ORIGINAL RESEARCH

Prediction of 30-Day Hospital Readmissions for All-Cause Dental Conditions using Machine Learning

Man Hung [1–5]
Wei Li [2]
Eric S Hon [6]
Sharon Su [1]
Weicong Su [7]
Yao He [8]
Xiaoming Sheng [9]
Richard Holubkov [10]
Martin S Lipsky [1]

[1]Roseman University of Health Sciences, College of Dental Medicine, South Jordan, UT, USA; [2]University of Utah, Department of Family and Preventive Medicine, Salt Lake City, UT, USA; [3]University of Utah, Department of Orthopaedics, Salt Lake City, UT, USA; [4]University of Utah, School of Business, Salt Lake City, UT, USA; [5]George E. Wahlen Veterans Affairs Medical Center, Salt Lake City, UT, USA; [6]University of Chicago, Department of Economics, Chicago, IL, USA; [7]University of Utah, Department of Mathematics, Salt Lake City, UT, USA; [8]University of Utah Alzheimer's Center, Salt Lake City, UT, USA; [9]University of Utah, College of Nursing, Salt Lake City, UT, USA; [10]University of Utah, Department of Pediatrics, Salt Lake City, UT, USA

Introduction: It is unknown whether patients admitted for all-cause dental conditions (ACDC) are at high risk for hospital readmission, or what are the risk factors for dental hospital readmission.

Objective: We examined the prevalence of, and risk factors associated with, 30-day hospital readmission for patients with an all-cause dental admission. We applied artificial intelligence to develop machine learning (ML) algorithms to predict patients at risk of 30-day hospital readmission.

Methods: This study used data extracted from the 2013 Nationwide Readmissions Database (NRD). There were a total of 11,341 cases for all-cause index admission for dental patients admitted to the hospitals. Descriptive statistics were used to analyze patient characteristics. This study applied five techniques to build risk prediction models and to identify risk factors. Model performance was evaluated using area under the receiver operating characteristic curve (AUC), and accuracy, sensitivity, specificity and precision.

Results: There were 11% of patients admitted for ACDC readmitted within 30 days of hospital discharge. On average, the total charge per patient was $131,004 for those with 30-day readmission (n=1254) and $69,750 for those without readmission (n=10,087). Factors significantly associated with 30-day hospital readmission included total charges, number of diagnoses, age, number of chronic conditions, length of hospital stays, number of procedures, Medicare insurance and Medicaid insurance, and severity of illness. Model performance from all methods was similar with the artificial neural network showing the highest AUC of 0.739.

Conclusion: Our results demonstrate that readmission after hospitalization with ACDC is fairly common. If one-third of the 30-day readmission cases can be avoided, there is a potential annual saving of over $25 million among the twenty-one states represented in the NRD. The ML algorithms can predict hospital readmission in dental patients and should be further tested to aid the reduction of hospital readmission and enhancement of patient-centered care.

Keywords: machine learning, dentistry, quality improvement, risk prediction, healthcare policy, precision medicine

Correspondence: Man Hung
Roseman University of Health Sciences College of Dental Medicine, 10894 S. River Front Parkway, South Jordan, UT 84095, USA
Tel +1801-878-1270
Email mhung@roseman.edu

Introduction

Hospital readmission is a common occurrence that can impose a multitude of burdens such as increased costs and an increased risk of mortality.[1] Studies indicate that the percentage of patients readmitted to hospitals in 30 days following discharge may be

as high as 20%.[2] Due to its cost and reoccurrence, in 2012 the Center for Medicare & Medicaid Services (CMS) instituted the Hospital Readmissions Reduction Program to reduce the frequency of rehospitalization by establishing financial penalties for hospitals with readmission rates deemed excessive.[3] The goal to reduce readmission spurred interest in studying and predicting hospital readmission.

Efforts to predict hospital readmission attempt to identify those at greatest risk and to uncover factors or trends leading to readmission. Due to the increasing availability of large databases, artificial intelligence (AI) and machine learning (ML) offer a potentially effective mean to predict risk factors influencing a variety of healthcare issues.[4] Unlike traditional modeling, ML utilizes higher-order and nonlinear interactions between predictors which could demonstrate better model performance,[5] and by identifying patterns in learning processes could help predict future events. Standard methods to predict hospital readmission typically rely on a limited set of clinical data using simple calculations such as a modified LACE score that exams at the length of hospital stay, acuity on admission, comorbidity and emergency department visits.[6] When comparing standard methods to ML, some researchers found that ML scores were better at predicting hospital readmissions.[6] However, others showed mixed results,[7] highlighting the need for further research to clarify discrepancies in predictive models.

Much of the published literature focused primarily on hospital readmission for surgical and other medical conditions, overlooking the importance of dental health. Globally, oral health and oral diseases such as dental caries, periodontal disease, lesions and oropharyngeal cancers are major public health concerns with poor oral health negatively affecting one's general health and quality of life.[8,9] In addition to its importance to overall well-being, in 2016 the United States spent more than $100 billion on dental expenditures, representing over a 3% increase from the previous year.[10,11] Despite potential clinical and financial impact, factors associated with hospital readmissions related to oral or dental issues remain unexplored.

This study aimed to explore the risk factors related to dental hospital readmission for patients with all-cause dental admissions (ACDA), and to build ML algorithms to predict whether an individual patient will be at risk of readmission within 30 days following hospital discharge for all-cause dental conditions (ACDC) . This predictive model created from artificial intelligence should be useful in decreasing risk and occurrences of hospital readmission following dental procedures.

Methods
Data
Data for this study came from the 2013 Nationwide Readmissions Database (NRD). The NRD is a part of the Healthcare Cost and Utilization Project and aims to provide nationally representative data on hospital readmissions for all types of payers and the uninsured. It contains data from approximately 85% of the State Inpatient Databases from 21 participating states in the United States. The 2013 NRD contains data for 14,325,172 hospital discharges. It has four distinct files (core file, severity file, diagnostic and procedure code file, hospital characteristics file) containing various data to enable more in-depth analyses.

In order to identify ACDA for inclusion into this study, we included patients with ICD-9-CM dental diagnostic codes (beginning with 52) and procedure codes (beginning with 23, 24 and 27, and some selected codes beginning with 96, 97 and 99). Appendix 1 shows the top dental diagnosis and procedure codes among all of them. This study excluded patients who died during a hospital stay, hospital stays less than one day, and those discharged after November 2013. Index events were identified by using inclusion and exclusion criteria stated above.

After applying the inclusion and exclusion criteria, 11,851 cases were identified for 30-day readmissions for ACDA patients in 2013. The initial dataset contained 236 variables, but some variables were eliminated resulting in a final total of 55 potential predictor variables and one outcome variable used for the analyses (Appendix 2). The eliminated variables included ID variables considered unlikely to be related to hospital readmission, variables showing extremely high correlation with existing ones, or variables used to weight for hospitals and discharges, as well as those that were used to filter specific patients. To prepare for further variable selection and preliminary analyses, the application of listwise deletion eliminated a small portion of cases with missing data (4.3%). A final total of 11,341 cases remained in the 30-day hospital readmission dataset for this study.

Outcome Variable
The primary outcome variable was dichotomous (yes/no on whether patients had readmissions within 30 days

following discharge of an index hospitalization with dental-related procedures). Altogether, 1254 cases (11.1%) had 30-day hospital readmissions for ACDA patients over the one-year study period.

Potential Predictor Variables

Based on literature review and consultation with an experienced dentist, the study defined three categories of information from the 55 potential predictor variables: (1) a demographic category such as age and gender, (2) a socioeconomic category such as median household income, primary payer, patient location, total charges, etc., and (3) a clinical category such as length of hospital stay, hospital urban-rural designation, emergency department service indicator, number of chronic conditions, number of diagnosis, number of procedures, AHRQ comorbidity measures, etc. Appendix 3 contains a detailed summary of the potential predictors across the Readmission groups.

Statistical Analyses

Traditional statistical analyses were used to further process the data in order to select a parsimonious set of variables for subsequent ML analyses. Continuous variables and categorical variables were analyzed to examine whether there were statistically significant differences at $\alpha=0.05$ between the readmission group and non-readmission group. For categorical variables, Fisher's exact test was used when the total sample size was less than 30 or the expected cell count was less than 5, or alternatively a simulated Chi-squared test applied when the dataset was too large for Fisher's exact test; otherwise the variables were tested using Chi-squared test. For continuous variables, a *t*-test was used if the variable was normally distributed; otherwise a Wilcoxon rank sum test was used. In sum, this selection process identified 49 variables (out of the 55) with statistically significant relationships to the outcome variable "Readmission" for use in subsequent feature selection by ML to build the 30-day readmission risk prediction models. All statistical analyses were performed using R 4.0.0 software.[12]

Machine Learning Methods

In preparing for the creation of preliminary risk prediction models using ML methods, the dataset was randomly split into two mutually exclusive sets, a training set with 70% of the data, and a testing set with 30% of the data. The training set was used for the machine (ie, computer) to learn the information contained in the data and then generate the risk prediction algorithms. The testing set was used to validate the performance of the prediction algorithms generated from the training set against the new data (eg, data unseen by the computer during the learning stage) contained in the testing set.

The data revealed an imbalanced binary outcome in which only 11.1% of patients had readmission and 88.9% did not. An imbalanced training set provides less information on the minority class (eg, 11.1%) and can bias prediction, thus potentially contributing to inaccuracy of risk prediction. In order to adjust for the imbalanced data and to increase accuracy, data resampling techniques were applied to balance the data. Specifically, oversampling of cases was applied to the minority class so that the dataset for ML has a balance of readmitted and did not readmit cases.

Data standardization and normalization are required for some ML methods such as k-Nearest Neighbor (k-NN), but not for the tree-based decision algorithms like decision trees (DT) or random forests (RF).[13] This study applied Z-score standardization for continuous predictors and dummy coding for the categorical variables. The standardized training dataset was utilized in the Support Vector Machine (SVM), k-Nearest Neighbor (k-NN) and Artificial Neural Network (ANN) ML methods as well as Logistic Regression (LR) in this study. We intentionally used LR (a traditional statistical technique) in this study so that we could compare its results with the ML methods.

RF, a tree-based ML algorithm, also known as an ensemble learning method, can be applied to both classification and regression tasks. A bagging (or bootstrap aggregating) technique was used in RF to build independent identically distributed trees, to grow deep trees, and to reduce the variance of an estimated prediction function. During the construction of RF, the out-of-bag error rate was calculated by averaging the prediction error of each unselected training sample through the bootstrap sampling for training. Aggregation of predictions can help reduce both bias and variance.

As a supervised learning technique (eg, the presence of predefined outcome variable to guide the learning process), one popular application of RF is important feature selection. Feature selection helps to reduce the dimensions without much loss of the information, leads to a decrease in training time and model complexity and prevents the data overfitting. This study used the randomForest package in R to select the important features for building risk prediction models using the ML models. As an ensemble

of individual DT, RF also uses the Gini Impurity to evaluate the importance of a variable on predicting the outcome across all of the individual trees in the forest. Relative importance of the features (eg, displaying the risk factors) was generated as a result.

The study applied five methods (DT, SVM, k-NN, ANN and LR) to build risk prediction algorithms. The 10-fold validation was then employed to estimate the model on test data and model performances on the testing set evaluated by using the confusion matrix (accuracy, sensitivity, specificity, precision, etc.) and the area under the receiver operating characteristic curve (AUC). According to United States federal regulations (45 CFR 46, category 4), this is a secondary analysis of existing data and the NRD data were de-identified and publicly available (see https://www.hcup-us.ahrq.gov/db/nation/nrd/NRD_Introduction_2013.jsp); thus, this research was exempt from Institutional Review Board approval. This study conforms to STROBE guidelines.

Results

Tables 1 and 2 summarize demographic data and compare key factors between the readmission and no readmission groups, respectively. Among the 11,341 patients who were discharged, 11.1% (n=1254) had hospital readmission within 30 days and 89.9% (n=10,087) did not. The sample consisted of 55.9% male (n=6338) and 44.1% female (n=5003). The mean age at admission was 42.4 years. The mean length of hospital stay was 7.9 days and the average total medical charges were $76,522. The average number of diagnoses for each patient was 9.7, and the average number of procedures was 3.8. The average number of chronic conditions was 3.7. There were 22.8% of patients with Medicare insurance (n=2590), 28.7% with Medicaid (n=3250) and 24.5% with private insurance (n=2781). Other groups included self-pay 15.8% (n=1792) and 8.1% other payers or with no hospital charges (n=928). Among the sample, 38.8% experienced minor loss of function (n=4399) and 33.9% (n=3845) moderate loss of function. Patients readmitted to the hospitals tended to be older, incur higher medical bills, have longer hospital lengths of stay, have more diagnoses, procedures and experience more chronic and severe illnesses. They also had a greater proportion of Medicare and/or Medicaid insurance (Table 2). The average hospital charge for a patient with 30-day readmission was $131,004, compared to $69,749 for those without readmission. If one-

Table 1 Demographic Characteristics of Patients

Variables	Summary (N=11,341)
Indicator of hospital readmission	
Yes	1254 (11.1%)
No	10,087 (88.9%)
Gender	
Male	6338 (55.9%)
Female	5003 (44.1%)
Age in years at admission	
Mean (SD)	42.4 (20.7)
Median (IQR)	43.0 (26.0, 57.0)
Range	(0.0, 90.0)
Length of stay	
Mean (SD)	7.9 (14.9)
Median (IQR)	4.0 (2.0, 8.0)
Range	(1.0, 340.0)
Total charges	
Mean (SD)	76,522.6 (141,182.8)
Median (IQR)	3,4260.0
	(18,978.0, 72,821.0)
Range	(473.0, 4580,711.0)
Number of diagnosis on this record:	
Mean (SD)	9.7 (6.4)
Median (IQR)	8.0 (5.0, 14.0)
Range	(1.0, 25.0)
Number of procedures on this record	
Mean (SD)	3.8 (3.1)
Median (IQR)	3.0 (2.0, 5.0)
Range	(1.0, 15.0)
Number of chronic conditions	
Mean (SD)	3.7 (3.3)
Median (IQR)	3.0 (1.0, 6.0)
Range	(0.0, 17.0)
Primary expected payer	
Medicare	2590 (22.8%)
Medicaid	3250 (28.7%)
Private insurance	2781 (24.5%)
Self-pay	1792 (15.8%)
No charge	209 (1.8%)
Other	719 (6.3%)
Severity of illness subclass	
Minor loss of function	4399 (38.8%)
Moderate loss of function	3845 (33.9%)
Major loss of function	2214 (19.5%)
Extreme loss of function	883 (7.8%)

(*Continued*)

Table 1 (Continued).

Variables	Summary (N=11,341)
Median household income	
$1 - $37,999	3851 (34%)
$38,000 - $47,999	2969 (26.2%)
$48,000 - $63,999	2529 (22.3%)
$64,000 or more	1992 (17.6%)

third of the readmitted cases (ie, 418 cases) could be prevented, it represents a cost saving of more than $25 million (ie, ($131,004 - $69,749) x 418 = $25,604,590) for the twenty-one states in the NRD.

Based on the largest contribution to decreasing Gini index, the top eight most important features selected by the random forest algorithm included total charges, number of diagnosis on this record, age in years at admission, number of chronic conditions, length of hospital stay, number of procedures on this record, primary expected payer, and severity of illness subclass (Figure 1). These eight selected features were then used to construct the risk prediction models using DT, SVM, k-NN, ANN and LR.

Figure 2 displays the 30-day hospital readmission risk prediction algorithms generated by DT. By using each individual patient's information about the top eight features, healthcare providers can follow the DT diagram in

Figure 2 to get an estimated readmission classification, and thus arrive at a conclusion of whether a patient will likely be readmitted 30 days within hospital discharge. Unlike DT, many ML methods produce highly complicated algorithms which act like a black-box and do not provide clarity as to how to arrive at the likelihood that a patient will be readmitted or not. Figure 2 also illustrates an example of such a black-box prediction algorithm developed from ANN.

Figure 3 depicts the AUC of the risk prediction models for the 30-day readmission from the five methods. High AUC value represents high prediction accuracy of the risk prediction models. The AUCs were similar across five methods using the eight predictors, with ANN having the highest AUC of 0.739, and LR having the second-highest AUC of 0.735. The precision was low across all methods at around 0.200. Figure 3 also graphically summarizes the performance indices of each method, including sensitivity, specificity, accuracy, precision and AUC.

Discussion

In 2009, Jencks et al[2] highlighted the issue of hospital readmissions in a milestone article. Using Medicare claims data, they found that about 1 in 5 patients discharged from the hospital were readmitted within 30 days. Both from a cost and quality perspective, the CMS made readmission

Table 2 Descriptive Summary of the Top Eight Predictors of 30-Day Readmission by Readmission and No Readmission Groups

Variable	Readmission (N=1254)	No Readmission (N=10,087)	P-value
TOTCHG (Total charges): Mean (SD)	131,004.0 (200,377.1)	69,749.6 (130,414.2)	<0.001[t]
AGE (Age in years at admission): Mean (SD)	49.3 (20.5)	41.5 (20.6)	<0.001[t]
LOS (Length of stay): Mean (SD)	13.7 (19.0)	7.2 (14.1)	<0.001[t]
NDX (Number of diagnosis on this record): Median (IQR)	14.0 (9.0, 19.0)	8.0 (4.0, 13.0)	<0.001[w]
NCHRONIC (Number of chronic conditions): Median (IQR)	6.0 (3.0, 8.0)	3.0 (1.0, 5.0)	<0.001[w]
NPR (Number of procedures on this record): Median (IQR)	4.0 (2.0, 7.0)	3.0 (2.0, 4.0)	<0.001[w]
PAY1 (Primary expected payer):			
Medicare	469 (37.4%)	2121 (21%)	<0.001[c]
Medicaid	394 (31.4%)	2856 (28.3%)	-
Private insurance	233 (18.6%)	2548 (25.3%)	-
Self-pay	98 (7.8%)	1694 (16.8%)	-
No charge	12 (1%)	197 (2%)	-
Other	48 (3.8%)	671 (6.7%)	-
APRDRG_Severity (Severity of illness subclass):			
Minor loss of function	187 (14.9%)	4212 (41.8%)	<0.001[c]
Moderate loss of function	376 (30%)	3469 (34.4%)	-
Major loss of function	464 (37%)	1750 (17.3%)	-
Extreme loss of function	227 (18.1%)	656 (6.5%)	-

Notes: [c]Chi-squared test; [t]t-test; [w]Wilcoxon rank sum test.
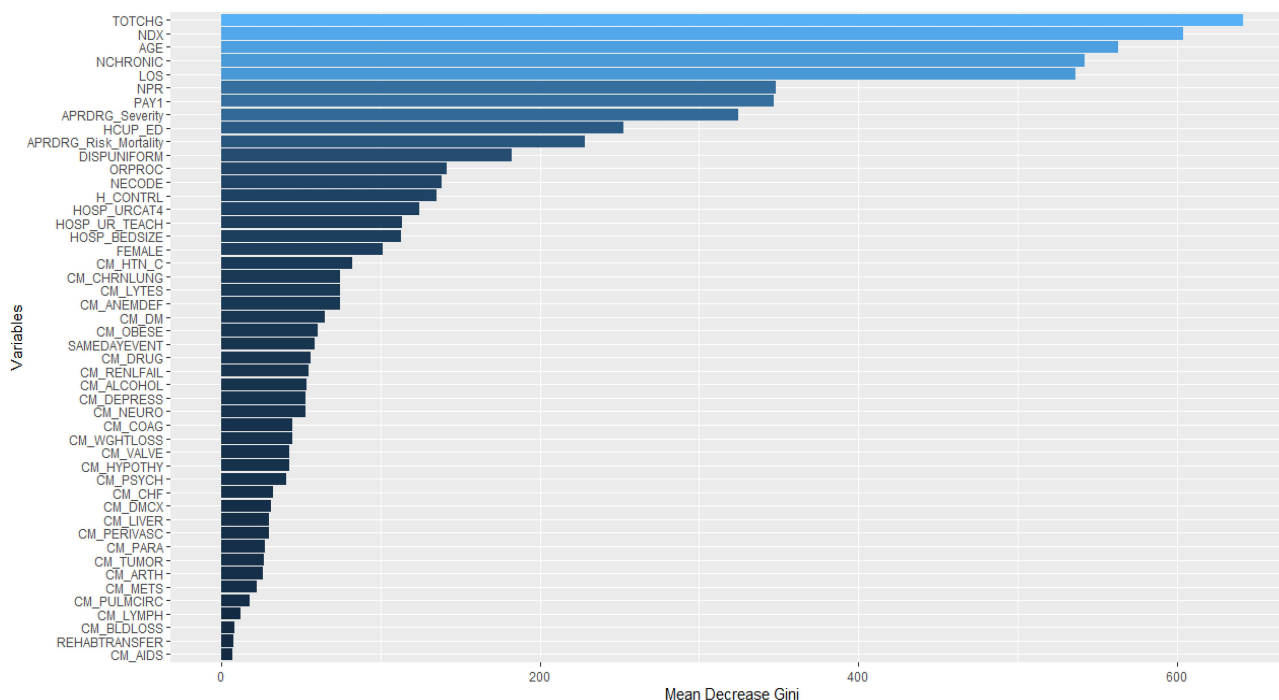
Figure 1 Important feature selection using random forest (The top eight features selected from random forest were used to build hospital readmissions risk prediction models. A higher "Mean Decrease in Gini" in x-axis indicates a higher purity (less noise, less bias) contributed by variable, and higher variable importance).

as focus, initially concentrating on heart disease and pneumonia, but more recently expanding to other conditions such as joint replacement surgery. Ideally, tools to forecast those at greatest risk for readmission could identify individuals most likely to benefit from interventions and to develop metrics to evaluate providers and organizations.

Little research exists about predictive models and ML tools for readmission for those hospitalized with ACDC. Likewise, research related to preventable dental related hospitalizations is also limited.[14] To our knowledge, this study represents the first in the United States to explore ACDC hospital readmissions and to apply AI to develop risk prediction tools. Our results indicate that about 11% of ACDC patients were readmitted within 30 days, a finding that helps establish a benchmark to assess interventions and for institutions and hospitals to evaluate their performance. In comparison, all-cause 30-day readmission rates vary between 7.1% and 14.4%.[15] Similar to other types of hospital readmission, those patients insured by Medicare and Medicaid experienced higher dental hospital readmission rates. The 11% 30-day dental hospital readmission rate found in this study certainly exceeds the 30-day rate for total hip arthroplasty (5.6%) and total knee arthroplasty (3.3%). Although the exact percent of preventable readmissions remains unclear, based on a median

estimate of 30% being preventable,[16] our study estimates that the annual cost-saving opportunity for 30-day dental hospital readmissions may exceed $25 million dollars for the twenty-one states in the NRD. With an average annual dental expenditure of $290.86 per person,[11] this $25 million represents dental care for over 88,000 individuals. Comparable to readmission research for heart disease and pneumonia, age, number of conditions, disease acuity and level of function correlated with an increased readmission risk for 30-day dental hospital readmissions. While intuitively it makes sense that "sicker" patients are at greater risk, this is the first study to document which co-morbidities correlate with ACDC hospital readmissions. Understanding risk factors can help guide strategies and research aimed at reducing readmission rates.

Despite numerous attempts to develop accurate prediction models for readmission, most models perform poorly.[17] In contrast, all of the five models in this study provided quite decent AUC, accuracy, sensitively and specificity. Precision was poor though, meaning that there could be large variation in future model performance. The poor precision is likely a result of the small number of predictors (ie, eight). Future effort can be focused on adding more predictors to increase the precision. After all, readmission likely represents a broad array of factors that
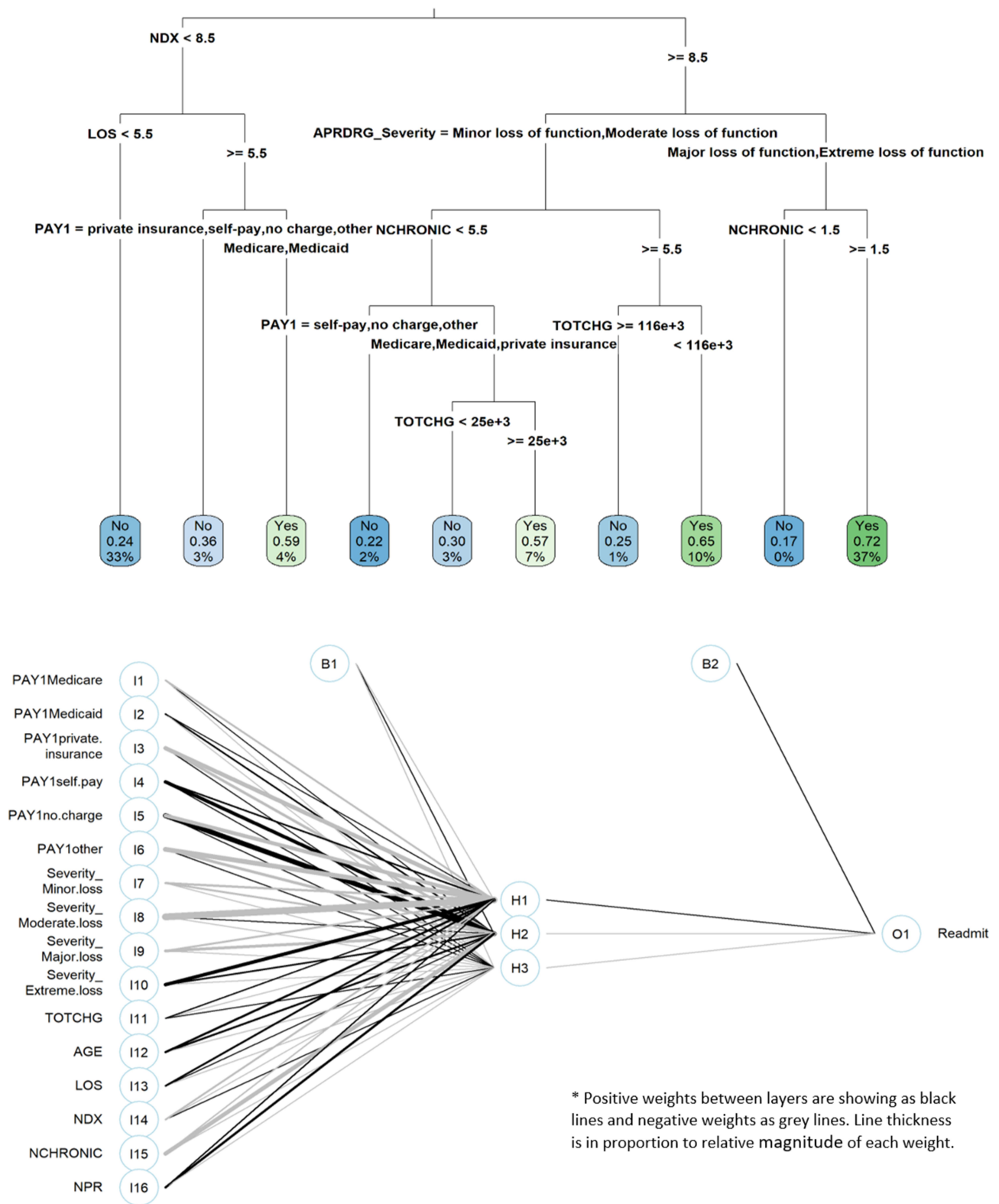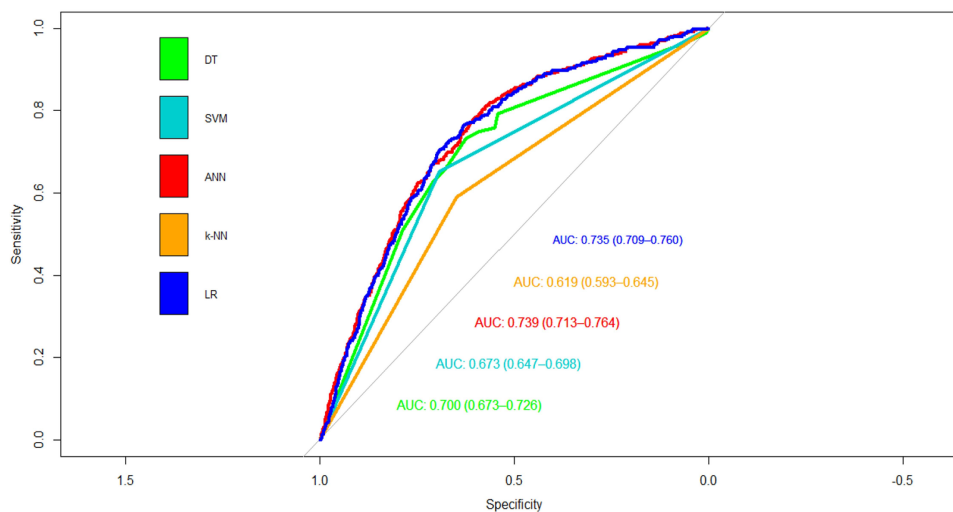
Figure 2 Risk prediction models for 30-day hospital readmission generated by decision tree (top diagram) and artificial neural network (bottom diagram).

interact in complex ways that challenge scientific discovery. Nonetheless, most predictive models rely on traditional analytic methods but this study is novel in its

application of ML methods from AI using only eight features to predict all-cause 30-day readmission for dental patients, and it can inform individualized patient treatment

ML Method	Accuracy	Sensitivity	Specificity	Precision	AUC
Decision Tree	0.636	0.733	0.624	0.196	0.700
Support Vector Machine	0.689	0.652	0.693	0.210	0.673
Artificial Neural Network	0.675	0.672	0.696	0.210	0.739
k-Nearest Neighbors	0.640	0.590	0.646	0.173	0.619
Logistic Regression	0.692	0.704	0.690	0.221	0.735

Figure 3 Performance metrics of the risk prediction models for 30-day hospital readmission when conducting model validation using various methods.

decision. The prediction models built by ANN provided better model performance. The DT, LR and other methods also had good AUCs. However, ML works best with large data sets and the small number of readmission cases and predictors might not provide robust enough data for ML techniques to outperform LR. We intentionally limited the final number of predictors to be very small in this study in order to produce a parsimonious model for practical use in clinical settings for all hospitals. Further studies using either a larger number of readmission cases or larger number of predictors in the final model will help confirm ML performance against traditional analyses such as LR and also help assess each model's applicability to other diagnoses. A surprising finding was that the length of stay for index admissions was unexpectantly long at almost eight days. However, a study examining hospital admissions in Florida observed admission rates and diagnostic codes for non-traumatic dental conditions (NTDC) (Morón et al, 2019) consistent with our findings suggesting that our study correctly identified dental admissions. Another study identified hospital charges of over $53,000 per NTDC admission (Tomar et al, 2019) suggesting a relatively high acuity level and complexity for ACDC admissions. Still the exact reasons for an average 8-day length of hospital stay remains uncertain and suggests an area of future research and perhaps new opportunities for identifying cost savings.

Limitations

Finding the best method to build a model depends on the setting, population and condition. While this study employed several ML methods, further fine-tuning using grid search to optimize hyperparameter combination (eg, using different or more predictors, adjusting the number of branches in a tree), or utilizing additional ML methods might improve in future models. Moreover, trying different feature selection sets and multiple feature combinations might improve prediction performance, but such a strategy would require extensive repeat testing and refinement. Future research can continue this effort to search and refine the models and algorithms. Another limitation was that because some numeric feature distributions were highly right skewed, log transformed data was used to produce better results. However, the trade-off in using log transformed data is the difficulty in how to directly interpret the results. Another limitation is that while the study examined risk factors, it did not explore which dental

conditions associated with readmission. While this study provides insights that can guide strategies to prevent readmissions, we have yet to carry out the next step in the future to fully realize the impact on clinical decisions. The next step will deploy the ML algorithms generated from this study by programming an app for mobile phone and/or computer desktop, and use the app for real-time clinical decision support in hospitals.

The NRD database represents a single year of data and so it is unknown if the data are consistent with other time periods, and any time trends cannot be identified. Another limitation is that the performance of the ML analyses was limited by the relatively small number of readmission cases. Also, the NRD is constructed from State Inpatient Databases and contains hospital readmission records from only twenty-one states, so readmissions occurring in other states are not captured. Another drawback is that race and ethnicity are not identified and thus the impact of these factors cannot be assessed.

Conclusion

The 30-day readmission rate for ACDC hospitalizations in our study is 11.1%. Readmission was associated with older age, higher number of diagnoses and chronic conditions, longer length of stay for the index hospitalization, higher severity of illness and having Medicare or Medicaid insurance. While intuitively it makes sense that "sicker" patients are at greater risk, this is the first study to provide empirical evidence that certain co-morbidities correlate with ACDC hospital readmissions. This is also the first study to establish a benchmark for institutions and hospitals to assess their performance using only a small number of eight predictors that are universal and applicable to many hospitals. Model performance from all methods was similar with ANN showing slightly better performance.

Disclosure

The authors report no conflicts of interest for this work.

References

1. Madrid-Garcia A, Font-Urgelles J, Vega-Barbas M, et al. Outpatient readmission in rheumatology: a machine learning predictive model of patient's return to the clinic. *J Clin Med*. 2019;8(8):1156. doi:10.3390/jcm8081156

2. Jencks SF, Williams MV, Coleman EA. Rehospitalizations among patients in the medicare fee-for-service program. *N Engl J Med*. 2009;360(14):1418–1428. doi:10.1056/NEJMsa0803563

3. Centers for Medicare & Medicaid Services. The Hospital Readmissions Reduction (HRR) program; 2017. Available from: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/HRRP/Hospital-Readmission-Reduction-Program.html. Accessed June 27, 2020.

4. Wiens J, Shenoy ES. Machine learning for healthcare: on the verge of a major shift in healthcare epidemiology. *Clin Infect Dis*. 2018;66 (1):149–153. doi:10.1093/cid/cix731

5. Golas SB, Shibahara T, Agboola S, et al. A machine learning model to predict the risk of 30-day readmissions in patients with heart failure: a retrospective analysis of electronic medical records data. *BMC Med Inform Decis Mak*. 2018;18(1):44.

6. Morgan DJ, Bame B, Zimand P, et al. Assessment of machine learning vs standard prediction rules for predicting hospital readmissions. *JAMA Netw Open*. 2019;2(3):e190348. doi:10.1001/jamanetworkopen. 2019.0348

7. Frizzell JD, Liang L, Schulte PJ, et al. Prediction of 30-day all-cause readmissions in patients hospitalized for heart failure: comparison of machine learning and other statistical approaches. *JAMA Cardiol*. 2017;2(2):204–209. doi:10.1001/jamacardio.2016.3956

8. Petersen PE, Bourgeois D, Ogawa H, Estupinan-Day S, Ndiaye C. The global burden of oral diseases and risks to oral health. *Bull World Health Organ*. 2005;83(9):661–669.

9. Hung M, Moffat R, Gill G, et al. Oral health as a gateway to overall health and well-being: surveillance of the geriatric population in the United States. *Spec Care Dentist*. 2019;39(4):354–361. doi:10.1111/scd.12385

10. Health Policy Institute. U. S. Dental Expenditures; 2017. Available from: https://www.ada.org/~/media/ADA/Science%20and%20Research/HPI/Files/HPIBrief_1217_1.pdf. Accessed June 28, 2020.

11. Hung M, Lipsky MS, Moffat R, et al. Health and dental care expenditures in the United States from 1996 to 2016. *PLoS One*. 2020;15 (6):e0234459. doi:10.1371/journal.pone.0234459

12. R Core Team. R: A language and environment for statistical computing; 2020. Available from: http://www.R-project.org. Accessed July 7, 2020.

13. Lesmeister C. *Mastering Machine Learning with R: Advanced Prediction, Algorithms, and Learning Methods with R*. 2nd ed. Birmingham, UK: Packt Publishing; 2017.

14. Acharya A, Khan S, Hoang H, Bettiol S, Goldberg L, Crocombe L. Dental conditions associated with preventable hospital admissions in Australia: a systematic literature review. *BMC Health Serv Res*. 2018;18(1):921. doi:10.1186/s12913-018-3733-2

15. Jain S, Khera R, Mortensen EM, Weissler JC. Readmissions of adults within three age groups following hospitalization for pneumonia: analysis from the nationwide readmissions database. *PLoS One*. 2018;13(9):e0203375. doi:10.1371/journal.pone.0203375

16. van Walraven C, Bennett C, Jennings A, Austin PC, Forster AJ. Proportion of hospital readmissions deemed avoidable: a systematic review. *CMAJ*. 2011;183(7):E391–E402. doi:10.1503/cmaj.101860

17. Kansagara D, Englander H, Salanitro A, et al. Risk prediction models for hospital readmission: a systematic review. *JAMA*. 2011;306 (15):1688–1698. doi:10.1001/jama.2011.1515