University of South Florida

Digital Commons @ University of South Florida

2020

Prediction of Chronic Disease-Related Inpatient Prolonged Length of Stay Using Machine Learning Algorithms

Hasan Symum
University of South Florida

José Zayas-Castro
University of South Florida, josezaya@usf.edu

Follow this and additional works at: https://digitalcommons.usf.edu/egs_facpub

HIR

Healthcare Informatics Research

Prediction of Chronic Disease-Related Inpatient Prolonged Length of Stay Using Machine Learning Algorithms

Hasan Symum[1], José L. Zayas-Castro[1,2]
[1]Department of Industrial and Management System Engineering, University of South Florida, Tampa, FL, USA
[2]College of Engineering, University of South Florida, Tampa, FL, USA

Objectives: The study aimed to develop and compare predictive models based on supervised machine learning algorithms for predicting the prolonged length of stay (LOS) of hospitalized patients diagnosed with five different chronic conditions. **Methods:** An administrative claim dataset (2008–2012) of a regional network of nine hospitals in the Tampa Bay area, Florida, USA, was used to develop the prediction models. Features were extracted from the dataset using the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) codes. Five learning algorithms, namely, decision tree C5.0, linear support vector machine (LSVM), k-nearest neighbors, random forest, and multi-layered artificial neural networks, were used to build the model with semi-supervised anomaly detection and two feature selection methods. Issues with the unbalanced nature of the dataset were resolved using the Synthetic Minority Over-sampling Technique (SMOTE). **Results:** LSVM with wrapper feature selection performed moderately well for all patient cohorts. Using SMOTE to counter data imbalances triggered a tradeoff between the model's sensitivity and specificity, which can be masked under a similar area under the curve. The proposed aggregate rank selection approach resulted in a balanced performing model compared to other criteria. Finally, factors such as comorbidity conditions, source of admission, and payer types were associated with the increased risk of a prolonged LOS. **Conclusions:** Prolonged LOS is mostly associated with pre-intraoperative clinical and patient socioeconomic factors. Accurate patient identification with the risk of prolonged LOS using the selected model can provide hospitals a better tool for planning early discharge and resource allocation, thus reducing avoidable hospitalization costs.

Keywords: Length of Stay, Chronic Disease, Inpatients, Machine Learning, Discharge Planning

I. Introduction

In recent years, congestive heart failure (CHF), acute myocardial infarction (AMI), chronic obstructive pulmonary disease (COPD), pneumonia (PN), and type 2 diabetes (DB) have become the top most costly hospitalized conditions in the United States [1]. The majority of these conditions are characterized by longer than national average length of stay (LOS) of 4.5 days [2]. Moreover, in 2013, the number of hospitalizations for these conditions equaled 3.621 million stays (10.2% of inpatient admissions) [1]. Likewise, the average inpatient treatment costs incurred for these conditions were

high, between $7,400 and $18,400 per stay, compared to the national average [3]. Due to the recent substantial increase in medical costs and hospital expenditures, predicting the likelihood of prolonged LOS has become increasingly important to reduce the waste of hospital resources and improve patient satisfaction. Determining influential risk factors for prolonged LOS is useful for planning interventions or care management for patients with multiple chronic conditions. Furthermore, the prediction of prolonged LOS can improve the process of arranging a continuum of care for the patients, thus allowing family members to prepare for the return of their loved one. Additionally, under the government's new inpatient progressive payment system (IPPS), reimbursements are paid in fixed payments based on the patient's diagnosis-related group (DRG) rather than the volume of services [4].

Several studies have explored the use of various predictive models to improve performance in predicting LOS [5]. Multiple variations of artificial neural network (ANN)-based models have been applied in a variety of hospital settings (e.g., emergency department, psychiatric, and intensive care unit) [6,7]. Several other classification algorithms (e.g., support vector machine, logistic regression, and random forest) have also been applied for predicting LOS, and they have achieved diverse levels of accuracy [8,9]. However, these models did not use machine learning-based feature selection, anomaly detection, and class imbalance techniques in a framework that might result in overfitting and a weak learner. Only one study attempted to apply a class imbalance technique in predicting prolonged emergency department (ED) LOS [10]. Hence, the use of machine learning algorithms to predict condition-specific prolonged LOS needs further exploration. Therefore, a prolonged LOS prediction model is crucial and indispensable to healthcare providers, especially those with an alternative payment contract (e.g., accountable care organizations) with the Centers for Medicare and Medicaid Services (CMS). Thus, there is a need to develop a predictive decision support system that (1) identifies patients with prolonged LOS risk and (2) helps to develop individual discharge planning to reduce inpatient usage and eventually improve quality of care.

This study constructed and compared predictive models based on supervised machine learning algorithms to identify patients with the risk of prolonged LOS hospitalized with chronic conditions. Condition-specific prolonged LOS prediction represents a significant benchmark in providing healthcare providers a better tool to plan for discharge planning and resource allocation to reduce LOS; therefore,

it can lower hospitalization costs. We developed a robust framework for prolonged LOS prediction using data mining algorithms to extract important features, handle missing values, eliminate multicollinearity, detect outlier observations, and balance imbalanced class. Based on previous studies, we chose five algorithms: decision tree C5.0, linear support vector machine (LSVM), k-nearest neighbors (KNN), random forest (RF), and multi-layered ANNs. Twenty different model combinations for each cohort were constructed and compared in terms of several performance metrics.

II. Methods

1. Study Design

Prediction models were constructed using an administrative claim dataset provided by a network of nine hospitals geographically localized within three adjacent counties in the Tampa Bay region, Florida, USA. The types of hospitals in the study included general, teaching, and specialized hospitals. The initial dataset included 594,751 patients accounting for 1,093,177 patient discharges from January 2008 through July 2012. The five disease cohorts included in this study were AMI, CHF, COPD, DB, and PN. These conditions were identified by a primary diagnosis ICD-9 code for the inpatient claims. ICD-9 codes are used to identify hospital admission for AMI (codes 410.*), CHF (codes 428.*, 402.01, 402.91, 404.01, 404.03, 404.11, 404.13, 404.91, 404.93), COPD (codes 491.0, 491.1, 491.2, 491.20, 491.21, 490, 492, 496), DB (codes 250.*2), and PN (codes 480–483, 485–486, 510, 511.0, 511.1, 511.9, 780.6, 786.00, 786.05, 786.06, 786.07, 786.2, 786.3, 786.4, 786.5, 786.51, 786.52, 786.7). The final subsets of AMI, CHF, COPD, DB, and PN cohorts consisted of 10,983, 9,194, 7,189, 3,476, and 21,317 inpatient admissions, respectively.

For each discharge claim, we extracted 82 common features (including patient demographics, hospital information, and comorbidity) and several disease-specific features from the inpatient diagnosis and revenue codes based on insights from previous studies [5,11,12]. Descriptive statistics for the data and variables (common and cohort specific) are shown in Tables 1 and 2, respectively. Features were extracted from the diagnosis codes using ICD-9 numeric, E and V codes. For example, one of the features, accidental fall, was identified from 30 diagnoses ICD-9 codes by filtering E88–E89. The severity index was calculated as the severity of illness (from 1 = minor to 4 = extreme) defined by 3M all-patient refined–diagnosis-related groups (APR-DRG) [13].

Table 1. Descriptive statistics for all common features

Variable type	Variable name	Level	CHF (n = 9,194)	AMI (n = 10,983)	COPD (n = 7,189)	PN (n = 21,317)	DB (n = 3,476)
Outcome	Prolonged LOS	Yes	1,287 (14)	1,614 (15)	863 (12)	1,918 (9)	452 (13)
Demographics	Age (yr)	N/A	70.4 ± 13.7	65.1 ± 13.1	65.8 ± 13.7	47.9 ± 27.1	53.2 ± 16.2
	Gender	Male	4,443 (48)	6,448 (59)	4,161 (58)	11,231 (53)	1,785 (51)
		Female	4,751 (52)	4,535 (41)	3,028 (42)	10,086 (47)	1,691 (49)
	Race	Caucasian	6,884 (75)	9,010 (82)	6,129 (85)	13,891 (65)	1,881 (54)
		African American	1,413 (15)	691 (6)	607 (8)	3,370 (16)	1,061 (31)
		Hispanic	758 (8)	914 (8)	357 (5)	3,387 (16)	474 (14)
		Others	139 (2)	368 (4)	96 (1)	669 (3)	60 (1)
	Language	English	6,381 (69)	8,616 (78)	5,670 (79)	15,815 (74)	2,664 (77)
		Spanish	209 (2)	233 (2)	89 (1)	738 (3)	128 (4)
		Others	2,604 (29)	2,134 (20)	1,430 (20)	4,764 (23)	684 (19)
	Marital status	Single	2,019 (22)	2,508 (23)	1,731 (24)	10,384 (49)	1,350 (39)
		Married	3,651 (40)	5,615 (51)	2,504 (35)	6,306 (30)	1,223 (35)
		Others	3,524 (18)	2,860 (26)	2,954 (41)	4,627 (11)	903 (16)
Socioeconomic factor	Payer class	Public	7,739 (84)	6,768 (62)	5,711 (79)	14,333 (67)	2,001 (58)
		Private	39 (0)	122 (1)	53 (1)	254 (1)	25 (1)
		Managed care	823 (9)	2,750 (25)	723 (10)	4,527 (21)	707 (20)
		Uninsured	593 (7)	1,343 (12)	702 (10)	2,203 (11)	743 (21)
Hospital and admission factors	Number of previous admissions	N/A	2.8 ± 2.6	1.9 ± 1.8	3.3 ± 3.6	2.3 ± 2.9	3.3 ± 5.1
	Admission type	Emergency	8,037 (87)	8,660 (79)	6,459 (90)	17,571 (82)	2,978 (86)
		Routine	636 (7)	1,034 (9)	291 (4)	1173 (6)	213 (6)
		Others	521 (6)	1,309 (14)	439 (6)	2,573 (12)	285 (8)
	Admission source	ED	7,999 (87)	8,623 (78)	6,428 (89)	17,522 (82)	2,976 (85)
		Acute care	54 (1)	1,349 (12)	19 (0)	318 (1)	15 (0)
		Others	1,141 (12)	1,011 (10)	742 (11)	3,477 (17)	485 (15)
	Severity index	Minor	898 (10)	2,827 (26)	1,627 (22)	5,622 (26)	956 (27)
		Moderate	4,358 (47)	4,592 (42)	3,417 (48)	10,298 (48)	1,384 (40)
		Major	3,402 (37)	2,549 (23)	1,907 (27)	4,592 (22)	1,001 (29)
		Extreme	536 (6)	1,015 (9)	238 (3)	805 (4)	135 (4)
	Number of PX	N/A	1.1 ± 1.5	3.5 ± 1.9	0.5 ± 1.0	0.6 ± 1.1	0.9 ± 1.4
	Weekend admission	Yes	2,727 (13)	2,765 (25)	1,786 (24)	5,295 (25)	774 (21)
	Hospital location	Rural	1,655 (18)	1,089 (10)	1,078 (15)	3,839 (18)	418 (12)
		Urban	7,539 (82)	9,894 (90)	6,111 (85)	17,478 (82)	3,058 (88)
	Hospitalist	Yes	2,358 (26)	2,697 (27)	2,037 (28)	6,382 (30)	1,109 (32)

2. Outcome Variable

We defined prolonged LOS in our study as >7 days by calculating the 85th percentile threshold for the entire study population cohort's LOS [14]. The uniform prolonged LOS criteria (>7 days) for all cohorts were applied to simplify hospital resource allocation in discharge planning and to reduce the hospital-wise risk of post-discharge complications. Hospital stays longer than seven days are associated with a higher risk of post-discharge adverse outcomes and complications than short stay (≤7 days) regardless of admission causes [15].

Table 1. Continued 1

Variable type	Variable name	Level	CHF (n = 9,194)	AMI (n = 10,983)	COPD (n = 7,189)	PN (n = 21,317)	DB (n = 3,476)
Comorbidities	Number of chronic diseases	N/A	3.3 ± 1.6	2.1 ± 1.4	1.8 ± 1.2	1.1 ± 1.2	1.6 ± 1.9
	Infectious disease	Yes	180 (2)	104 (1)	262 (4)	742 (4)	228 (7)
	Neoplasm	Yes	21 (0)	25 (0)	8 (0)	0 (0)	2 (0)
	Immunity disorder	Yes	7,471 (82)	8,620 (78)	5,059 (70)	11,870 (56)	3,326 (96)
	Blood forming disease	Yes	3,285 (36)	2,732 (25)	1,927 (27)	5,737 (27)	998 (29)
	Mental disorder	Yes	3,633 (40)	4,606 (42)	4,558 (63)	7,658 (36)	1,647 (47)
	Nerve disease	Yes	2,127 (23)	1,763 (16)	1,841 (26)	4,905 (23)	1,400 (40)
	Rheumatic heart disease	Yes	947 (10)	376 (3)	101 (1)	310 (1)	28 (1)
	Hypertension	Yes	6,999 (76)	7,353 (67)	4,433 (62)	8,539 (40)	2,202 (63)
	Ischemic heart disease	Yes	2,400 (26)	9,949 (91)	794 (11)	1518 (7)	255 (7)
	Pulmonary disorder	Yes	1,551 (17)	418 (4)	461 (6)	549 (3)	49 (1)
	Other heart diseases	Yes	7,116 (77)	5,007 (46)	2,336 (32)	4,754 (22)	564 (16)
	Cerebrovascular disease	Yes	311 (3)	440 (4)	151 (2)	285 (1)	76 (2)
	Diseases of arteries	Yes	1,195 (13)	1,236 (11)	558 (8)	855 (4)	552 (16)
	Circulatory disorder	Yes	720 (8)	923 (8)	278 (4)	850 (4)	187 (5)
	Respiratory infections	Yes	435 (5)	268 (2)	572 (8)	2,335 (11)	129 (4)
	Influenza	Yes	792 (9)	455 (4)	909 (13)	16,474 (77)	57 (2)
	Chronic pulmonary disease	Yes	3,555 (39)	2,189 (20)	6,325 (88)	4,558 (21)	570 (16)
	Pneumoconiosis	Yes	1,842 (20)	1,348 (12)	1,465 (20)	4,558 (21)	134 (4)
	Stomach disease	Yes	1,672 (18)	1,806 (16)	1,634 (23)	3,650 (17)	850 (24)
	Oral disorder	Yes	25 (0)	25 (0)	23 (0)	91 (0)	26 (1)
	Gastroenteritis	Yes	667 (7)	592 (5)	632 (9)	1600 (8)	318 (9)
	Digestive system disorder	Yes	411 (4.47)	407 (3)	268 (4)	824 (4)	252 (7)
	Nephritis	Yes	0 (0)	12 (0)	7 (0)	0 (0)	0 (0)
	Kidney disease	Yes	1,097 (12)	751 (7)	458 (6)	1,525 (7)	338 (11)
	Diseases of productive organs	Yes	634 (7)	563 (5)	390 (5)	930 (4)	169 (5)
	Pregnancy complication	Yes	4 (0)	0 (0)	0 (0)	3 (0)	0 (0)
	Skin disease	Yes	692 (7)	357 (3)	283 (4)	1,083 (5)	1,046 (30)
	Musculoskeletal disease	Yes	2,524 (27)	2,317 (21)	2,297 (32)	4,824 (23)	1,135 (33)
	Congenital anomalies	Yes	144 (2)	86 (1)	45 (0)	633 (3)	32 (1)
	Perinatal abnormalities	Yes	3 (0)	7 (0)	3 (0)	133 (1)	0 (0)
	Ill conditions	Yes	2,411 (26)	1,941 (18)	2,282 (32)	9,740 (46)	1,119 (32)
	Abnormal symptoms	Yes	1,017 (11)	843 (8)	836 (12)	2,052 (10)	218 (6)
	Fracture	Yes	59 (1)	78 (1)	83 (1)	150 (1)	40 (1)
	Intracranial injury	Yes	39 (0)	44 (0)	4 (0)	37 (0)	5 (0)
	Open wounds	Yes	30 (0)	44 (0)	19 (0)	56 (0)	22 (1)

Table 1. Continued 2

Variable type	Variable name	Level	CHF (n = 9,194)	AMI (n = 10,983)	COPD (n = 7,189)	PN (n = 21,317)	DB (n = 3,476)
	Injury to blood vessels	Yes	0 (0)	2 (0)	0 (0)	0 (0)	0 (0)
	Superficial injury	Yes	16 (0)	25 (0)	11 (0)	34 (0)	14 (0)
	Entrance of foreign body	Yes	12 (0)	17 (0)	19 (0)	102 (0)	3 (0)
	Burns	Yes	1 (0)	4 (0)	2 (0)	4 (0)	2 (0)
	Spinal cord injury	Yes	2 (0)	12 (0)	3 (0)	10 (0)	0 (0)
	Poisoning	Yes	14 (0)	23 (0)	30 (0)	35 (0)	5 (0)
	Surgical care complication	Yes	207 (2)	851 (8)	50 (1)	188 (1)	69 (2)
	Cancer	Yes	25 (0)	13 (0)	4 (0)	17 (0)	10 (0)
	Artery disease	Yes	5,266 (57)	2,169 (20)	2,090 (29)	3,794 (18)	637 (18)
	Heart failure	Yes	9,194 (100)	8,303 (76)	1,165 (16)	1,997 (9)	261 (8)
	Vascular disease	Yes	1,090 (12)	2,363 (22)	512 (7)	758 (4)	440 (13)
	Chronic liver disease	Yes	516 (6)	348 (3)	171 (2)	373 (2)	1,730 (50)
	Diabetics with organ damage	Yes	833 (9)	413 (4)	132 (2)	309 (1)	184 (5)
	Renal failure	Yes	1,101 (12)	660 (6)	491 (7)	1,582 (7)	195 (6)
	Appendicitis	Yes	11 (0)	19 (0)	10 (0)	32 (0)	15 (0)
External causes and patient history	Accident	Yes	290 (0)	484 (4)	297 (4)	738 (1)	132 (3.8)
	Misadventure during care	Yes	160 (2)	680 (6)	49 (1)	231 (1)	55 (2)
	Falls	Yes	97 (1)	149 (1)	110 (1.5)	229 (1)	65 (2)
	Drug adverse effects	Yes	564 (6)	440 (4)	708 (10)	1,137 (5)	127 (4)
	Communicable disease history	Yes	182 (2)	153 (1)	166 (2)	451 (2)	45 (1)
	Treatment isolation	Yes	93 (1)	52 (0)	66 (1)	367 (2)	91 (3)
	History of cancer	Yes	665 (7)	584 (5)	485 (7)	1,041 (5)	83 (2)
	Mental disorder history	Yes	15 (0)	11 (0)	11 (0)	23 (0)	5 (0)
	Previous heart attack	Yes	361 (4)	285 (2)	215 (3)	444 (2)	72 (2)
	Allergic to drug	Yes	690 (8)	669 (6)	701 (10)	1,443 (7)	260 (7)
	Tobacco users	Yes	1,399 (15)	1,565 (14)	1,607 (22)	2,061 (10)	244 (7)
	History of non-compliance	Yes	772 (8)	407 (4)	312 (4)	359 (2)	635 (18)
	Organ replacement	Yes	587 (6)	264 (2)	220 (3)	546 (3)	58 (2)
	Machine dependence	Yes	3,439 (37)	2,253 (21)	1,690 (25)	3,338 (16)	402 (12)
	Heart surgery history	Yes	2,180 (24)	1,862 (20)	871 (15)	1,929 (9)	292 (8)
	Current drug users	Yes	1,819 (20)	1,230 (11)	1,062 (15)	1,880 (9)	1,054 (30)
	Homeless	Yes	143 (2)	180 (1)	182 (3)	301 (1)	98 (3)

Values are presented as number (%) or mean ± standard deviation. For the binary (Yes or No) type variable, descriptive statistics is shown for "Yes" level only.

CHF: congestive heart failure, AMI: acute myocardial infarction, COPD: chronic obstructive pulmonary disease, PN: pneumonia, DB: type 2 diabetes, LOS: length of stay.

Table 2. Descriptive statistics for disease cohort–specific variables

Disease cohorts	Variable name	Level	Total number (%)
CHF (n = 9,194)	Unspecified CHF	Yes	635 (7)
	Malignant CHF	Yes	134 (1)
	Systolic heart failure	Yes	1,383 (15)
	Diastolic heart failure	Yes	1,339 (15)
	Combined CHF	Yes	297 (3)
COPD (n = 7,189)	Farmer lungs	Yes	10 (0)
	Hazardous chemical exposure	Yes	7 (0)
	Pneumopathy	Yes	2 (0)
	Emphysema	Yes	155 (2)
	Extrinsic allergic alveolitis	Yes	12 (0)
	COPD with acute bronchitis	Yes	55 (1)
	COPD with asthma	Yes	65 (1)
	Exacerbation of COPD	Yes	6,499 (90)
AMI (n = 10,983)	STEMI elevated AMI	Yes	6,817 (62)
	Non-STEMI AMI	Yes	2,699 (25)
	Inferior wall AMI	Yes	1,568 (15)
	Anterior wall AMI	Yes	1,084 (10)
	Subendocardial infraction	Yes	6,817 (62)
	Initial episode of care	Yes	9,949 (91)
DB (n = 3,476)	Hyperglycemia	Yes	2 (0)
	Sepsis	Yes	39 (1)
	Urinary tract infection	Yes	395 (11)
	Wound/graft infection	Yes	10(0)
	Postoperative bleeding	Yes	35 (1)
	Hyperlipidemia	Yes	1,368 (40)
	Diabetes with neurological manifestations	Yes	1,342 (38)
PN (n = 21,317)	Sepsis	Yes	117 (0)
	Ventilator dependence	Yes	37 (0)
	Coma	Yes	2 (0)
	Dyspnea	Yes	222 (1)
	Dialysis	Yes	224 (1)
	Chronic condition with asthma	Yes	749 (3.5)
	Bacterial pneumonia	Yes	916 (4)
	Influenzas vaccine	Yes	415 (2)
	Pleural effusion	Yes	1,750 (8)
	Lung abscess	Yes	17 (0)
	Acute respiratory distress syndrome	Yes	3 (0)
	Hospital-acquired pneumonia	Yes	13,540 (63)

For the binary (Yes or No) type variable, descriptive statistics is shown for "Yes" level only.

CHF: congestive heart failure, COPD: chronic obstructive pulmonary disease, AMI: acute myocardial infarction, DB: type 2 diabetes, PN: pneumonia.

3. Modeling Framework

The modeling framework comprised three major steps: data preprocessing, model training, and performance evaluation. The data preprocessing comprised missing value handling, zero-variance test, correlation test, novelty detection, and feature selection. Figure 1 illustrates the steps involved in data preprocessing. Using the software RStudio caret packages [16], we performed the data preprocessing steps. First, missing values in the patient records were handled using established strategies. If the feature contained over 15% missing cases, we excluded the attributed feature. If less than 15% of the records were missing, the mean or median value replaced the blanks for the continuous and ordinal feature respectively.

We used the one-class support vector machine (O-SVM) to identify outliers from noisy observations [17]. The O-SVM
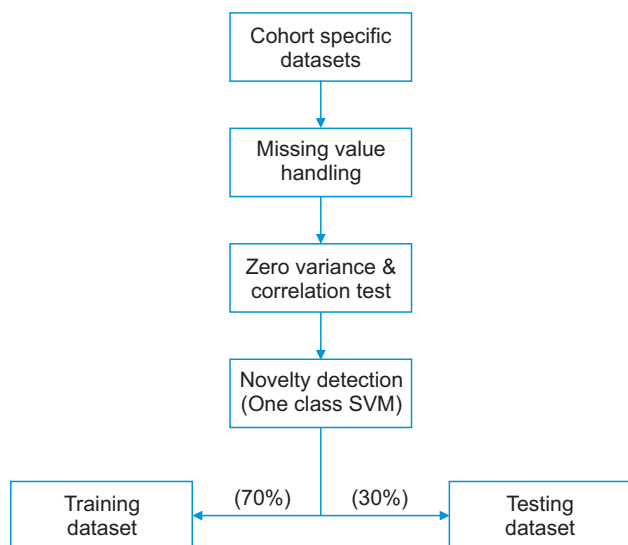
identified a similar proportion of anomalies (1.92%–2.44%) and excluded them from the dataset. Then we identified correlated features using the Pearson correlation and chi-square test for the continuous and nominal features, respectively, with a 0.05 level of significance. Among the correlated pairs of features, we dropped those with the highest variance inflation factor. Finally, features with in-class imbalances or zero variances were dropped after the zero-variance test with a 1.0% cutoff. Table 3 summarizes the results obtained from the data preprocessing steps.

We separated the records into training (70%), and testing (30%) sets for each cohort. Figure 2 illustrates the process of model building and the evaluation process. Next, using the same training dataset for each patient cohort, two different types of feature selection methods, chi-square filtering, and the SVM-based wrapper algorithm were applied to identify significant variables [18]. In the chi-square filtering method, features were selected at a 0.05 level of significance. For the wrapper algorithm, we limited our algorithm to a maximum of 200 iterations for each training model. The selected features from chi-square filtering and wrapper feature selection methods are shown in Supplementary Tables S1 and S2. After selecting features from both methods, we trained C5.0, LSVM, KNN, RF, and multi-layered ANN models for each cohort.

While training these models, we also explored the issues with the imbalance nature of the data. When training with imbalanced data, the algorithm tends to learn more from the majority class than the minority class, resulting in a weak learner with limited predictability. For the five cohorts, we had a varying imbalance ratio (0.09 to 0.15). To resolve this issue, we over-sampled the training data set using the Synthetic Minority Over-sampling Technique (SMOTE) and created new balanced data [19]. We trained 20 different models for each cohort and compared the performance of

Figure 1. Data preprocessing steps for building predictive model. SVM: support vector machine.

Table 3. Summary statistics in data preprocessing steps

Patient cohort	Number of observations	Anomalies	Total features	Correlation test rejection	Zero variance test rejection	Remained features
CHF	9,194	217 (2.36)	87	3	25	59
AMI	10,983	259 (2.35)	89	5	25	59
COPD	7,189	172 (2.39)	90	2	29	59
PN	21,317	410 (1.92)	97	2	36	59
DB	3,476	85 (2.44)	93	3	32	58

Values are presented as number (%).
CHF: congestive heart failure, AMI: acute myocardial infarction, COPD: chronic obstructive pulmonary disease, PN: pneumonia, DB: type 2 diabetes.
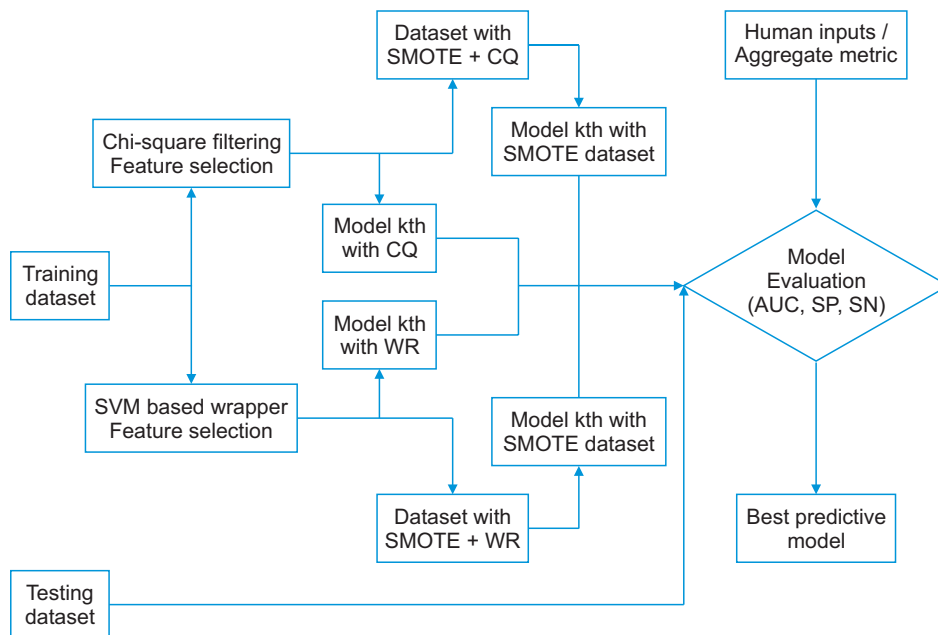
Figure 2. Flowchart of the predictive model building and best performing model selection. CQ: chi-square feature selection, WR: support vector machine-based wrapper feature selection, AUC: area under the curve, SP: specificity, SN: sensitivity, SMOTE: Synthetic Minority Oversampling Technique.

the models using the testing dataset under several metrics. Although the area under the curve (AUC) metric was unaffected by imbalances, the AUC tends to mask poor performance [20]. Therefore, we considered two other performance metrics, namely, sensitivity and specificity, with the AUC to minimize imbalance biases.

We propose a new rank average aggregate metric approach for selecting the best performing model to deal with the dilemma of performance tradeoff initiated by data imbalance. In our approach, the performance of each model was ranked separately for the three metrics of AUC, sensitivity, and specificity, and each was assigned a score (between 1 to 20) based on the rank. For example, if the LSVM model was ranked third by the AUC, we assigned a score of 18 out of 20. These three scores were multiplied by the assigned weights and summed to obtain a single aggregate metric where the summation of all weights must be equal to 1. Finally, we selected a single model by comparing the composite weighted sum metrics among the 20 different models, and the steps were repeated for each cohort.

III. Results

1. Assessment of the Prediction Models
We fitted 20 different model combinations comprising two feature selection methods, with or without SMOTE, and five learning algorithms for each cohort. Table 4 summarizes the performance of the learning algorithms for each cohort. As shown, the KNN models did not outperform any of the other algorithms. LSVM outperformed every other algo-

rithm in all cohorts with AUC, while it only outperformed CHF in terms of specificity. RF models outperformed for the AMI and DB cohorts, whereas for CHF and PN, the ANN models worked better according to the specificity metric. However, it is evident that selecting a model solely based on the AUC masked a model's poor specificity. For example, the best model for CHF under AUC is LSVM using the wrapper feature selection, with 0.81 AUC and 0.28 specificity. A specificity of 0.28, meaning only a 28% chance of detecting a true negative, represents poor model performance, which was completely shadowed by the AUC. Between the two feature selection methods, the SVM-based wrapper method yielded better prediction by AUC, whereas chi-square filtering methods achieved better true-negative rates. SMOTE used with feature selection did not improve the model's AUC. However, chi-square feature selection with SMOTE resulted in the highest specificity in all cohorts.

Figure 3 illustrates the changes in sensitivity and specificity with and without using SMOTE for the chi-square feature selection method. As shown, there is a significant tradeoff between the sensitivity and specificity. Furthermore, all the learning algorithms showed a positive tradeoff, specificity increase, and sensitivity decrease, except for the RF models. SMOTE yielded the highest and lowest increment of sensitivity for the C5.0 and KNN algorithms, respectively. The performance of each model depends on the algorithm as well as the feature selection and data balancing technique. Additionally, the tradeoff between the metrics due to the dataset imbalance makes it more challenging to select the best performing model. Table 5 shows the selected models based on

Table 4. Performance comparison of predictive models

Prediction model	AMI		CHF		COPD		DB		PN	
	AUC	SP	AUC	SP	AUC	SP	AUC	SP	AUC	SP
C5.0										
WR	0.83	0.44	0.75	0.24	0.72	0.22	0.82	0.46	0.84	0.40
WR+ST	0.84	0.58	0.74	0.46	0.72	0.40	0.82	0.67	0.82	0.64
CQ	0.80	0.43	0.75	0.26	0.72	0.15	0.82	0.49	0.78	0.35
CQ+ST	0.84	0.73	0.71	0.76[b]	0.75	0.63	0.80	0.76	0.81	0.81
KNN										
WR	0.77	0.29	0.65	0.21	0.60	0.02	0.81	0.44	0.75	0.19
WR+ST	0.73	0.36	0.65	0.09	0.57	0.12	0.81	0.45	0.70	0.25
CQ	0.78	0.39	0.62	0.25	0.70	0.18	0.83	0.56	0.75	0.27
CQ+ST	0.75	0.67	0.73	0.67	0.63	0.36	0.81	0.52	0.79	0.70
LSVM										
WR	0.88[a]	0.46	0.81[a]	0.28	0.79[a]	0.19	0.86[a]	0.56	0.88[a]	0.40
WR+ST	0.87	0.51	0.79	0.35	0.77	0.21	0.86	0.57	0.85	0.42
CQ	0.86	0.51	0.79	0.35	0.77	0.21	0.86	0.57	0.85	0.42
CQ+ST	0.87	0.74	0.79	0.73	0.77	0.64[b]	0.86	0.74	0.85	0.76
RF										
WR	0.86	0.66	0.76	0.69	0.71	0.46	0.83	0.66	0.84	0.71
WR+ST	0.82	0.49	0.72	0.26	0.64	0.17	0.82	0.43	0.78	0.47
CQ	0.83	0.71	0.74	0.67	0.74	0.61	0.81	0.80[b]	0.78	0.77
CQ+ST	0.80	0.84[b]	0.75	0.68	0.71	0.63	0.81	0.70	0.76	0.75
NN										
WR	0.85	0.46	0.81[a]	0.27	0.74	0.05	0.85	0.55	0.87	0.40
WR+ST	0.83	0.58	0.76	0.34	0.70	0.26	0.82	0.50	0.82	0.43
CQ	0.83	0.47	0.76	0.28	0.70	0.13	0.82	0.50	0.82	0.35
CQ+ST	0.86	0.78	0.79	0.70	0.76	0.60	0.85	0.79	0.85	0.80[b]

AMI: acute myocardial infarction, CHF: congestive heart failure, COPD: chronic obstructive pulmonary disease, DB: type 2 diabetes, PN: pneumonia, AUC: area under the curve, SP: specificity, KNN: k-nearest neighbor, LSVM: linear support vector machine, RF: random forest, NN: multi-layer neural network, WR: support vector machine-based wrapper method, WR+ST: wrapper method with SMOTE (Synthetic Minority Over-sampling Technique), CQ: chi-square filtering method, CQ+ST: chi-square with SMOTE.
[a]Best model based on AUC, [b]best model based on specificity.

AUC, sensitivity, F1 score, proposed aggregate rank, and a custom rule for each cohort. Based on our proposed metric, we selected several variations of LSVM models that showed balanced performance in every metric. In healthcare decision making, administrators or decision makers select the best model either empirically or by custom decision criteria favorable to the budgetary constraints. Therefore, we tested a custom rule comprising a minimum 0.75 specificity and the maximum for the AUC metric. If there was no model associated with more than 0.75 specificity, we selected the

final model with the highest specificity. In general, LSVM with the wrapper feature selection was selected based on the AUC criteria, while different algorithms with SMOTE were selected based on the specificity metric. The selected models obtained by custom rules comprised different machine learning algorithms with SMOTE, and the results showed a moderate performance across all metrics.

2. Important Features
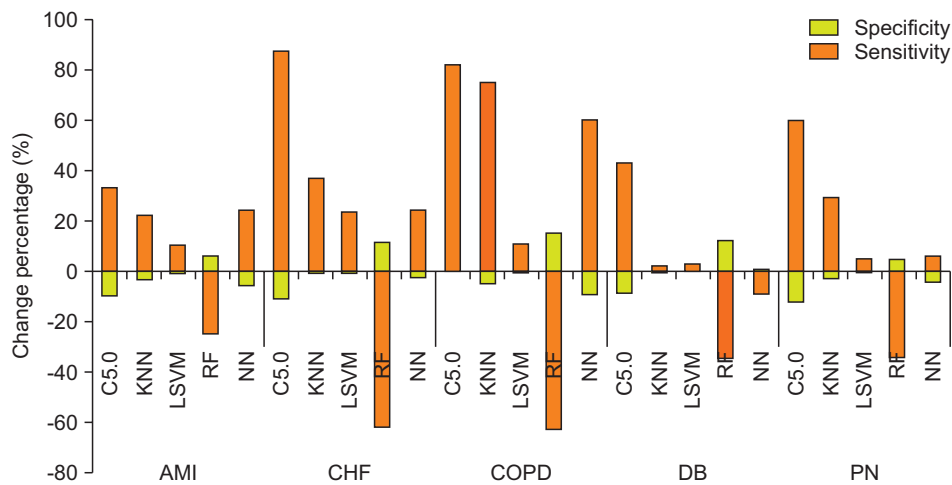Tables 6 and 7 show significant features using the LSVM

Figure 3. Performance metric changes (%) with and without SMOTE balancing. AMI: acute myocardial infarction, CHF: congestive heart failure, COPD: chronic obstructive pulmonary disease, DB: type 2 diabetes, PN: pneumonia, SP: specificity, SN: sensitivity, KNN: k-nearest neighbor, LSVM: linear support vector machine, RF: random forest, NN: multi-layer neural network, SMOTE: Synthetic Minority Over-sampling Technique.

Table 5. Best performing model based on several criteria

Selection criteria	Selected model for disease (AUC, SP, SN)				
	AMI	CHF	COPD	DB	PN
AUC	LSVM+WR	LSVM+WR	LSVM+WR	LSVM+WR	LSVM+WR
	(0.88, 0.46, 0.93)	(0.81, 0.28, 0.89)	(0.79, 0.19, 0.94)	(0.86, 0.56, 0.92)	(0.88, 0.40, 0.93)
SP	RF+CQ+ST	C5.0+CQ+ST	LSVM+CQ+ST	RF+CQ	NN+CQ+ST
	(0.80, 0.84, 0.80)	(0.71, 0.76, 0.85)	(0.77, 0.64, 0.86)	(0.81, 0.80, 0.88)	(0.85, 0.80, 0.83)
F1-score	LSVM+WR	LSVM+CQ+ST	LSVM+CQ+ST	KNN+WR+ST	LSVM+WR
	(0.88, 0.46, 0.93)	(0.79, 0.73, 0.91)	(0.77, 0.64 ,0.86)	(0.81, 0.52, 0.95)	(0.88, 0.40, 0.93)
Aggregate ranking	LSVM+CQ+ST	LSVM+CQ+ST	LSVM+CQ+ST	LSVM+CQ+ST	LSVM+WR
	(0.87, 0.74, 0.87)	(0.79, 0.35, 0.91)	(0.77, 0.64, 0.86)	(0.86, 0.74, 0.91)	(0.88, 0.40, 0.93)
Custom rule	NN+CQ+ST	C5.0+CQ+ST	LSVM+CQ+ST	LSVM+CQ+ST	NN+CQ+ST
	(0.86, 0.78, 0.86)	(0.71, 0.76, 0.85)	(0.77, 0.64, 0.86)	(0.86, 0.74, 0.91)	(0.85, 0.80, 0.83)

AMI: acute myocardial infarction, CHF: congestive heart failure, COPD: chronic obstructive pulmonary disease, DB: type 2 diabetes, PN: pneumonia, AUC: area under the curve, SP: specificity, SN: sensitivity, LSVM: linear support vector machine, RF: random forest, NN: multi-layer neural network, KNN: k-nearest neighbor, WR: support vector machine-based wrapper method, WR+ST: wrapper method with SMOTE (Synthetic Minority Over-sampling Technique), CQ: chi-square filtering method, CQ+ST: chi-square with SMOTE.
Example, a+b+c; a = machine learning method, b = feature selection technique, c = presence of SMOTE balancing; LSVM+CQ+ST = linear support vector mechanics with chi-square feature selection and SMOTE date balancing technique.

algorithm and regression analysis, respectively. The most important variable in all disease cohorts except DB for making a prolonged LOS prediction was the disease severity index with varying relative weights. In addition, the presence of different types of comorbidity was a strong predictor of prolonged LOS. For example, in AMI, COPD, and PN, the presence of comorbidities related to blood and blood-forming organ diseases resulted in a longer LOS. In addi-

tion, AMI and CHF patients admitted with pneumoconiosis and other lung-related conditions tended to stay longer in hospital inpatient settings. Several non-comorbidity-related features, such as the number of PX, source of admission, and payer class were also associated with a prolonged LOS. The number of tests (PX) required to assess patient condition was highly associated with prolonged inpatient stays in all cohorts expect AMI, and the higher the number of tests, the

Table 6. Important features extracted by LSVM algorithm

Rank	Cohort specific important features (relative importance weight)				
	AMI	CHF	COPD	DB	PN
1	Severity index (0.22)	Severity index (0.21)	Severity index (0.22)	Number PX (0.25)	Severity index (0.21)
2	Blood forming disease (0.15)	Number of PX (0.14)	Number of PX (0.21)	Severity index (0.21)	Number of PX (0.15)
3	Pneumoconiosis (0.10)	Pneumoconiosis (0.09)	Admission source (0.12)	Musculoskeletal disease (0.12)	Payer class (0.12)
4	Surgery complication (0.07)	Influenza (0.08)	Blood forming disease (0.09)	Skin disease (0.09)	Blood forming disease (0.1)
5	Non-disease related abnormal symptoms (0.04)	Skin disease (0.05)	Number of chronic diseases (0.08)	Chronic liver disease (0.06)	Number of chronic diseases (0.06)

LSVM: linear support vector machine, AMI: acute myocardial infarction, CHF: congestive heart failure, COPD: chronic obstructive pulmonary disease, DB: type 2 diabetes, PN: pneumonia.

greater the risk of a prolonged LOS. Specifically, PN patients with non-commercial payers and COPD patients admitted through the ED showed a greater likelihood of prolonged LOS.

IV. Discussion

In this study, we analyzed prolonged inpatient stays using an administrative claim dataset, performed extensive data preprocessing, and then developed and compared several variants of predictive models for the five disease cohorts. We identified several important factors that increase the risk of a prolonged stay for each disease. We found that prolonged LOS is associated with blood-forming and skin-disease-related comorbidities in most of the chronic conditions. The finding of several previous studies also support this result [21]. Some other studies have reported that patient demographics, gender, and hospital locations were contributors to identifying the risk of a prolonged LOS. However, our results do not conform to the findings in those studies [22,23]. One possible reason for this discrepancy is that the previous studies have mostly used homogenous data consisting of a single hospital or specific type of operation (e.g., knee replacement, heart surgery). When examining data from nine different hospitals over eight years, other factors had more weight than demographic factors (e.g., race, gender) in terms of the prediction of prolonged LOS.

Significant factors found in our study could be used to formulate individual disease-specific treatment pathways and early discharge planning to decrease inpatient LOS. Identifying patients with risk of prolonged LOS at the time of admission or inpatient care, the hospital can assign a dedicated

hospitalist and prepare a plan for the advanced discharged planning process. Studies show that having a dedicated hospitalist after four days of inpatient care and effective early discharged planning with a continuum of care can significantly reduce inpatient LOS [24,25]. Additionally, prioritizing laboratory tests and avoiding duplication of tests using hospital information exchange (HIE) can effectively decrease the LOS [26]. Furthermore, implementing improved care management and care coordination for patients with specific comorbidities in accountable care organizations (ACO) could reduce inpatient care utilization [27]. In addition, we found that the type of payer or insurance, which are typically considered to be significant socioeconomic factors, significantly affect the likelihood of a prolonged LOS. This insight agrees with the claims made in previous studies that social deprivation or economic inequality has a negative effect on the expected length of hospital stays of admitted patients [28]. Furthermore, an individual prolonged LOS risk profile can be used as a decision-making aid to the physician's subjective judgment while adjusting a patient's LOS [21]. This study of disease-specific prolonged LOS prediction may also assist in reducing the financial burden of the numerous outlier claims under CMS IPPS resulting from extended hospital stays. [29]. Outlier payments exert tremendous pressure on Medicare expenditures and are responsible for an average of $4.04 billion each year [30].

The prediction model we developed was compared to other published models in terms of predictive power and robustness. The selected cohort-specific models showed a variation of prediction performance depending on the model evaluation criteria. We found that although predictive power (AUC) was similar across certain methods, the range in de-

Table 7. Important features extracted using regression analysis

Variable name	Level	Odds ratio (mean and lower-upper bounds)				
		CHF	AMI	COPD	PN	DB
Number of previous admissions	N/A					1.3 (1.1–1.4)
Admission source	ED			1		
	Acute care			1.4 (1.1–1.8)		
	Others			0.6 (0.3–0.9)		
Severity Index	Minor	1	1	1	1	1
	Moderate	1.5 (1.2–1.8)	2.1 (1.6–2.7)	1.3 (1.1–1.6)	0.6 (0.4–0.8)	1.2 (1.0–1.4)
	Major	2.3 (2.1–2.6)	3.4 (3.2–4.1)	1.6 (1.3–1.9)	1.3 (1.1–1.5)	1.2 (1.0–1.5)
	Extreme	3.6 (3.1–4.1)	4.4 (3.8–5.1)	2.1 (1.9–2.5)	1.9 (1.7–2.1)	2.3 (1.9–2.7)
Number of PX	N/A			1.6 (1.5–1.7)		1.8 (1.6–1.9)
Payer class	Public				1	
	Private				0.8 (0.6–1.1)	
	Managed care				1.3 (1.0–1.6)	
	Uninsured				1.6 (1.4–1.8)	
Number of chronic diseases	Yes	1.2 (1.0–1.4)	1.1 (1.0–1.3)	1.3 (1.2–1.4)	1.2 (1.0–1.5)	1.2 (1.0–1.4)
Infectious disease	Yes			2.1 (1.8–2.5)		
Blood forming disease	Yes	1.7 (1.5–1.9)	4.0 (3.6–4.6)	1.6 (1.3–1.9)	2.7 (2.5–3.1)	1.3 (1.1–1.5)
Nerve disease	Yes		1.2 (1.1–1.4)			
Hypertension	Yes		1.8 (1.5–2.1)		1.9 (1.8–2.1)	1.8 (1.4–2.1)
Ischemic heart disease	Yes			1.4 (1.2–1.7)	1.3 (1.1–1.4)	
Pulmonary disorder	Yes			1.4 (1.2–1.6)		
Circulatory disorder	Yes	1.5 (1.1–1.5)				
Influenza	Yes	1.7 (1.4–1.9)	1.8 (1.4–2.3)			
Pneumoconiosis	Yes	1.9 (1.7–2.1)	4.2 (3.6–4.8)	1.7 (1.3-1.9)	1.7 (1.5–1.9)	
Digestive system disorder	Yes	1.4 (1.2–1.6)				
Skin disease	Yes	1.2 (1.0–1.5)				1.9 (1.5–2.4)
Musculoskeletal disease	Yes					1.8 (1.4–2.2)
Ill conditions	Yes	1.5 (1.3–1.8)	1.9 (1.7–2.2)			
Surgical care complication	Yes		2.8 (2.4–3.4)			
Artery disease	Yes					
Heart failure	Yes		1.3 (1.1–1.6)	1.12 (1.0–1.4)		
Vascular disease	Yes					1.2 (1.1–1.4)
Chronic Liver disease	Yes				1.6 (1.4–1.9)	1.3 (1.1–1.5)
DB with organ damage	Yes	1.3 (1.1–1.5)				
Drug adverse effects	Yes			1.5 (1.2–1.7)		
Bacterial pneumonia	Yes				5.7 (4.9–6.8)	
DB with neurological manifestations	Yes					1.9 (1.6–2.4)

CHF: congestive heart failure, AMI: acute myocardial infarction, COPD: chronic obstructive pulmonary disease, DB: type 2 diabetes, PN: pneumonia, ED: emergency department.

tecting true-positive and true-negative events varied greatly. Analyzing multiple aspects of models provides the health administrators or decision-makers a stronger understanding of those models and real-time applicability. LSVM models with wrapper feature selection showed overall better performance for all cohorts. Furthermore, integration of O-SVM for outlier detection in data preprocessing also improved model robustness when dealing with noisy observations. Implementation of the SMOTE technique along with feature-selection algorithms showed a significant tradeoff between sensitivity and specificity in all prediction models except RF, which made the final model selection based on a single performance metric difficult. Moreover, our results showed that using only the AUC as a baseline metric may mask a model's poor prediction performance, especially regarding the true-negative rate. Our proposed aggregate rank-based selection approach resolves this tradeoff dilemma by choosing a model with balanced performance, and it can provide a decision support tool to health administrators when comparing predictive models.

In conclusions, the accurate prediction of a prolonged LOS and prognosis of the risks associated with chronic disease are challenging. We adapted five machine-based learning techniques with feature selection, anomaly detection, and SMOTE balancing to predict prolonged LOS. The performance of the methods varies in complex ways, including discrimination and predictive range. We found that LSVM models performed better in terms of AUC and sensitivity. We also found that clinical and socioeconomic factors are the main features driving patient prolonged LOS. Designing predictive models would help to accelerate the stratification of patients according to prolonged LOS risk for improved care. The proposed prolonged LOS prediction model can be used to plan for advanced discharge planning, healthcare personnel allocation, and care coordination programs to reduce the usage of inpatient care. Some limitations of the present study should be addressed because they may restrict generalizability and are indicative of the need for further research. Our research did not include potential pathological (e.g., hemoglobin level) and sociocultural (e.g., education) features due to data availability, which might be useful for improving accuracy.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

ORCID

Hasan Symum (http://orcid.org/0000-0002-1525-624X)
José L. Zayas-Castro (http://orcid.org/0000-0001-7775-428X)

Supplementary Materials

References

1. Torio CM, Moore B. National inpatient hospital costs: the most expensive conditions by payer, 2013 [Internet]. Rockville (MD): Agency for Healthcare Research and Quality; 2006 [cited at 2020 Jan 10]. Available from: https://www.hcup-us.ahrq.gov/reports/statbriefs/sb204-Most-Expensive-Hospital-Conditions.pdf.

2. Weiss AJ, Elixhauser A. Overview of hospital stays in the United States, 2012 [Internet]. Rockville (MD): Agency for Healthcare Research and Quality; 2014 [cited at 2020 Jan 10]. Available from: https://www.hcup-us.ahrq.gov/reports/statbriefs/sb180-Hospitalizations-United-States-2012.pdf.

3. Pfuntner A, Wier LM, Steiner C. Costs for hospital stays in the United States, 2010 [Internet]. Rockville (MD): Agency for Healthcare Research and Quality; 2013 [cited at 2020 Jan 10]. Available from: https://www.hcup-us.ahrq.gov/reports/statbriefs/sb146.pdf.

4. Centers for Medicare & Medicaid Services (CMS), HHS. Medicare Program; Medicare Shared Savings Program; Accountable Care Organizations: pathways to success and extreme and uncontrollable circumstances policies for performance year 2017: final rules. Fed Regist 2018; 83(249):67816-8082.

5. Lu M, Sajobi T, Lucyk K, Lorenzetti D, Quan H. Systematic review of risk adjustment models of hospital length of stay (LOS). Med Care 2015;53(4):355-65.

6. Rowan M, Ryan T, Hegarty F, O'Hare N. The use of artificial neural networks to stratify the length of stay of cardiac patients based on preoperative and initial postoperative factors. Artif Intell Med 2007;40(3):211-21.

7. Wrenn J, Jones I, Lanaghan K, Congdon CB, Aronsky D. Estimating patient's length of stay in the Emergency Department with an artificial neural network. AMIA Annu Symp Proc 2005;2005:1155.

8. Morton A, Marzban E, Giannoulis G, Patel A, Aparasu R, Kakadiaris IA. A comparison of supervised machine learning techniques for predicting short-term in-hospital length of stay among diabetic patients. Proceedings of 2014 13th International Conference on Machine Learning and Applications; 2014 Dec 3-6; Detroit, MI. p. 428-31.

9. Jiang X, Qu X, Davis LB. Using data mining to analyze patient discharge data for an urban hospital. Proceedings of the 2010 International Conference on Data Mining (DMIN); 2010 Jul 12-15; Las Vegas, NV. p. 139-44.

10. Azari A, Janeja VP, Levin S. Imbalanced learning to predict long stay Emergency Department patients. Proceedings of 2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM); 2015 Nov 9-12; Washington, DC. p. 807-14.

11. Salah H. Predicting inpatient length of stay in western New York health service area using machine learning algorithms [dissertation]. Binghamton (NY): State University of New York at Binghamton; 2017.

12. Li JS, Tian Y, Liu YF, Shu T, Liang MH. Applying a BP neural network model to predict the length of hospital stay. In: Huang G, Liu X, He J, Klawonn F, Yao G, editors. Health information science. Heidelberg, Germany: Springer; 2013. p. 18-29.

13. Averill RF, Goldfield N, Hughes JS, Bonazelli J, McCullough EC, Mullin R, et al. 3M APR DRG classification system: methodology overview [Internet]. Wallingford (CT): 3M Health Information Systems; 2008 [cited at 2020 Jan 10]. Available from: https://www.hcup-us.ahrq.gov/db/nation/nis/v261_aprdrg_meth_ovrview.pdf.

14. Thompson B, Elish K, Steele R. Machine learning-based prediction of prolonged length of stay in newborns. Proceedings of 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA); 2018 Dec 17-20; Orlando, FL. p. 1454-9.

15. Allard JP, Keller H, Jeejeebhoy KN, Laporte M, Duerksen DR, Gramlich L, et al. Decline in nutritional status is associated with prolonged length of stay in hospitalized patients admitted for 7 days or more: a prospective cohort study. Clin Nutr 2016;35(1):144-52.

16. Kuhn M. Building predictive models in R using the caret package. J Stat Softw 2008;28(5):1-26.

17. Amer M, Goldstein M, Abdennadher S. Enhancing one-class support vector machines for unsupervised anomaly detection. Proceedings of the ACM SIGKDD Workshop on Outlier Detection and Description; 2013

Aug 11; Chicago, IL. p. 8-15.

18. Guyon I, Elisseeff A. An introduction to variable and feature selection. J Mach Learn Res 2003;3:1157-82.

19. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res 2002;16:321-57.

20. Jeni LA, Cohn JF, De La Torre F. Facing imbalanced data: recommendations for the use of performance metrics. Proceedings of 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction; 2003 Sep 2-5; Geneva, Switzerland. pp. 245-51.

21. Chuang MT, Hu YH, Lo CL. Predicting the prolonged length of stay of general surgery patients: a supervised learning approach. Int Trans Oper Res 2018;25(1):75-90.

22. Wang Y, Stavem K, Dahl FA, Humerfelt S, Haugen T. Factors associated with a prolonged length of stay after acute exacerbation of chronic obstructive pulmonary disease (AECOPD). Int J Chron Obstruct Pulmon Dis 2014;9:99-105.

23. Hachesu PR, Ahmadi M, Alizadeh S, Sadoughi F. Use of data mining techniques to determine and predict length of stay of cardiac patients. Healthc Inform Res 2013;19(2):121-9.

24. Gonçalves-Bradley DC, Lannin NA, Clemson LM, Cameron ID, Shepperd S. Discharge planning from hospital. Cochrane Database Syst Rev 2016;(1):CD000313.

25. Chandra S, Wright SM, Howell EE. The Creating Incentives and Continuity Leading to Efficiency staffing model: a quality improvement initiative in hospital medicine. Mayo Clin Proc 2012;87(4):364-71.

26. Menachemi N, Rahurkar S, Harle CA, Vest JR. The benefits of health information exchange: an updated systematic review. J Am Med Inform Assoc 2018;25(9):1259-65.

27. Kaufman BG, Spivack BS, Stearns SC, Song PH, O'Brien EC. Impact of accountable care organizations on utilization, care, and outcomes: a systematic review. Med Care Res Rev 2019;76(3):255-90.

28. Hasan O, Orav EJ, Hicks LS. Insurance status and hospital care for myocardial infarction, stroke, and pneumonia. J Hosp Med 2010;5(8):452-9.

29. Bai G, Anderson GF. US Hospitals are still using chargemaster markups to maximize revenues. Health Aff (Millwood) 2016;35(9):1658-64.

30. Krinsky S, Ryan AM, Mijanovich T, Blustein J. Variation in payment rates under Medicare's inpatient prospective payment system. Health Serv Res 2017;52(2):676-96.