policy/docs/ssb/v37n3/v37n3p35.pdf

3. Gruber L, Shadle M, CL P. From Movement To Industry: The Growth Of HMOs. Health Aff 1988;7(3):197–208.

4. Kohn LT, Corrigan J, Donaldson MS. Institute of Medicine. To Err Is Human: Building a Safer Health System. [Internet]. In: To Err Is Human: Building a Safer Health System. Washginton DC: National Academies Press; 1999. p. 8.Available from: http://www.journals.cambridge.org/abstract_S095026880100509X

5. America I of M (US) C on Q of HC in. Crossing the Quality Chasm [Internet]. National Academies Press (US); 2001 [cited 2017 Mar 17]. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25057539

6. Report to the Congress: Reforming the Delivery System [Internet]. Washginton DC: 2008 [cited 2018 Jan 2]. Available from: http://www.medpac.gov/docs/default-source/reports/Jun08_EntireReport.pdf?sfvrsn=0

7. Joynt KE, Jha AK. A path forward on Medicare readmissions. N Engl J Med 2013;368(13):1175–7.

8. Patient Protection and Affordable Care Act [Internet]. United States: 2010 [cited 2018 Jan 2]. Available from: https://www.gpo.gov/fdsys/pkg/STATUTE-124/pdf/STATUTE-124-Pg119.pdf

9. Report to the Congress: Promoting Greater Efficiency in Medicare [Internet]. Washington DC: 2007 [cited 2018 Jan 2]. Available from: http://www.medpac.gov/docs/default-source/reports/Jun07_EntireReport.pdf?sfvrsn=0

10. Jencks SF, Williams M V., Coleman EA. Rehospitalizations among Patients in the Medicare Fee-for-Service Program. N Engl J Med [Internet] 2009;360(14):1418–28. Available from: http://www.nejm.org/doi/abs/10.1056/NEJMsa0803563

11. Boccuti C, Casillas G. Aiming for Fewer Hospital U-turns: The Medicare Hospital Readmission Reduction Program [Internet]. Henry J. Kaiser Fam. Found. 2017 [cited 2018 Jan 2];Available from: http://kff.org/medicare/issue-brief/aiming-for-fewer-hospital-u-turns-the-medicare-hospital-readmission-reduction-program/

12. National Diabetes Statistics Report | Data & Statistics | Diabetes | CDC [Internet]. [cited 2018 Jan 2];Available from: https://www.cdc.gov/diabetes/pdfs/data/statistics/national-diabetes-statistics-report.pdf

13. Draznin B, Gilden J, Golden SH, et al. Pathways to Quality Inpatient Management of Hyperglycemia and Diabetes: A Call to Action. Diabetes Care [Internet] 2013 [cited 2018 Jan 2];36(7):1807–14. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23801791

14. Fraze T, Jiang HJ, Burgess J. Hospital Stays for Patients with Diabetes, 2008: Statistical Brief #93 [Internet]. Agency for Healthcare Research and Quality (US); 2006 [cited 2018 Jan 2]. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21413213

15. Ramezani M, Ghoddousi K, Hashemi M, et al. Diabetes as the Cause of End-Stage Renal Disease Affects the Pattern of Post Kidney Transplant Rehospitalizations. Transplant Proc 2007;39(4):966–9.

16. Steuer J, Blomqvist P, Granath F, et al. Hospital readmission after coronary artery bypass grafting: Are women doing worse? Ann Thorac Surg 2002;73(5):1380–6.

17. Whang W, Bigger JT. Diabetes and outcomes of coronary artery bypass graft surgery in patients with severe left ventricular dysfunction: Results from the CABG Patch Trial database. J Am Coll Cardiol 2000;36(4):1166–72.

18. Billings J, Mijanovich T. Improving the management of care for high-cost Medicaid patients. Health Aff (Millwood) [Internet] [cited 2016 Apr 19];26(6):1643–54. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17978384

19. Howell S, Coory M, Martin J, Duckett S. Using routine inpatient data to identify patients at risk of hospital readmission. BMC Health Serv Res [Internet] 2009 [cited 2018 Jan 2];9:96. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19505342

20. Phillips RS, Safran C, Cleary PD, Delbanco TL. Predicting emergency readmissions for patients discharged from the medical service of a teaching hospital. J Gen Intern Med [Internet] [cited 2018 Jan 2];2(6):400–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/3694300

21. Ferraris VA, Ferraris SP, Harmon RC, Evans BD. Risk factors for early hospital readmission after cardiac operations. J Thorac Cardiovasc Surg [Internet] 2001 [cited 2018 Jan 2];122(2):278–86. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0022522301441651

22. Greenberg BH, Abraham WT, Albert NM, et al. Influence of diabetes on characteristics and outcomes in patients hospitalized with heart failure: a report from the Organized Program to Initiate Lifesaving Treatment in Hospitalized Patients with Heart Failure (OPTIMIZE-HF). Am Heart J [Internet] 2007 [cited 2018 Jan 2];154(2):277.e1-8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17643576

23. Keenan PS, Normand S-LT, Lin Z, et al. An administrative claims measure suitable for profiling hospital performance on the basis of 30-day all-cause readmission rates among patients with heart failure. Circ Cardiovasc Qual Outcomes [Internet] 2008 [cited 2018 Jan 2];1(1):29–37. Available from: http://circoutcomes.ahajournals.org/cgi/doi/10.1161/CIRCOUTCOMES.108.802686

24. Dunlay SM, Weston SA, Killian JM, Bell MR, Jaffe AS, Roger VL. Thirty-day rehospitalizations after acute myocardial infarction: a cohort study. Ann Intern Med [Internet] 2012 [cited 2018 Jan 2];157(1):11–8. Available from: http://annals.org/article.aspx?doi=10.7326/0003-4819-157-1-201207030-00004

25. Heller RF, Fisher JD, D'Este CA, Lim LL, Dobson AJ, Porter R. Death and readmission in the year after hospital admission with cardiovascular disease: the Hunter Area Heart and Stroke Register. Med J Aust [Internet] 2000 [cited 2018 Jan 4];172(6):261–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10860090

26. Kennedy BS. Does race predict stroke readmission? An analysis using the truncated negative binomial model. J Natl Med Assoc [Internet] 2005 [cited 2018 Jan 4];97(5):699–713. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15926648

27. Sun Y, Toh MPHS. Impact of Diabetes Mellitus (DM) on the Health-Care Utilization and Clinical Outcomes of Patients with Stroke in Singapore. Value Heal [Internet] 2009 [cited 2018 Jan 4];12:S101–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20586969

28. Berman K, Tandra S, Forssell K, et al. Incidence and Predictors of 30-Day Readmission Among Patients Hospitalized for Advanced Liver Disease. Clin Gastroenterol Hepatol [Internet] 2011 [cited 2018 Jan 4];9(3):254–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21092762

29. Najafian A, Selvarajah S, Schneider EB, et al. Thirty-day readmission after lower extremity bypass in diabetic patients. J Surg Res [Internet] 2016 [cited 2018 Jan 11];200(1):356–64. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26216749

30. Billings J, Dixon J, Mijanovich T, Wennberg D. Case finding for patients at risk of readmission to hospital: development of algorithm to identify high risk patients.

31. Rubin DJ. Hospital readmission of patients with diabetes. Curr Diab Rep [Internet] 2015 [cited 2016 Apr 19];15(4):17. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25712258

32. Fingar KR, Barrett ML, Jiang HJ. A Comparison of All-Cause 7-Day and 30-Day Readmissions, 2014 #230 [Internet]. Rockville, MD: 2017 [cited 2018 Jan 3]. Available from: https://www.hcup-us.ahrq.gov/reports/statbriefs/sb230-7-Day-Versus-30-Day-Readmissions.jsp

33. Benbassat J, Taragin M. Hospital readmissions as a measure of quality of health care: advantages and limitations. Arch Intern Med [Internet] 2000 [cited 2018 Jan 4];160(8):1074–81. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10789599

34. Budnitz DS, Lovegrove MC, Shehab N, Richards CL. Emergency Hospitalizations for Adverse Drug Events in Older Americans. N Engl J Med [Internet] 2011 [cited 2018 Jan 2];365(21):2002–12. Available from: http://www.nejm.org/doi/abs/10.1056/NEJMsa1103053

35. Healy SJ, Black D, Harris C, Lorenz A, Dungan KM. Inpatient diabetes education is associated with less frequent hospital readmission among patients with poor glycemic control. Diabetes Care [Internet] 2013 [cited 2018 Jan 12];36(10):2960–7. Available from: http://care.diabetesjournals.org/cgi/doi/10.2337/dc13-0108

36. Drincic A, Pfeffer E, Luo J, Goldner WS. The effect of diabetes case management and Diabetes Resource Nurse program on readmissions of patients with diabetes mellitus. J Clin Transl Endocrinol [Internet] 2017 [cited 2018 Jan 10];8:29–34. Available from: http://linkinghub.elsevier.com/retrieve/pii/S2214623717300212

37. Wei NJ, Wexler DJ, Nathan DM, Grant RW. Intensification of diabetes medication and risk for 30-day readmission. Diabet Med [Internet] 2013 [cited 2018 Feb 27];30(2):e56–62. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23126686

38. Wu EQ, Zhou S, Yu A, et al. Outcomes associated with post-discharge insulin continuity in US patients with type 2 diabetes mellitus initiating insulin in the hospital. Hosp Pract (1995) [Internet] 2012 [cited 2018 Feb 27];40(4):40–8. Available from: http://www.tandfonline.com/doi/full/10.3810/hp.2012.10.1002

39. Lee PH, Franks AS, Barlow PB, Farland MZ. Hospital readmission and emergency department use based on prescribing patterns in patients with severely uncontrolled type 2 diabetes mellitus. Diabetes Technol Ther [Internet] 2014 [cited 2018 Feb 27];16(3):150–5. Available from: http://online.liebertpub.com/doi/abs/10.1089/dia.2013.0168

40. Leppin AL, Gionfriddo MR, Kessler M, et al. Preventing 30-day hospital readmissions: a systematic review and meta-analysis of randomized trials. JAMA Intern Med [Internet] 2014;174(7):1095–107. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24820131

41. Hansen LO, Young RS, Hinami K, Leung A, Williams M V. Interventions to reduce 30-day rehospitalization: a systematic review. Ann Intern Med [Internet] 2011 [cited 2017 Mar 18];155(8):520–8. Available from: http://annals.org/article.aspx?doi=10.7326/0003-4819-155-8-201110180-00008

42. Amarasingham R, Patel PC, Toto K, et al. Allocating scarce resources in real-time to reduce heart failure readmissions: a prospective, controlled study. BMJ Qual Saf 2013;22(12):998–1005.

43. Allaudeen N, Schnipper JL, Orav EJ, Wachter RM, Vidyarthi AR. Inability of Providers to Predict Unplanned Readmissions. J Gen Intern Med 7(26):771–6.

44. Kansagara D, Englander H, Salanitro A, et al. Risk prediction models for hospital readmission: a systematic review. JAMA [Internet] 2011;306(15):1688–98. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3603349&tool=pmcentrez&rendertype=abstract

45. Zhou H, Della PR, Roberts P, Goh L, Dhaliwal SS. Utility of models to predict 28-day or 30-day unplanned hospital readmissions: an updated systematic review. BMJ Open [Internet] 2016 [cited 2018 Jan 9];6(6):e011060. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27354072

46. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. Eur Heart J [Internet] 2014 [cited 2018 Jan 9];35(29):1925–31. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24898551

47. Hosmer DW, Lemeshow S, Sturdivant RX. Applied logistic regression. [Internet]. [cited 2018 Jan 9]. Available from: https://books.google.com/books?hl=en&lr=&id=64JYAwAAQBAJ&oi=fnd&pg=PA313&dq=hosmer+lemeshow+applied+logistic+regression+john+wiley+and+sons:+new+york+2000&ots=DscP7W4snP&sig=Z-fvJANHqNk4neb3bXtAzFto_YM#v=snippet&q=roc&f=false

48. Futoma J, Morris J, Lucas J. A comparison of models for predicting early hospital readmissions. J Biomed Inform [Internet] 2015 [cited 2018 Jan 8];56:229–38. Available from: https://www.sciencedirect.com/science/article/pii/S1532046415000969

49. Yu S, Farooq F, van Esbroeck A, Fung G, Anand V, Krishnapuram B. Predicting readmission risk with institution-specific prediction models. Artif Intell Med [Internet] 2015 [cited 2018 Jan 5];65(2):89–96. Available from: http://www.sciencedirect.com/science/article/pii/S0933365715000949

50. Krumholz HM, Wang Y, Mattera JA, et al. An administrative claims model suitable for profiling hospital performance based on 30-day mortality rates among patients with heart failure. Circulation [Internet] 2006 [cited 2018 Jan 7];113(13):1693–701. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16549636

51. Walsh C, Hripcsak G. The effects of data sources, cohort selection, and outcome definition on a predictive model of risk of thirty-day hospital readmissions. J Biomed Inform [Internet] 2014 [cited 2018 Jan 4];52:418–26. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25182868

52. Hanchate AD, Stolzmann KL, Rosen AK, et al. Does adding clinical data to administrative data improve agreement among hospital quality measures? Healthcare

[Internet] 2017 [cited 2018 Jan 8];5(3):112–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27932261

53. Grossman L, Reeder R, Walsh CG, Kansagara D, Salmasian H. Improving and Implementing Models to Predict 30-Day Hospital Readmissions: A Survey of Leading Researchers [Internet]. In: Academy Health. New Orleans: 2017 [cited 2018 Jan 12]. Available from: https://academyhealth.confex.com/academyhealth/2017arm/meetingapp.cgi/Paper/17304

54. van Walraven C, Dhalla IA, Bell C, et al. Derivation and validation of an index to predict early death or unplanned readmission after discharge from hospital to the community. CMAJ [Internet] 2010 [cited 2018 Jan 9];182(6):551–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20194559

55. van Walraven C, Wong J, Forster AJ. LACE+ index: extension of a validated index to predict early death or urgent readmission after hospital discharge using administrative data. Open Med [Internet] 2012 [cited 2018 Jan 12];6(3):e80-90. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23696773

56. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis [Internet] 1987 [cited 2018 Jan 16];40(5):373–83. Available from: http://www.ncbi.nlm.nih.gov/pubmed/3558716

57. Baldwin L-M, Klabunde CN, Green P, Barlow W, Wright G. In Search of the Perfect Comorbidity Measure for Use With Administrative Claims Data. Med Care [Internet] 2006 [cited 2018 Jan 16];44(8):745–53. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16862036

58. Goldstein LB, Samsa GP, Matchar DB, Horner RD. Charlson Index Comorbidity Adjustment for Ischemic Stroke Outcome Studies. Stroke [Internet] 2004 [cited 2018 Jan 16];35(8):1941–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15232123

59. Hemmelgarn BR, Manns BJ, Quan H, Ghali WA. Adapting the Charlson Comorbidity Index for use in patients with ESRD. Am J Kidney Dis [Internet] 2003 [cited 2018 Jan 16];42(1):125–32. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12830464

60. Lee DS, Donovan L, Austin PC, et al. Comparison of coding of heart failure and comorbidities in administrative and clinical data for use in outcomes research. Med Care [Internet] 2005 [cited 2018 Jan 16];43(2):182–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15655432

61. Myers RP, Quan H, Hubbard JN, Shaheen AAM, Kaplan GG. Predicting in-hospital mortality in patients with cirrhosis: Results differ across risk adjustment methods. Hepatology [Internet] 2009 [cited 2018 Jan 16];49(2):568–77. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19085957

62. Poses RM, McClish DK, Smith WR, Bekes C, Scott WE. Prediction of survival of critically ill patients by admission comorbidity. J Clin Epidemiol [Internet] 1996 [cited 2018 Jan 16];49(7):743–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8691223

63. Quan H, Sundararajan V, Halfon P, et al. Coding Algorithms for Defining Comorbidities in ICD-9-CM and ICD-10 Administrative Data. [cited 2018 Jan 16];Available from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.466.4756&rep=rep1&type=pdf

64. Backes Y, Schwartz MP, ter Borg F, et al. Multicentre prospective evaluation of real-time optical diagnosis of T1 colorectal cancer in large non-pedunculated colorectal polyps using narrow band imaging (the OPTICAL study). Gut [Internet] 2018 [cited 2018 Jan 11];gutjnl-2017-314723. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29298873

65. Song J, Yan Q, Zhang H, et al. Five key lncRNAs considered as prognostic targets for predicting Pancreatic Ductal Adenocarcinoma. J Cell Biochem [Internet] 2017 [cited 2018 Jan 11];Available from: http://www.ncbi.nlm.nih.gov/pubmed/29239017

66. Ramsay IS, Ma S, Fisher M, et al. Model selection and prediction of outcomes in recent onset schizophrenia patients who undergo cognitive training. Schizophr Res Cogn [Internet] 2018 [cited 2018 Jan 11];11:1–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29159134

67. Gedeborg R, Svennblad B, Byberg L, Michaëlsson K, Thiblin I. Prediction of mortality risk in victims of violent crimes. Forensic Sci Int [Internet] 2017 [cited 2018 Jan 11];281:92–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29125989

68. Endo A, Baer HJ, Nagao M, Weaver MJ. Prediction Model of In-Hospital Mortality After Hip Fracture Surgery. J Orthop Trauma [Internet] 2018 [cited 2018 Jan 11];32(1):34–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29076984

69. Zhang Z, Hong Y. Development of a novel score for the prediction of hospital mortality in patients with severe sepsis: the use of electronic healthcare records with LASSO regression. Oncotarget [Internet] 2017 [cited 2018 Jan 11];8(30):49637–45. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28548951

70. Buccheri S, Capodanno D, Barbanti M, et al. A Risk Model for Prediction of 1-Year Mortality in Patients Undergoing MitraClip Implantation. Am J Cardiol [Internet] 2017 [cited 2018 Jan 11];119(9):1443–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28274574

71. Gupta P, Rettiganti M, Gossett JM, Daufeldt J, Rice TB, Wetzel RC. Development and Validation of an Empiric Tool to Predict Favorable Neurologic Outcomes Among PICU Patients*. Crit Care Med [Internet] 2018 [cited 2018 Jan 11];46(1):108–15. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28991830

72. Uematsu H, Yamashita K, Kunisawa S, Otsubo T, Imanaka Y. Prediction of pneumonia hospitalization in adults using health checkup data. PLoS One [Internet] 2017 [cited 2018 Jan 11];12(6):e0180159. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28662167

73. Yan S, Tsurumi A, Que Y-A, et al. Prediction of Multiple Infections After Severe Burn Trauma. Ann Surg [Internet] 2015 [cited 2018 Jan 11];261(4):781–92. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24950278

74. Qian J, Payabvash S, Kemmling A, Lev MH, Schwamm LH, Betensky RA. Variable selection and prediction using a nested, matched case-control study: Application to hospital acquired pneumonia in stroke patients. Biometrics [Internet] 2014 [cited 2018 Jan 11];70(1):153–63. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24320930

75. Moon H, Zhao Y, Pluta D, Ahn H. Subgroup analysis based on prognostic and predictive gene signatures for adjuvant chemotherapy in early-stage non-small-cell lung cancer patients. J Biopharm Stat [Internet] 2017 [cited 2018 Jan 11];1–13. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29157115

76. Dente CJ, Bradley M, Schobel S, et al. Towards precision medicine. J Trauma Acute Care Surg [Internet] 2017 [cited 2018 Jan 11];83(4):609–16. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28538622

77. Schneider J, Hapfelmeier A, Thöres S, et al. Mortality Risk for Acute Cholangitis (MAC): a risk prediction model for in-hospital mortality in patients with acute cholangitis. BMC Gastroenterol [Internet] 2016 [cited 2018 Jan 11];16(1):15. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26860903

78. Taylor RA, Pare JR, Venkatesh AK, et al. Prediction of In-hospital Mortality in Emergency Department Patients With Sepsis: A Local Big Data-Driven, Machine Learning Approach. Acad Emerg Med [Internet] 2016 [cited 2018 Jan 11];23(3):269–78. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26679719

79. Mansoor H, Elgendy IY, Segal R, Bavry AA, Bian J. Risk prediction model for in-hospital mortality in women with ST-elevation myocardial infarction: A machine learning approach. Hear Lung J Acute Crit Care [Internet] 2017 [cited 2018 Jan 11];46(6):405–11. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28992993

80. LaBarbera FD, Nikiforov I, Parvathenani A, Pramil V, Gorrepati S. A prediction model for Clostridium difficile recurrence. J community Hosp Intern Med Perspect [Internet] 2015 [cited 2018 Jan 11];5(1):26033. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25656667

81. Liu G, Xu Y, Wang X, et al. Developing a Machine Learning System for Identification of Severe Hand, Foot, and Mouth Disease from Electronic Medical Record Data. Sci Rep [Internet] 2017 [cited 2018 Jan 11];7(1):16341. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29180702

82. Ambale-Venkatesh B, Yang X, Wu CO, et al. Cardiovascular Event Prediction by Machine LearningNovelty and Significance. Circ Res [Internet] 2017 [cited 2018 Jan 11];121(9):1092–101. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28794054

83. Vezzoli M, Ravaggi A, Zanotti L, et al. RERT: A Novel Regression Tree Approach to Predict Extrauterine Disease in Endometrial Carcinoma Patients. Sci Rep [Internet] 2017 [cited 2018 Jan 11];7(1):10528. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28874808

84. Pan L, Liu G, Lin F, et al. Machine learning applications for prediction of relapse in childhood acute lymphoblastic leukemia. Sci Rep [Internet] 2017 [cited 2018 Jan 11];7(1):7402. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28784991

85. Thottakkara P, Ozrazgat-Baslanti T, Hupf BB, et al. Application of Machine Learning Techniques to High-Dimensional Clinical Data to Forecast Postoperative Complications. PLoS One [Internet] 2016 [cited 2018 Jan 11];11(5):e0155705. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27232332

86. Kumar N, Shahjaman M, Mollah MNH, Islam SMS, Hoque MA. Serum and Plasma Metabolomic Biomarkers for Lung Cancer. Bioinformation [Internet] 2017 [cited 2018 Jan 11];13(6):202–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28729763

87. Kuo P-J, Wu S-C, Chien P-C, et al. Derivation and validation of different machine-learning models in mortality prediction of trauma in motorcycle riders: a cross-sectional retrospective study in southern Taiwan. BMJ Open [Internet] 2018 [cited 2018 Jan 11];8(1):e018252. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29306885

88. Ribas VJ, Lopez JC, Ruiz-Sanmartin A, et al. Severe sepsis mortality prediction with relevance vector machines [Internet]. In: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE; 2011 [cited 2018 Jan 11]. p. 100–3.Available from: http://www.ncbi.nlm.nih.gov/pubmed/22254260

89. Wang G, Lam K-M, Deng Z, Choi K-S. Prediction of mortality after radical cystectomy for bladder cancer by machine learning techniques. Comput Biol Med [Internet] 2015 [cited 2018 Jan 11];63:124–32. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26073099

90. Lotfnezhad Afshar H, Ahmadi M, Roudbari M, Sadoughi F. Prediction of Breast Cancer Survival Through Knowledge Discovery in Databases. Glob J Health Sci [Internet] 2015 [cited 2018 Jan 11];7(4):392–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25946945

91. van Walraven C, Dhalla IA, Bell C, et al. Derivation and validation of an index to predict early death or unplanned readmission after discharge from hospital to the community. Can Med Assoc J [Internet] 2010 [cited 2016 Apr 19];182(6):551–7. Available from: http://www.cmaj.ca/cgi/doi/10.1503/cmaj.091117

92. Jamei M, Nisnevich A, Wetchler E, Sudat S, Liu E. Predicting all-cause risk of 30-day hospital readmission using artificial neural networks. PLoS One [Internet] 2017 [cited 2018 Jan 9];12(7):e0181173. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28708848

93. Frizzell JD, Liang L, Schulte PJ, et al. Prediction of 30-Day All-Cause Readmissions in Patients Hospitalized for Heart Failure. JAMA Cardiol [Internet] 2017 [cited 2018 Jan 9];2(2):204. Available from: http://cardiology.jamanetwork.com/article.aspx?doi=10.1001/jamacardio.2016.3956

94. Mortazavi BJ, Downing NS, Bucholz EM, et al. Analysis of Machine Learning Techniques for Heart Failure Readmissions. Circ Cardiovasc Qual Outcomes [Internet] 2016 [cited 2017 Mar 19];9(6):629–40. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28263938

95. Jiang HJ, Stryer D, Friedman B, Andrews R. Multiple hospitalizations for patients with diabetes. Diabetes Care [Internet] 2003 [cited 2018 Jan 2];26(5):1421–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12716799

96. Nyenwe E, Loganathan R, Blum S, et al. Admissions for diabetic ketoacidosis in ethnic minority groups in a city hospital. Metabolism [Internet] 2007 [cited 2018 Jan 12];56(2):172–8. Available from: http://linkinghub.elsevier.com/retrieve/pii/S002604950600343X

97. McCoy RG, Lipska KJ, Herrin J, Jeffery MM, Krumholz HM, Shah ND. Hospital Readmissions among Commercially Insured and Medicare Advantage Beneficiaries with Diabetes and the Impact of Severe Hypoglycemic and Hyperglycemic Events. J Gen Intern Med [Internet] 2017 [cited 2018 Jan 2];32(10):1097–105. Available from: http://link.springer.com/10.1007/s11606-017-4095-x

98. Enomoto LM, Shrestha DP, Rosenthal MB, Hollenbeak CS, Gabbay RA. Risk factors associated with 30-day readmission and length of stay in patients with type 2 diabetes. J Diabetes Complications [Internet] 2017 [cited 2018 Jan 2];31(1):122–7. Available from: http://www.sciencedirect.com/science/article/pii/S1056872716307383?via%3Dihub

99. Caughey GE, Pratt NL, Barratt JD, Shakib S, Kemp-Casey AR, Roughead EE. Understanding 30-day re-admission after hospitalisation of older patients for diabetes: identifying those at greatest risk. Med J Aust [Internet] 2017 [cited 2018 Jan 12];206(4):170–5. Available from: https://www.mja.com.au/journal/2017/206/4/understanding-30-day-re-admission-after-hospitalisation-older-patients-diabetes

100. Ries Z, Rungprai C, Harpole B, et al. Incidence, Risk Factors, and Causes for Thirty-Day Unplanned Readmissions Following Primary Lower-Extremity Amputation in Patients with Diabetes. J Bone Joint Surg Am [Internet] 2015 [cited 2018 Jan 12];97(21):1774–80. Available from: http://insights.ovid.com/crossref?an=00004623-201511040-00006

101. Raval AD, Zhou S, Wei W, Bhattacharjee S, Miao R, Sambamoorthi U. 30-Day Readmission Among Elderly Medicare Beneficiaries with Type 2 Diabetes. Popul Health Manag [Internet] 2015 [cited 2016 Apr 19];18(4):256–64. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25608114

102. Albrecht JS, Hirshon JM, Goldberg R, et al. Serious mental illness and acute hospital readmission in diabetic patients. Am J Med Qual [Internet] 2012 [cited 2018 Jan 12];27(6):503–8. Available from: http://journals.sagepub.com/doi/10.1177/1062860612436576

103. Bennett KJ, Probst JC, Vyavaharkar M, Glover SH. Lower rehospitalization rates among rural Medicare beneficiaries with diabetes. J Rural Health [Internet] 2012 [cited 2018 Jan 12];28(3):227–34. Available from: http://doi.wiley.com/10.1111/j.1748-0361.2011.00399.x

104. Zapatero A, Gómez-Huelgas R, González N, et al. Frequency of Hypoglycemia and its Impact on Length of Stay, Mortality, and Short-Term Readmission in Patients with Diabetes Hospitalized in Internal Medicine Wards. Endocr Pract [Internet] 2014 [cited 2018 Jan 12];20(9):870–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24641928

105. Robbins JM, Webb DA. Diagnosing diabetes and preventing rehospitalizations: the urban diabetes study. Med Care [Internet] 2006 [cited 2018 Jan 2];44(3):292–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16501402

106. Rubin DJ, Handorf EA, Golden SH, Nelson DB, McDonnell ME, Zhao H. DEVELOPMENT AND VALIDATION OF A NOVEL TOOL TO PREDICT HOSPITAL READMISSION RISK AMONG PATIENTS WITH DIABETES. Endocr Pract [Internet] 2016 [cited 2018 Jan 10];22(10):1204–15. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27732098

107. Rubin DJ, Golden SH, McDonnell ME, Zhao H. Predicting readmission risk of patients with diabetes hospitalized for cardiovascular disease: a retrospective cohort study. J Diabetes Complications [Internet] 2017 [cited 2018 Jan 10];31(8):1332–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28571933

108. Collins J, Abbass IM, Harvey R, et al. Predictors of all-cause 30 day readmission among Medicare patients with type 2 diabetes. Curr Med Res Opin [Internet] 2017 [cited 2018 Jan 10];33(8):1517–23. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28498094

109. Denny JC, Bastarache L, Ritchie MD, et al. Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data. Nat Biotechnol [Internet] 2013 [cited 2018 Jan 15];31(12):1102–11. Available from: http://www.nature.com/articles/nbt.2749

110. Denny JC, Ritchie MD, Basford MA, et al. PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. Bioinformatics [Internet] 2010 [cited 2018 Jan 14];26(9):1205–10. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20335276

111. Averill RF, Goldfield N, Hughes JS, et al. All Patient Refined Diagnosis Related Groups (APR-DRGs) Version 20.0: Methodology Overview [Internet]. 2003 [cited 2018 Jan 15]. Available from: https://www.hcup-us.ahrq.gov/db/nation/nis/APR-DRGsV20MethodologyOverviewandBibliography.pdf

112. APR-DRG Service Intensity Weights and Average Length of Stay July 1, 2014 [Internet]. [cited 2018 Jan 15];Available from: https://www.health.ny.gov/facilities/hospital/reimbursement/apr-drg/weights/siw_alos_2014.htm

113. Roden D, Pulley J, Basford M, et al. Development of a Large-Scale De-Identified DNA Biobank to Enable Personalized Medicine. Clin Pharmacol Ther [Internet] 2008 [cited 2018 Feb 3];84(3):362–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18500243

114. Python Software Foundation [Internet]. [cited 2018 Feb 3];Available from: https://www.python.org/

115. Team RC. R: A Language and Environment for Statistical Computing. 2016

116. Health Innovation Program. Area Deprivation Index [Internet]. 2014 [cited 2018 Jan 15];Available from: https://www.hipxchange.org/ADI

117. Dumitrascu AG, Burton MC, Dawson NL, et al. Patient portal use and hospital outcomes. J Am Med Informatics Assoc [Internet] 2017 [cited 2018 Feb 3];Available from: http://academic.oup.com/jamia/advance-article/doi/10.1093/jamia/ocx149/4781349

118. Griffin A, Skinner A, Thornhill J, Weinberger M. Patient Portals. Appl Clin Inform [Internet] 2016 [cited 2018 Feb 3];7(2):489–501. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27437056

119. Harrell FE. Package “Hmisc” [Internet]. 2018 [cited 2018 Jan 18];Available from: https://cran.r-project.org/web/packages/Hmisc/Hmisc.pdf

120. Tibsharani R. Regression shrinkage and selection via the lasso.

121. Breiman L. Random Forests. Mach Learn [Internet] 2001 [cited 2018 Jan 21];45(1):5–32. Available from: http://link.springer.com/10.1023/A:1010933404324

122. Wright MN, Wager S, Probst P, Maintainer], De> <cran@wrig. Package “ranger” A Fast Implementation of Random Forests. 2018 [cited 2018 Jan 21];Available from:

https://cran.r-project.org/web/packages/ranger/ranger.pdf

123. Cortes C, Vapnik V. Support-vector networks. Mach Learn [Internet] 1995 [cited 2018 Jan 21];20(3):273–97. Available from: http://link.springer.com/10.1007/BF00994018

124. Meyer D, Dimitriadou E, Hornik K, et al. Misc Functions of the Department of Statistics, Probability Theory Group. 2017 [cited 2018 Jan 21];Available from: https://cran.r-project.org/web/packages/e1071/e1071.pdf

125. Steyerberg EW, Harrell FE, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J Clin Epidemiol [Internet] 2001 [cited 2018 Feb 4];54(8):774–81. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11470385

126. Jr FEH. rms: Regression Modeling Strategies [Internet]. 2015;Available from: http://cran.r-project.org/package=rms

127. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med [Internet] 1996 [cited 2018 Jan 23];15(4):361–87. Available from: http://doi.wiley.com/10.1002/%28SICI%291097-0258%2819960229%2915%3A4%3C361%3A%3AAID-SIM168%3E3.0.CO%3B2-4

128. Marshall A, Altman DG, Holder RL, Royston P. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. BMC Med Res Methodol [Internet] 2009 [cited 2018 Jan 21];9:57. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19638200

129. Chakravarti, Laha, Roy. Handbook of Method of Applied Statistics, Volume I [Internet]. John Wiley and Sons; 1967 [cited 2018 Jan 21]. Available from: http://www.itl.nist.gov/div898/handbook/eda/section4/eda43.htm#Chakravart

130. Sing T, Sander O, Beerenwinkel N, Lengauer T, Sing MT. Visualizing the Performance of Scoring Classifiers. 2015 [cited 2018 Jan 23];Available from: http://rocr.bioinf.mpi-sb.mpg.de/

131. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology [Internet] 2010 [cited 2018 Jan 21];21(1):128–38. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20010215

132. Drummond C, Holte RC. What ROC Curves Can't Do (and Cost Curves Can). [cited 2018 Jan 23];Available from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.204.7433&rep=rep1&type=pdf

133. Drummond C, Holte RC. Explicitly Representing Expected Cost: An Alternative to ROC Representation. [cited 2018 Jan 23];Available from: https://pdfs.semanticscholar.org/7d0f/75f61161666c4476bee122263c3ac52239e8.pdf

134. Bach FR, Org FB. Bolasso: Model Consistent Lasso Estimation through the Bootstrap. [cited 2018 Jan 25];Available from: https://www.di.ens.fr/~fbach/fbach_bolasso_icml2008.pdf

135. Nathan DM, Cleary PA, Backlund J-YC, et al. Intensive Diabetes Treatment and Cardiovascular Disease in Patients with Type 1 Diabetes. N Engl J Med [Internet] 2005 [cited 2018 Feb 10];353(25):2643–53. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16371630

136. Exner D V, Dries DL, Domanski MJ, Cohn JN. Lesser response to angiotensin-converting-enzyme inhibitor therapy in black as compared with white patients with left ventricular dysfunction. N Engl J Med [Internet] 2001 [cited 2018 Feb 10];344(18):1351–7. Available from: http://www.nejm.org/doi/abs/10.1056/NEJM200105033441802

137. Dries DL, Strong MH, Cooper RS, Drazner MH. Efficacy of angiotensin-converting enzyme inhibition in reducing progression from asymptomatic left ventricular dysfunction to symptomatic heart failure in black and white patients. J Am Coll Cardiol [Internet] 2002 [cited 2018 Feb 10];40(2):311–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12106937

138. Rathore SS, Foody JM, Wang Y, et al. Race, quality of care, and outcomes of elderly patients hospitalized with heart failure. JAMA [Internet] 2003 [cited 2018 Feb 10];289(19):2517–24. Available from: http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.289.19.2517

139. Mathew J, Wittes J, McSherry F, et al. Racial differences in outcome and treatment effect in congestive heart failure. Am Heart J [Internet] 2005 [cited 2018 Feb 10];150(5):968–76. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0002870305004059

140. Foraker RE, Rose KM, Suchindran CM, Chang PP, McNeill AM, Rosamond WD. Socioeconomic status, Medicaid coverage, clinical comorbidity, and rehospitalization or death after an incident heart failure hospitalization: Atherosclerosis Risk in Communities cohort (1987 to 2004). Circ Heart Fail [Internet] 2011 [cited 2018 Feb 20];4(3):308–16. Available from: http://circheartfailure.ahajournals.org/cgi/doi/10.1161/CIRCHEARTFAILURE.110.959031

141. Kind AJH, Jencks S, Brock J, et al. Neighborhood socioeconomic disadvantage and 30-day rehospitalization: a retrospective cohort study. Ann Intern Med [Internet] 2014 [cited 2018 Feb 20];161(11):765–74. Available from: http://annals.org/article.aspx?doi=10.7326/M13-2946

142. Kangovi S, Barg FK, Carter T, et al. Challenges faced by patients with low socioeconomic status during the post-hospital transition. J Gen Intern Med [Internet] 2014 [cited 2018 Feb 20];29(2):283–9. Available from: http://link.springer.com/10.1007/s11606-013-2571-5

143. Jiang X, Osl M, Kim J, Ohno-Machado L. Calibrating predictive model estimates to support personalized medicine. J Am Med Inform Assoc [Internet] 2012 [cited 2018 Feb 19];19(2):263–74. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21984587