

Original Paper

# A Machine Learning Approach to Support Urgent Stroke Triage Using Administrative Data and Social Determinants of Health at Hospital Presentation: Retrospective Study

Min Chen<sup>1</sup>, BA, MPP, PhD; Xuan Tan<sup>2</sup>, BSc, MSc, PhD; Rema Padman<sup>3</sup>, BTECH, MS, PhD

<sup>1</sup>Department of Information Systems & Business Analytics, College of Business, Florida International University, Miami, FL, United States

<sup>2</sup>Department of Information Systems and Analytics, Leavey School of Business, Santa Clara University, Santa Clara, CA, United States

<sup>3</sup>The H John Heinz III College of Information Systems and Public Policy, Carnegie Mellon University, Pittsburgh, PA, United States

**Corresponding Author:**

Rema Padman, BTECH, MS, PhD

The H John Heinz III College of Information Systems and Public Policy

Carnegie Mellon University

4800 Forbes Avenue

Hamburg Hall 2101D

Pittsburgh, PA, 15213

United States

Phone: 1 412 268 2180

Email: [rpadman@cmu.edu](mailto:rpadman@cmu.edu)

## Abstract

**Background:** The key to effective stroke management is timely diagnosis and triage. Machine learning (ML) methods developed to assist in detecting stroke have focused on interpreting detailed clinical data such as clinical notes and diagnostic imaging results. However, such information may not be readily available when patients are initially triaged, particularly in rural and underserved communities.

**Objective:** This study aimed to develop an ML stroke prediction algorithm based on data widely available at the time of patients' hospital presentations and assess the added value of social determinants of health (SDoH) in stroke prediction.

**Methods:** We conducted a retrospective study of the emergency department and hospitalization records from 2012 to 2014 from all the acute care hospitals in the state of Florida, merged with the SDoH data from the American Community Survey. A case-control design was adopted to construct stroke and stroke mimic cohorts. We compared the algorithm performance and feature importance measures of the ML models (ie, gradient boosting machine and random forest) with those of the logistic regression model based on 3 sets of predictors. To provide insights into the prediction and ultimately assist care providers in decision-making, we used TreeSHAP for tree-based ML models to explain the stroke prediction.

**Results:** Our analysis included 143,203 hospital visits of unique patients, and it was confirmed based on the principal diagnosis at discharge that 73% (n=104,662) of these patients had a stroke. The approach proposed in this study has high sensitivity and is particularly effective at reducing the misdiagnosis of dangerous stroke chameleons (false-negative rate <4%). ML classifiers consistently outperformed the benchmark logistic regression in all 3 input combinations. We found significant consistency across the models in the features that explain their performance. The most important features are age, the number of chronic conditions on admission, and primary payer (eg, Medicare or private insurance). Although both the individual- and community-level SDoH features helped improve the predictive performance of the models, the inclusion of the individual-level SDoH features led to a much larger improvement (area under the receiver operating characteristic curve increased from 0.694 to 0.823) than the inclusion of the community-level SDoH features (area under the receiver operating characteristic curve increased from 0.823 to 0.829).

**Conclusions:** Using data widely available at the time of patients' hospital presentations, we developed a stroke prediction model with high sensitivity and reasonable specificity. The prediction algorithm uses variables that are routinely collected by providers and payers and might be useful in underresourced hospitals with limited availability of sensitive diagnostic tools or incomplete data-gathering capabilities.

(*J Med Internet Res* 2023;25:e36477) doi: [10.2196/36477](https://doi.org/10.2196/36477)

**KEYWORDS**

stroke; diagnosis; triage; decision support; social determinants of health; prediction; machine learning; interpretability; medical decision-making; retrospective study; claims data

## Introduction

### Background

Diagnostic errors have emerged as a major public health problem, contributing to preventable patient harm and excess health spending. A recent US National Academies report titled “Improving Diagnosis in Healthcare” suggested that medical misdiagnosis is likely to affect almost everyone at least once in their lifetime, sometimes with devastating consequences [1]. Misdiagnosis accounts for at least 40,000 to 80,000 hospital deaths and probably a comparable amount of disability annually in the United States [2]. Physician-reported errors and closed malpractice claims indicate that stroke is among the most common and dangerous misdiagnosed medical conditions [3-5]. Preventable deaths from stroke due to diagnostic errors occur  $\geq 30$  times more often than deaths from myocardial infarction [6,7].

The diagnosis of stroke is complicated by the abundance of stroke mimics and stroke chameleons. Approximately 30% of patients admitted to hospitals with typical stroke symptoms ended up having nonstroke conditions (ie, stroke mimics) [8]. A wide range of other medical conditions can exhibit symptoms that mimic strokes, such as seizures, migraines, psychiatric disorders, and drug or alcohol intoxication [8,9]. Mistaking a mimic for acute stroke may expose patients to unnecessary diagnostics and therapy, waste limited resources, and incur additional costs. Conversely, and more dangerously, stroke chameleons are actual stroke conditions presenting with atypical or underrecognized stroke symptoms and masquerading as nonstroke medical conditions. Approximately 25% of patients who had a stroke do not present with typical “face, arm, speech” symptoms at onset, and it is challenging for emergency medical services to identify stroke in such patients [10]. Misdiagnosis of chameleons can lead to significant delays in identifying and treating patients with actual strokes. Approximately one-third of potentially eligible patients failed to receive alteplase (tissue plasminogen activator), the gold-standard treatment for acute ischemic stroke [10]. This is because of either the lack of available specialists to perform appropriate clinical assessments or delays in the process of referring to health care facilities with the required stroke-handling capabilities [11]. In particular, Black people, Hispanic people, women, older people on Medicare, and people in rural areas are more prone to misdiagnosis and delay in receiving tissue plasminogen activator after having a stroke [10]. Furthermore, it is particularly challenging to accurately diagnose stroke in emergency departments (EDs) because of the time-sensitive and dynamic nature of emergency conditions, the fast-paced environment, frequent interruptions, the prevalence of information gaps, and high workload [12-15]. An automated screening tool that can be seamlessly integrated into the clinical workflow to quickly analyze the available information and suggest a diagnosis of stroke (“Stroke Alert” pop-up) could be very helpful [16].

Machine learning (ML), a crucial branch of artificial intelligence, has the potential to identify hidden insights from a large volume of data and generate predictions on unseen data (ie, test data) by iteratively learning from example inputs (ie, training data). ML problems can generally be divided into 3 main types: classification and regression, which are known as supervised learning, and unsupervised learning, which in the context of ML applications often refers to clustering. In the literature on stroke research, ML algorithms have been applied in different tasks, such as identifying factors associated with future stroke risk [17-19], developing stroke severity measures [20,21], and predicting stroke outcomes [22,23]. To improve diagnosis, researchers have focused on developing (electronic health record [EHR]-based) algorithms to determine stroke subtypes [24-26] and applying deep learning methods to facilitate imaging evaluation [27,28]. The recent advances in phenotyping algorithms and deep learning models have significantly improved the prediction for stroke by using multiple types of EHR data, especially clinical notes and advanced diagnostic tests. However, only a few investigations have focused on the application of diagnostic algorithms using ML in emergency triage when detailed clinical assessments and diagnostic tests are not readily available.

The first brain imaging for most patients with suspected stroke is a noncontrast computed tomography (CT) scan, which is completed within minutes of the arrival of the patient to the ED. However, a noncontrast CT scan is not sufficient to diagnose acute stroke, as the head CT test cannot reveal a hyperacute stroke in most cases, and it has reduced sensitivity for lacunar strokes [29]. More sensitive diagnostic tools such as diffusion-weighted magnetic resonance imaging can show ischemic changes very early. Despite the recent increase in the use of advanced neuroimaging, the use of magnetic resonance imaging to diagnose stroke in ED is still limited, especially when diagnosis is urgently required [30]. Moreover, patients who present to EDs can be susceptible to having information gaps because they are usually acutely ill, report quickly to the hospital at irregular hours, and often go to the ED without their primary physician’s knowledge. These factors make it difficult for the attending emergency physician to obtain all the information (eg, clinical notes, reports, and diagnostic test results) needed for making a timely and accurate diagnosis.

Besides medical risk factors, social determinants of health (SDoH) have been shown to be associated with the risk of stroke and many other diseases [31,32]. SDoH include various community and social factors, such as “conditions in which people are born, grow, work, live, and age” and “the fundamental drivers of these conditions” [33]. According to a widely used population health model, only 20% of an individual’s health is tied to clinical care, which includes access to care and the quality of health care services. The other 80% of an individual’s health is tied to their physical environment, social determinants, and behavioral factors such as exercise or smoking [34,35]. In recent years, the increasing focus on

population health has led to efforts to address upstream SDoH factors such as access to healthy food and viable transportation options. There is a substantial body of literature devoted to investigating the correlation between various SDoH factors and stroke risk, which has been well documented [36-39]. However, only a few studies have incorporated SDoH information into their prediction model and explicitly evaluated the added value of SDoH information for stroke diagnosis and triage [40]. There is a call, both in the literature and in the practitioner community, to explicitly evaluate whether and how SDoH data can contribute to improving patient risk stratification and prediction [40,41].

### Goal of This Study

In this study, we aimed to develop an ML stroke prediction algorithm based on data widely available at the time of patients' hospital presentations and to assess the added value of SDoH in stroke prediction. Because the prediction model does not require clinical notes or diagnostic test results, it might be particularly useful in addressing the misdiagnosis challenges faced when dealing with walk-in patients with stroke with milder and atypical symptoms; in low-volume or nonstroke centers' EDs, where emergency providers have limited daily exposure to stroke [16]; and in rural areas and small communities where there is limited availability of sensitive diagnostic tools and incomplete or unreliable data-gathering capabilities [3,5]. The model could also be applied in emergency medical services and telemedicine to seamlessly triage patients in real time and alert the provider and care team. In addition, we analyzed the most influential driving features helping the diagnosis of each patient and, specifically, the role of SDoH in prediction. The findings can provide insights into the value of prediction models in this critical setting and ultimately assist emergency care providers in making more informed decisions.

## Methods

### Ethics Approval

The secondary hospital discharge data this study examined was from the Healthcare Cost and Utilization Project State-specific databases, Agency for Healthcare Research and Quality. Healthcare Cost and Utilization Project databases conform to the definition of a limited data set, and review by an institutional review board is not required for use of limited data sets [42].

### Data Sources

Our data were obtained from 2 primary sources. We obtained longitudinal administrative data that contained encounter-level information on inpatient stays and ED visits from hospitals in the state of Florida. The second data source was the American Community Survey (ACS) conducted by the US Census Bureau [43]. The ACS data offered zip code-level SDoH information, such as demographic, social, housing, transportation, and other socioeconomic factors.

### Data Extraction and Synthesis

#### *The Stroke and Stroke Mimic Cohorts*

We adopted a case-control design, and the initial phase of our approach was to create representative examples for model training and ensure that stroke cases and controls have clear

separation. We retrospectively extracted 127,114 hospitalization records from 2012 to 2014 with a principal diagnosis of acute cerebrovascular disease in Florida using the clinical classification tool developed by the Agency for Healthcare Research and Quality [44]. Because we wanted to provide timely prediction of the likelihood of a patient's condition being stroke at the time of hospital presentation, we restricted attention to those variables that care providers can garner when patients first arrive at the hospital (eg, age, gender, race, admission time, primary payer, the number of chronic conditions on admission, etc). Thus, we excluded additional information that can be acquired only during hospitalization or at discharge (eg, procedures performed, length of stay, and total charges).

The key for a model to accurately predict stroke is to distinguish between stroke and stroke-like conditions ("stroke mimics"). We carefully created a stroke mimic data set to simulate tricky diagnostic decision-making and distinguish between actual stroke events and stroke-like events. Using all the records involving patients with nonstroke conditions to construct a prediction model will result in the inclusion of completely irrelevant cases, such as childbirth and hip replacement, and create a highly unbalanced data set. Hence, we consulted physicians about what conditions may show initial symptoms similar to those of a patient with stroke. On the basis of their suggestions, we obtained a list of conditions using Epocrates, a mobile app that health care providers use at the point of care for clinical reference information [45]. The stroke mimics included in the list were brain tumors, conversion and somatization disorders, Wernicke encephalopathy, seizure and postictal deficits, complicated migraines (hemiplegic migraines and migraines with aura), hypoglycemia, and hypertensive encephalopathy. Next, we searched the medical literature to confirm the validity of the list of stroke mimics and built a crosswalk between each stroke mimic and its corresponding International Classification of Diseases, ninth revision, codes. We then used the crosswalk to extract patients whose reasons for visits were one or more of the stroke mimics but subsequent discharge diagnoses were not stroke.

We pooled the stroke and stroke mimic data sets and retained only the data collected during the first admission of the patients. We performed data deduplication once again after combining stroke data and stroke mimic data because a patient may have been first admitted with stroke and readmitted with a stroke mimic condition and vice versa. If a patient appeared in both data sets, we kept only the first occurrence. Because patients may have returned to the hospital multiple times, providers may have obtained more information about patients who are readmitted. Retaining only the index encounter of the patients ensures that our models predict stroke based solely on the information available at the time of a patient's initial presentation at the hospital. We obtained data from 2010 to 2014, and hence we have 2 years before 2012 as our "cushion period." The patients included in the analysis were those with no records in 2010 or 2011. The "confirmed stroke" data set contains all the patients whose hospital discharge records confirmed that they had a stroke; thus, it includes not only patients with typical stroke symptoms but also those with mild and atypical symptoms. The stroke mimic data set includes

patients with general presentations similar to those of patients with actual stroke, including patients with a discharge diagnosis of epilepsy, diabetes, alcohol, and drug withdrawal. [Multimedia Appendix 1](#) lists the distribution of the top 20 principal diagnoses in the final analysis data set.

### Feature Extraction and Selection From SDoH Data

The original SDoH data we extracted from the ACS contained a large number of features. We adopted several methods to reduce noise and dimensionality and avoid overfitting. First, we conducted exploratory data analysis such as the principal component analysis to understand the feature distribution and identify patterns and multicollinearity among features. We then combined domain knowledge and a sparse regression method (least absolute shrinkage and selection operator) to remove irrelevant features and merge highly sparse features.

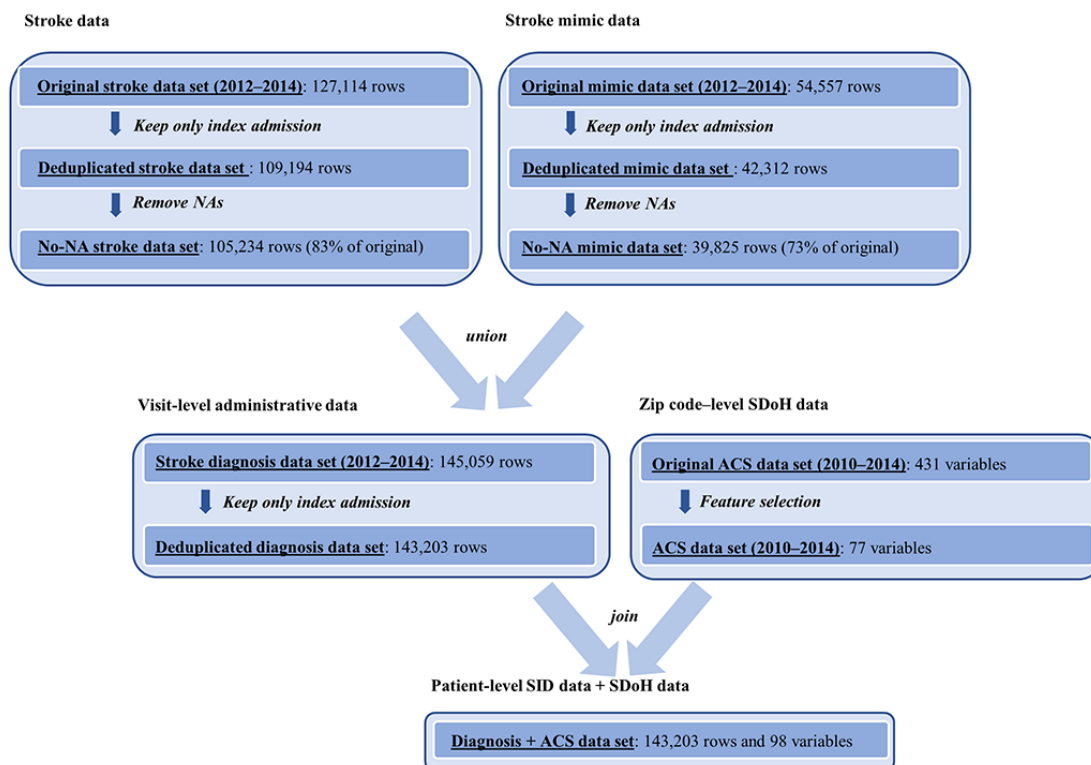
Overall, 4 categories were constructed from a large set of 431 variables in the ACS data for the 983 zip codes in Florida. These categories represent social, economic, housing, occupation, health insurance, and demographic characteristics referenced in the literature as being associated with stroke-related and cardiovascular health status ([Multimedia Appendix 2](#)). For example, low income, low education, and poverty have been shown to result in a higher risk of stroke [31,46,47], and low income and low education have been associated with lower heart health and higher risk of heart failure and death [48,49]. Occupation type and education level have been linked to the risk of heart disease [50]. Health insurance status and type have

also been linked to cardiovascular health [51,52]. Together, these interlinked socioeconomic factors determine a person’s overall socioeconomic status and, unsurprisingly, have a relationship with health over time. Some of the ACS variables included in the analysis are direct representations of socioeconomic status (eg, average household income and percentage of the population with at least high school-level education), whereas others serve as proxies (eg, percentage of housing units with no vehicles and percentage of the population with a non-English language spoken at home).

We also performed a Markov blanket feature selection method to determine a minimal subset of relevant features that yields the optimal classification performance [53]. Note that tree-based ML algorithms (eg, random forest [RF]) have a built-in feature selection function and inherently eliminate irrelevant features during model training.

The final analysis data set was formed by merging the patient-level data with the community-level ACS data based on the patients’ zip code information. [Figure 1](#) presents a flowchart of the data processing pipeline. In our final input data set, the number of stroke cases was significantly larger than the number of controls (ie, 73% of the patients were discharged with a confirmed stroke diagnosis, and 27% of the patients ended up having stroke mimics). To address the unbalanced distribution of stroke events in the real-world data, we adopted adaptive synthetic sampling, an oversampling technique for minority class (eg, nonstroke “control”) in the training data [25,54-56].

**Figure 1.** Data processing pipeline. ACS: American Community Survey; NA: not available; SDoH: social determinants of health; SID: State Inpatient Database.



### Data Modeling and Validation

We started by using the patient-level information available at the time of hospital presentation to predict a binary outcome

that indicates whether the patient’s final diagnosis at discharge is stroke. We ran three different models that are well established in the literature for the training process: (1) logistic regression, (2) RF, and (3) gradient boosting machine (GBM). Each model

was run with different combinations of predictor variables to assess the added predictive value of the different variables.

Logistic regression is a popular method for modeling the relationship between a set of predictor variables and a binary outcome variable and for benchmarking [57]. RF is a supervised learning algorithm that fits multiple decision trees on different subsamples of data to classify outcomes to prevent the issue of overfitting [58,59]. The predictive accuracy is the average of all the decision trees. It also provides insights into relative feature importance. Parameter tuning helped identify the number of trees and the allowed depth for each tree in the RF that provided the best performance. GBM is similar to RF, as it also constructs multiple decision trees for prediction; however, the difference lies in the way GBM builds the trees and the way it combines the results from the decision trees [60].

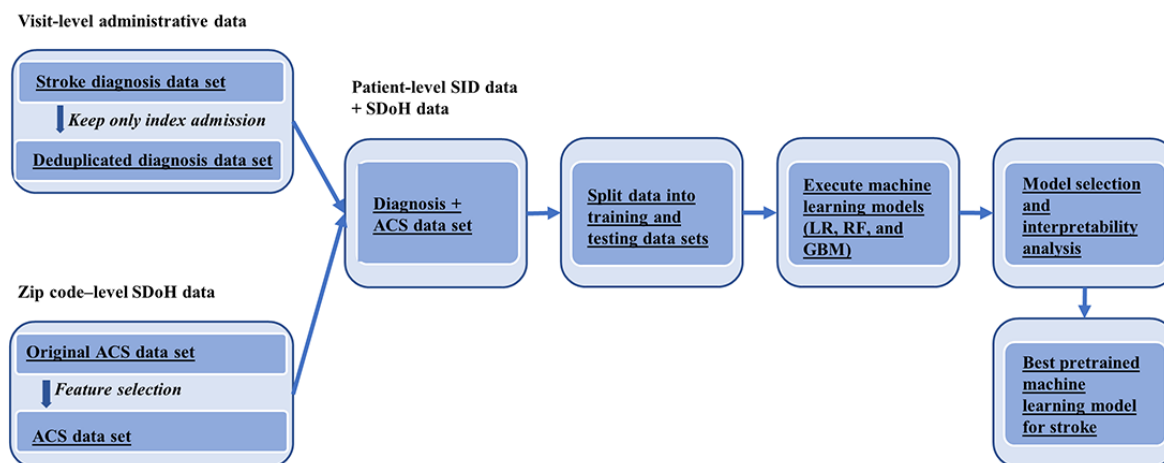
We first tuned the hyperparameters of all 3 models to find the optimal configurations using a grid search and 5-fold cross-validation on the entire data set. The evaluation metric used in the cross-validation was the area under the receiver operating characteristic curve (AUC). We used an 80-20 random split on the data set because this is a standard split method used in ML models and is typically performed to test the model performance in designing the ML-enabled diagnostic tool for providers in EDs [16]. We adopted the adaptive synthetic sampling technique to generate synthetic data for the minority class (eg, nonstroke “control”) in the training data to address the unbalanced distribution of stroke events in the real-world data. Using the optimal configurations of hyperparameters, we then developed and assessed our models using a balanced training data set with repeated 5-fold cross-validation and cost-sensitive classification to avoid overfitting. For each fold,

the models were evaluated on the performance metrics, including AUC, accuracy, precision, sensitivity or recall, specificity, and  $F_1$ -score, using the test data set. The logistic regression and RF models were implemented in Python (version 3.9.12, Python Software Foundation) using scikit-learn (version 1.0.2; David Cournapeau). GBM was implemented in Python 3.9.12 using CatBoost (version 1.0.6, Yandex LLC). The configurations of the key hyperparameters for each of our ML models are listed in [Multimedia Appendix 3](#).

As a robustness check, we adopted an alternative data split method by using historical 2012 data to predict for 2013 and using both 2012 and 2013 data to predict for 2014.

Although ML models can produce accurate predictions, they are often treated as black-box models that lack interpretability. This is an important problem, especially in medical care because clinicians are often unwilling to accept machine recommendations without clarity regarding the underlying reasoning [57]. However, according to a recent review, the number of ML studies in the medical domain that addressed explainability is very limited [58]. In this study, we followed the approach outlined by Saarela and Jauhiainen in their 2021 paper [59] to conduct a comparison of feature importance measures to enhance the interpretability or explainability of our models’ results. To provide insights into prediction and ultimately assist care providers in decision-making, we used TreeSHAP for tree-based ML models to explain the stroke prediction for each patient (refer to the details provided in the *Results* section). [Figure 2](#) demonstrates the investigation path followed to develop our models based on the synthesized data and compare and interpret the models to derive the best pretrained ML model for stroke prediction.

**Figure 2.** Analysis pipeline. ACS: American Community Survey; GBM: gradient boosting machine; LR: logistic regression; RF: random forest; SDoH: social determinants of health; SID: State Inpatient Database.



## Results

### Descriptive Statistics of the Data Set

In the final data set, there were 143,203 hospital visits of unique patients, and it was confirmed based on the hospital discharge records that 73% (n=104,662) of them had a stroke. The prediction models included 12 patient-level features from the hospital administrative data set, joined by 16 community-level

features from the ACS data set. We summarized the patient-level predictors into 3 categories: patient demographics, visit-level features, and individual-level SDoH; their summary statistics are presented in the following table ([Table 1](#)). Patients who ended up being diagnosed with stroke tended to be older, have more chronic conditions, and have Medicare as the primary payer. The results of 2-sample *t* tests (2-tailed) with Bonferroni correction showed that all the patient-level predictors were statistically different at a significance level of .05 between the

patients with stroke and those with stroke mimics (indicated by the *P* values in the last column of Table 1). The 16 community-level SDoH features were summarized into categories: area demographics, socioeconomic status,

occupation, and health insurance coverage at the population level. Multimedia Appendix 2 contains detailed information on the community-level predictors.

**Table 1.** Descriptive statistics of the patient-level predictors.

Features	Total sample (n=143,203), mean (SD)	Stroke cohort (n=104,662), mean (SD)	Stroke mimic cohort (n=38,541), mean (SD)	<i>P</i> value
<b>Patient demographics</b>				
Age (years)	65.2843 (19.97)	71.1259 (14.68)	49.4207 (23.49)	<.001
Sex (female)	0.5019 (0.50)	0.5014 (0.50)	0.5031 (0.50)	.03
Number of chronic conditions	6.5066 (3.21)	7.1200 (3.00)	4.8410 (3.17)	<.001
<b>Race and ethnicity</b>				
White	0.6594 (0.47)	0.6736 (0.47)	0.6209 (0.49)	<.001
Black	0.1802 (0.38)	0.1706 (0.38)	0.2064 (0.40)	<.001
Hispanic	0.1348 (0.34)	0.1302 (0.34)	0.1472 (0.35)	<.001
Other races	0.0256 (0.16)	0.0257 (0.16)	0.0255 (0.16)	.04
<b>Visit-level features</b>				
Emergency admission	0.9030 (0.30)	0.9094 (0.29)	0.8859 (0.32)	<.001
Elective admission	0.0403 (0.20)	0.0214 (0.14)	0.0914 (0.29)	<.001
Transfer in indicator	0.0913 (0.37)	0.0929 (0.37)	0.0869 (0.36)	<.001
Night shift <sup>a</sup>	0.3409 (0.47)	0.3257 (0.47)	0.3821 (0.49)	<.001
Weekend indicator	0.2558 (0.44)	0.2581 (0.44)	0.2496 (0.43)	<.001
<b>Individual-level SDoH<sup>b</sup></b>				
Urban residence	0.9529 (0.21)	0.9515 (0.21)	0.9567 (0.20)	<.001
<b>Primary payer</b>				
Medicare	0.6239 (0.48)	0.7027 (0.46)	0.4099 (0.49)	<.001
Medicaid	0.1103 (0.31)	0.0714 (0.26)	0.2159 (0.41)	<.001
Private insurance	0.1505 (0.36)	0.1331 (0.34)	0.1980 (0.40)	<.001
Other payers	0.1153 (0.32)	0.0929 (0.29)	0.1762 (0.38)	<.001
<b>Median household income</b>				
0-25th percentile	0.4025 (0.49)	0.3984 (0.49)	0.4134 (0.49)	<.001
26th-50th percentile	0.3261 (0.47)	0.3289 (0.47)	0.3186 (0.47)	<.001
51st-75th percentile	0.1992 (0.40)	0.1994 (0.40)	0.1986 (0.40)	.04
76th-100th percentile	0.0722 (0.26)	0.0733 (0.26)	0.0694 (0.25)	<.001

<sup>a</sup>Admission between 7 PM and 7 AM.

<sup>b</sup>SDoH: social determinants of health.

## Model Performance and Selection

Table 2 shows the algorithm performance measured on the test set for the 9 models run on 3 input combinations and 3 classifiers (logistic regression, RF, and GBM). ML classifiers consistently outperformed the benchmark logistic regression in all 3 input combinations. More specifically, the GBM classifier consistently outperformed logistic regression and RF in the first 2 input combinations (ie, when patient- and visit-level feature sets were used). When the patient-, visit-, and community-level variables were included as inputs (ie, the most complete input

combination), ML models dominated the logistic regression. Inclusion of the individual-level SDoH features improved the performance for all 3 classifiers, especially the GBM model, where AUC increased from 0.694 (model 3) to 0.823 (model 6). Further inclusion of the community-level SDoH features improved the overall predictive performance measures, AUC, sensitivity, and specificity of the 2 ML models (models 8 and 9).

We based our model selection on both the performance metrics and the clinical needs in actual care settings. Note that the cost

of misdiagnosis is asymmetrical. Misdiagnosis of a stroke (labeling a true stroke as a nonstroke condition) might have more severe adverse consequences for both patients and providers than overdiagnosis (ie, false-positive stroke diagnosis). Hence, the selected model should provide high sensitivity while maintaining specificity within a reasonable range. Both ML models (RF and GBM) correctly detected at least 97% (101,522/104,662) of all the patients that did have a stroke and thus significantly outperformed the prehospital stroke prediction scales (ranging between 0.38 and 0.62) [61] by a large margin. The Youden index, calculated by deducting 1 from the sum of the test's sensitivity and specificity, was used to evaluate the overall discriminative power of the diagnostic test. The Youden

index was not included in Table 2 because of space limitations; however, it can be easily calculated using sensitivity and specificity, both of which are included in the table. According to several recent literature reviews, the Youden index of the stroke prediction scales used in the emergency medical services, ambulances, and emergency room settings ranged from 0.30 to 0.54 [61,62], whereas that of our stroke prediction models ranged from 0.56 to 0.62.

Multimedia Appendix 4 presents the results of using the alternative data split method by using the historical 2012 data to train the model. Our model still demonstrated good overall performance with a high sensitivity rate of >90% and  $F_1$ -score in the range of 0.83 to 0.88.

**Table 2.** Performance of stroke prediction models.

Input combinations and model number	Classifier	Accuracy	AUC <sup>a</sup>	Sensitivity	Specificity	$F_1$ -score
<b>Patient demographics and visit information</b>						
1	Logit	0.828	0.693	0.960	0.626	0.893
2	RF <sup>b</sup>	0.804	0.680	0.928	0.632 <sup>c</sup>	0.877
3	GBM <sup>d</sup>	0.832	0.694	0.968	0.619	0.896
<b>Patient demographics, visit information, and individual SDoH<sup>e</sup></b>						
4	Logit	0.830	0.810	0.960	0.630	0.895
5	RF	0.794	0.724	0.899	0.656	0.868
6	GBM	0.835	0.823	0.965	0.631	0.898
<b>Patient demographics, visit information, individual SDoH, and community-level SDoH</b>						
7	Logit	0.822	0.810	0.967	0.629	0.891
8	RF	0.831	0.828	0.972	0.626	0.896
9	GBM	0.834	0.829	0.970	0.647	0.898

<sup>a</sup>AUC: area under the receiver operating characteristic curve.

<sup>b</sup>RF: random forest.

<sup>c</sup>For each input combination, the best performance among the 3 classifiers has been italicized.

<sup>d</sup>GBM: gradient boosting machine.

<sup>e</sup>SDoH: social determinants of health.

## Feature Importance Analysis

We found consistency across the 3 models in the most important features that explain their performance (Figure 3; glossary of the terms used as well as the variable definitions can be found in Multimedia Appendix 5). The top 5 features deal with age, the number of chronic conditions on admission, and primary payer (eg, Medicare or private insurance). The top 2 features were age and the number of chronic conditions across all 3 models. Both the RF and GBM models identified patients' age as one of the most important features. This is consistent with the findings of a recent study that used a new hybrid feature selection model that integrated various filter and wrapper methods to detect stroke risk [63]. Older age increases the predicted stroke probability, and younger age decreases the predicted probability. The second most important feature contributing to the models' performance was the number of chronic conditions on admission. A higher number of chronic

conditions on admission increases the predicted stroke probability.

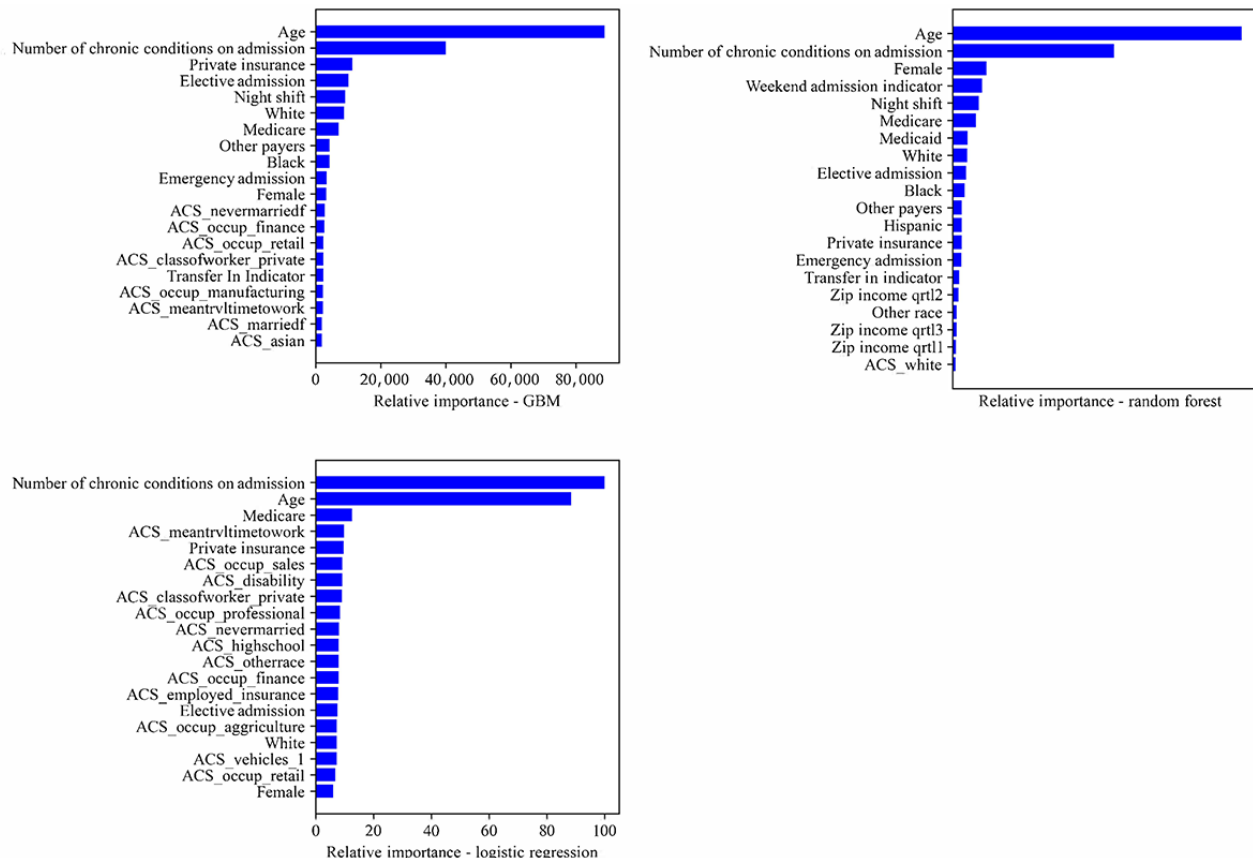
It is interesting to note that the patients' admission type (eg, whether it is an emergency or elective admission) and timing of admission (ie, whether they were admitted during the night shift) contributed to the accuracy of stroke prediction. Existing studies have investigated the presence of a "weekend effect" on mortality [64-66] and the differences in the quality of treatment that patients receive based on their time of hospital arrival or admission [67,68]. In general, these studies primarily focused on emergency admissions. If adverse patient outcomes such as mortality are related to different work practices and staff availability during off-hour periods, then the diagnosis of acute conditions is likely to be similarly affected, and our current findings confirm this hypothesis.

In addition to age, other patient-level demographic and socioeconomic factors, including gender, race, and primary

payer (ie, whether the medical expenses were covered by Medicare, Medicaid, private insurance, or other payers), contribute to the models' prediction. These findings complement the recently observed diverging stroke risk patterns among different racial and gender groups [69,70]. For instance, Howard et al [69] found that when aged between 45 and 74 years, White women were less likely to have a stroke than White men; however, there was no difference in stroke risk between White men and women when the latter were aged  $\geq 75$  years. By contrast, they found that Black women were at a lower risk of stroke than Black men when they were aged  $\leq 64$  years and

experienced a similar stroke risk thereafter [69]. Another study found that Black women had a greater risk of stroke than White women, and the racial disparities were greatest among women aged 50 to 60 years [70]. In addition, our findings revealed that health insurance status is not only associated with health care use but also an important predictor of stroke. These findings have important implications and suggest that a cookie-cutter approach may not work well for stroke prevention. For instance, interventions targeting socially disadvantaged individuals without Medicare coverage may provide the greatest benefit in reducing disparities.

**Figure 3.** Comparison of feature importance: 20 most important features for gradient boosting machine (GBM; upper left), random forest (upper right), and logistic regression (bottom). ACS: American Community Survey; Qrtl: quartile.



Some community-level SDoH variables (eg, percentage of single women; percentage of people with occupations closely related to finance, retail, and manufacturing industries; and mean travel time to work) were also among the top 20 features. However, the magnitude of their impact on stroke prediction was much less than that of patient-level demographic and socioeconomic features. This is consistent with the literature [40] and with the predictive performance reported in Table 2. Although both the individual- and community-level SDoH features helped improve the predictive performance, the inclusion of the individual-level SDoH features led to a much larger improvement (AUC increased from 0.694 to 0.823) than the inclusion of the community-level SDoH features (AUC increased from 0.823 to 0.829). Only adding the community-level SDoH features to the visit-level data (in the absence of individual-level SDoH features) increased the AUC from 0.694 to 0.724.

Ablation studies are commonly used for assigning importance scores to features [71-73]. In this method, the importance of a feature is decided based on the reduction in performance that its removal causes. We performed the ablation analysis as follows. First, we trained the GBM model on the training data set and calculated the base score on the testing data set using the accuracy metric. Second, we removed one feature from the training data set, trained the GBM model again, and then calculated the score of the model on the testing data set. This was repeated for each feature included in the model. Finally, we ranked the features based on the difference between the score calculated in their absence and the base score (calculated when all the features were present). Consistent with the findings from the feature importance comparison analysis, the top 2 highest-ranked features based on the ablation analysis were age and the number of chronic conditions, followed by the individual-level SDoH features and then the community-level SDoH features (Multimedia Appendix 6).



### Individual Prediction Explanations

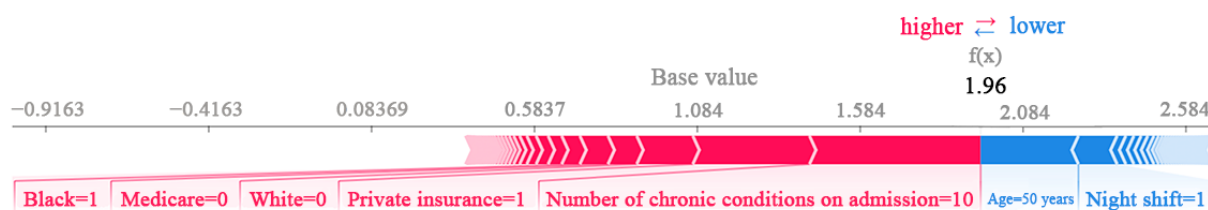
To provide insights into prediction and ultimately assist care providers in decision-making, we sought to explain the stroke prediction model using TreeSHAP [74], a variant of Shapley Additive Explanations (SHAP) for tree-based ML models. The SHAP method computes Shapley values from the coalitional game theory to quantify the contributions of each feature to the prediction [75-77]. TreeSHAP uses conditional expectation to estimate the effects for a single tree, and the Shapley values of a tree ensemble are the weighted average of the Shapley values of the individual trees.

Figure 4 shows SHAP values to explain the stroke prediction of 2 example cases (glossary of the terms used as well as the variable definitions can be found in Multimedia Appendix 5). We visualized feature attributions as “forces,” and each feature value is a force that either increases or decreases the prediction starting from the baseline. The base value or the expected value is the average of the model output over the training data and equals 1.084 [75]. Features that push the prediction higher (to

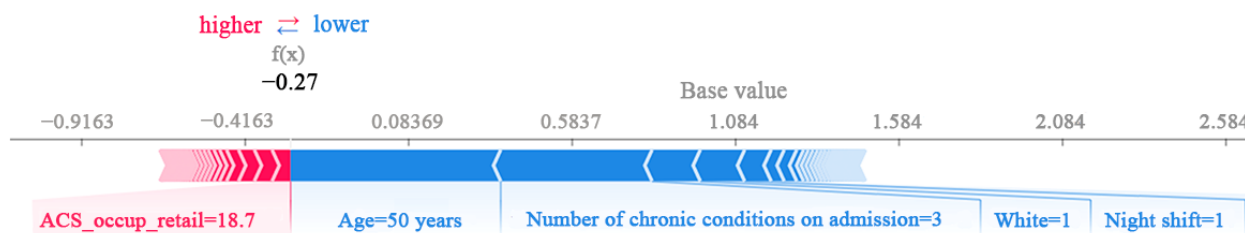
the right) are shown in red, and those pushing the prediction lower are in blue. The first example (prediction demonstration example 1) obtained an output value (ie, prediction for this observation) of 1.96, higher than the base value and hence, this example was labeled by the prediction model as stroke. Being Black, having 10 chronic conditions on admission, and having private insurance as the primary payer pushed the stroke prediction higher. This is consistent with the literature suggesting that the odds of a probable misdiagnosis of a stroke in the EDs were lower among Medicare or Medicaid recipients than among privately insured patients [7]. In comparison, we also looked at the SHAP values for another example (prediction demonstration example 2), where the model successfully predicts a stroke mimic. This second example obtained a low output value of -0.27. Similar to demonstration example 1, demonstration example 2 was aged 50 years; however, being White person, having 3 chronic conditions on admission, and being admitted during a night shift pushed the stroke prediction lower.

Figure 4. Shapley Additive Explanations values for example patients. ACS: American Community Survey.

#### Prediction demonstration example 1



#### Prediction demonstration example 2



These examples demonstrate that individual-level predictors of stroke can differ significantly from one case to another and can be used for personalized diagnostic and treatment decisions at the point of care, whereas the population-level analysis provides an overall ranking of the important predictors of stroke at hospital presentation and can be used to develop best practice guidelines and patient management programs.

## Discussion

### Principal Findings

In this study, we developed an ML-based approach using routinely collected administrative data to help reduce stroke misdiagnosis. Our findings suggest that before obtaining diagnostic imaging or laboratory test results, it is possible to predict stroke based on patients’ demographics and SDoH

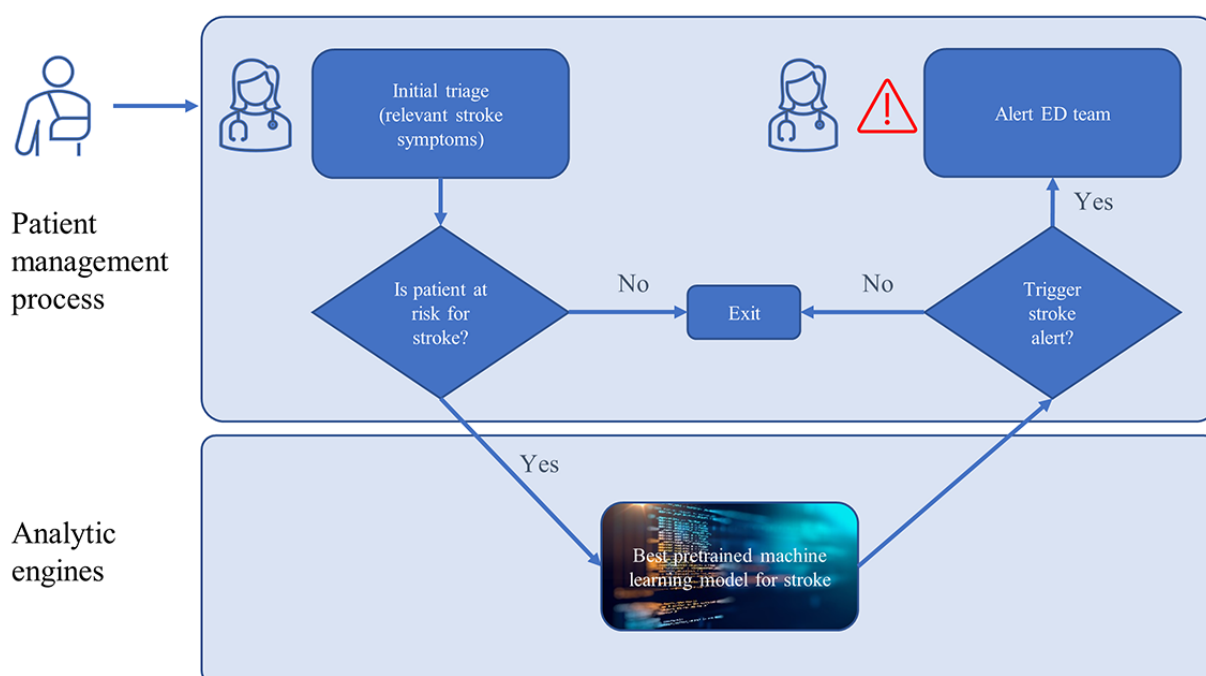
information available at the time of hospital presentation. The algorithm had an AUC of 83%, provided accurate results (high precision of 84%), and returned a supermajority (101,522/104,662, 97%) of all positive results (high sensitivity).

This study fills a critical gap in the current efforts to support stroke triage, which either focuses on improving specificity in the prehospital setting or requires detailed neurological assessments and imaging results. On the one hand, advanced ML techniques have been applied to assist in automatically interpreting clinical notes and imaging, but this is based on the availability of these information sources. On the other hand, because emergency medical service personnel lack the necessary time and training to perform detailed neurological assessments, short and simple clinical methods known as prehospital stroke scales have been developed to support the initial triage in the

field, such as the Cincinnati Prehospital Stroke Scale, Los Angeles Prehospital Stroke Scale, and Conveniently Grasped Field Assessment Stroke Triage. These scales have demonstrated wide performance variability in clinical practice; however, in general, they were found to have acceptable-to-good specificity but low sensitivity [62,78-80]. Literature reviews that compared studies with different prehospital stroke scales found that these scales varied in their accuracy and misdiagnosed up to 30% of acute strokes in the field. Depending on the sample and study site, Los Angeles Prehospital Stroke Scale and Cincinnati Prehospital Stroke Scale had similar diagnostic capabilities with sensitivity ranging from 0.79 to 0.91, and the sensitivity of using the Conveniently Grasped Field Assessment Stroke Triage in detecting large vessel occlusion stroke was 0.62. This means that these scales help detect false positives and thus reduce the wasteful use of medical resources. However, their low sensitivity has led to concerns that these scales will miss a substantial

percentage of people with stroke. Hence, it is important to have an additional screening or decision support tool to supplement clinical assessment and provide valuable information to increase the sensitivity in detecting stroke at hospital presentation and thus reduce missed diagnoses [81]. In crowded hospitals, and with shortage of medical resources and clinical staff, the ML-based model we proposed can help quickly prioritize patients for appropriate intervention. If a patient presents with stroke or stroke-like symptoms, an automated, computer-assisted screening tool will be triggered to quickly analyze all the patient's information available at the time of hospital presentation and suggest a diagnosis based on the best pretrained ML model for stroke. If the model predicts that the patient is at a high risk for stroke, a stroke pop-up will be triggered to alert the ED team. Figure 5 illustrates when and how this decision support prediction can be implemented in the field.

**Figure 5.** Decision support for stroke prediction. ED: emergency department.



This model can be integrated with other AI-enabled prediction or decision support systems based on EHRs in the ED to further improve stroke triage and diagnosis. Although EHR data contain rich and detailed clinical information, certain social and behavioral determinants that can also be important risk factors (eg, race) are both poorly represented (including a category for “Unknown”) and inadequately characterized in the EHR [82]. Furthermore, various obstacles such as the lack of interoperability have limited the full use of EHR data to improve the delivery of care. Consequently, existing studies are mostly based on patient data confined to a single EHR system within a single geographic area [83]. By contrast, administrative data, such as claims data, follow specific standards for both the structure and meaning of the variables contained within a claim, and nearly every health care provider must submit electronic claims in the same format to their payers or clearinghouses. Hence, such administrative data provide an efficient way to complement EHR data in measuring many important aspects

of health care delivery and provide solutions. We obtained the best of both worlds by leveraging the widely available administrative data with SDoH information to screen and quickly prioritize patients at hospital presentation and then using EHR data with rich clinical documentation and diagnostic test results to further assess and stratify patients based on risk.

### Comparison With Prior Work

It is important to consider specific clinical needs and care settings when comparing the various forms of performance measures reported across studies. In the case of strokes, misdiagnosis of a stroke (labeling a true stroke as a nonstroke condition) usually leads to more severe adverse patient outcomes than overdiagnosis. Although false-positive stroke mimics rarely lead to legal consequences, false negatives can cause delays in critical treatments and often give rise to accusations of medical errors. Moreover, given the inherent trade-off between sensitivity and specificity, the prehospital stroke scales' focus

on specificity (ie, reducing overdiagnosis) may result in a substantial number of misdiagnoses of strokes that need to be addressed at patients' hospital presentations. Therefore, minimizing the false-negative rate or maximizing the sensitivity is paramount in acute care settings for both patients and providers. Several recent studies have compared the currently available clinical assessment tools such as the field stroke triage scale, National Institute of Health Stroke Scale, Los Angeles Motor Scale, and Rapid Arterial Occlusion Evaluation, which incorporate cortical signs (eg, gaze deviation, aphasia, and neglect) as well as motor dysfunction, and found that these tools had better diagnostic accuracy for detecting patients with large vessel occlusion than for distinguishing between acute stroke and stroke mimics [81]. Many studies were designed such that patients with hemorrhage or stroke-mimicking conditions were excluded [81]. Clinical assessment tools aimed at distinguishing between acute stroke and stroke mimics demonstrated modest diagnostic accuracy with low sensitivity, ranging from 38% to 62%, in the prehospital setting [61]. The stroke diagnostic tools designed for ED settings, such as the Recognition of Stroke in the Emergency Room scale and the FABS scoring system, are found to have a higher sensitivity than the prehospital scales (up to 93%), and they require clinical assessments from neurologists, brain CT findings, or additional clinical information, such as atrial fibrillation [81,84]. To our knowledge, the sensitivity of the algorithms in this study, without relying on the availability of additional clinical information or imaging findings, outperforms any scoring scales used in the prehospital or ED settings.

This study is also one of the first large-scale studies to systematically assess the added value of SDoH information in a population-based risk-prediction setting using administrative data. Although many studies have shown that various social or behavioral factors are associated with health outcomes, very few have explicitly examined whether the knowledge of these factors improves the prediction of clinical events or health outcomes. Our results are consistent with the findings of nascent studies that link SDoH data with EHR data to predict ED visits [85] or the need for various social service referrals [86]. However, because EHR systems have not achieved full interoperability yet, these studies are mostly confined to patient data from a single EHR system within a single geographic area. This study extends the literature by leveraging the routinely collected data that span different health care systems and regions to complement some of the necessary first steps associated with population health analytics. Moreover, the development of electronic health information exchanges helps bring together information from multiple sources and combine administrative claims data with clinical data. Such progress makes it possible to create an integrated profile of a patient at the time of hospital presentation and further empowers our predictive analytics.

### Limitations and Future Research

This study has room for further improvement, which is left for future research. First, this was a retrospective study, and confirmation of stroke cases relied on International Classification of Diseases codes. It is desirable to have patients' complicated medical records reviewed to ascertain stroke diagnosis; however, this process is labor intensive and

expensive, especially when it is a large-scale study with hundreds of thousands of patients across different health systems. Our results require further validation but have the potential for improving stroke triage and diagnosis.

Second, the algorithm we proposed should not be considered as the gold standard for stroke diagnosis. Rather, we believe that the algorithm complements the existing stroke scoring systems used in the prehospital or emergency room settings and can be integrated into ML-enabled decision support systems that combine patients' medical history, SDoH, and clinical data. Such a decision system would have the advantage of being agile and iterative, in the sense that the model outcome can be reassessed at regular intervals as more data are collected in the ED, as well as the integration of variables with the most promising relevance.

Third, the focus of this study was to predict stroke solely based on the information available at the time of a patient's initial presentation at the hospital. This is because first-time or new patients with stroke make up the supermajority (77%) of the yearly US patient population with stroke [87], and it is more challenging to make stroke predictions accurately for those patients who show up at the ED for the first time with no historical data. Patients with repeated readmissions and single-visit patients may follow different trajectories with different underlying social and behavioral determinants [88]. Future research may continue to explore how to best incorporate past information to improve prediction and identify key risk factors for repeated patients.

Finally, our findings are limited to the SDoH variables available in administrative data, suggesting the importance of developing standards and tools to routinely collect and screen individual-level SDoH data and effectively integrate them into both EHR and structured claims data. Our current prediction does not require any additional effort to collect additional individual-level SDoH. The community-level ACS variables have already been incorporated as part of the best pretrained model. The patient-level details used in our prediction are (1) basic demographics including age, gender, race and ethnicity, and primary payer (ie, Medicare, Medicaid, private insurance, or others); (2) arrival information (eg, whether it was an emergency or elective admission and whether the patient was admitted during a weekend or night shift); and (3) whether the patient resided in an urban or a rural area and the quartile in which their median household income fell (Table 1), both of which are based on the zip code variable. All 3 categories of information are routinely collected by triage nurses at hospital EDs during the initial triage. For instance, the zip code can be obtained from the patient's home address, and the primary payer can be identified from the insurance information. Hence, patients do not need to answer any additional SDoH-related questions for the currently proposed prediction. To include more patient-level SDoH and further improve the predictive performance, more efforts are needed to develop SDoH screening and collection tools. National efforts are underway starting with primary care, such as the Protocol for Responding to and Assessing Patients' Assets, Risks, and Experiences survey, a standardized patient risk assessment tool consisting of a set of national core measures for addressing patients' SDoH. Future

research can also leverage more advanced ML algorithms (eg, deep learning) to facilitate a more comprehensive and efficient analysis of the large, high-dimensional data sets with claims, EHR, and SDoH data.

### Conclusions

Stroke is among the most common and dangerous misdiagnosed medical conditions. Black people, Hispanic people, women, older people on Medicare, and people in rural areas are less likely to be diagnosed in time for treatment after having a stroke. Timely detection is the key to effective management and improved patient outcomes.

We developed a high-performance ML-based stroke prediction algorithm that outperforms the existing early warning scoring

systems. The algorithm is based on variables routinely collected and readily available at the time of patients' hospital presentations and may provide an opportunity for enhanced patient monitoring and stroke triage and improved health outcomes. Because the prediction model does not require clinical notes or diagnostic test results, it can be particularly useful in underresourced EDs in rural and underserved communities with limited availability of sensitive diagnostic tools and incomplete data-gathering capabilities. Moreover, the algorithm can be incorporated into an automated, AI-enabled decision support system that combines administrative data widely available at the time of ED presentation and subsequently available clinical notes and diagnostic test results to further improve stroke diagnosis, triage, and management.

---

### Acknowledgments

The authors thank the Agency for Healthcare Research and Quality and its partner organization the Florida Agency for Health Care Administration for providing access to the State Inpatient Databases through the Healthcare Cost and Utilization Project.

They are grateful for the comments and suggestions from the 3 anonymous reviewers and participants of the 2019 Institute for Operations Research and the Management Sciences Healthcare Conference, 2020 Workshop on Information Technologies and Systems, Healthcare Information and Management Systems Society 2020 Big Data Symposium, and Production and Operations Management Society 31st Annual Conference, where earlier versions of this work were presented, and to graduate students at the Heinz College at Carnegie Mellon University for their help with data gathering and preliminary analyses.

---

### Authors' Contributions

MC and RP conceived the idea for this study. MC, RP, and XT designed the study. MC and XT conducted the analysis and drafted the manuscript with extensive input and critical suggestions from RP. All the authors interpreted the results, revised the manuscript, and read and approved the final manuscript.

---

### Conflicts of Interest

None declared.

---

### Multimedia Appendix 1

The top 20 principal diagnoses in the analysis sample.

[\[DOCX File , 16 KB-Multimedia Appendix 1\]](#)

---

### Multimedia Appendix 2

Community-level social determinants of health variables.

[\[DOCX File , 17 KB-Multimedia Appendix 2\]](#)

---

### Multimedia Appendix 3

Tuned hyperparameters in machine learning models.

[\[DOCX File , 17 KB-Multimedia Appendix 3\]](#)

---

### Multimedia Appendix 4

Performance of the stroke prediction models based on the alternative data split method.

[\[DOCX File , 16 KB-Multimedia Appendix 4\]](#)

---

### Multimedia Appendix 5

Glossary of the terms used as well as the variable definitions.

[\[DOCX File , 15 KB-Multimedia Appendix 5\]](#)

---

## Multimedia Appendix 6

Results of the ablation analysis.

[\[DOCX File, 17 KB-Multimedia Appendix 6\]](#)

## References

1. Ball JR, Balogh E. Improving diagnosis in health care: highlights of a report from the National Academies of Sciences, Engineering, and Medicine. *Ann Intern Med* 2016 Jan 05;164(1):59-61. [doi: [10.7326/M15-2256](https://doi.org/10.7326/M15-2256)] [Medline: [26414299](https://pubmed.ncbi.nlm.nih.gov/26414299/)]
2. Saber Tehrani AS, Lee H, Mathews SC, Shore A, Makary MA, Pronovost PJ, et al. 25-Year summary of US malpractice claims for diagnostic errors 1986-2010: an analysis from the National Practitioner Data Bank. *BMJ Qual Saf* 2013 Aug;22(8):672-680. [doi: [10.1136/bmjqs-2012-001550](https://doi.org/10.1136/bmjqs-2012-001550)] [Medline: [23610443](https://pubmed.ncbi.nlm.nih.gov/23610443/)]
3. Venkat A, Cappelen-Smith C, Askar S, Thomas PR, Bhaskar S, Tam A, et al. Factors associated with stroke misdiagnosis in the emergency department: a retrospective case-control study. *Neuroepidemiology* 2018;51(3-4):123-127. [doi: [10.1159/000491635](https://doi.org/10.1159/000491635)] [Medline: [30092562](https://pubmed.ncbi.nlm.nih.gov/30092562/)]
4. Schiff GD, Hasan O, Kim S, Abrams R, Cosby K, Lambert BL, et al. Diagnostic error in medicine: analysis of 583 physician-reported errors. *Arch Intern Med* 2009 Nov 09;169(20):1881-1887. [doi: [10.1001/archinternmed.2009.333](https://doi.org/10.1001/archinternmed.2009.333)] [Medline: [19901140](https://pubmed.ncbi.nlm.nih.gov/19901140/)]
5. Tarnutzer AA, Lee SH, Robinson KA, Wang Z, Edlow JA, Newman-Toker DE. ED misdiagnosis of cerebrovascular events in the era of modern neuroimaging: a meta-analysis. *Neurology* 2017 Apr 11;88(15):1468-1477 [FREE Full text] [doi: [10.1212/WNL.0000000000003814](https://doi.org/10.1212/WNL.0000000000003814)] [Medline: [28356464](https://pubmed.ncbi.nlm.nih.gov/28356464/)]
6. Dubois RW, Brook RH. Preventable deaths: who, how often, and why? *Ann Intern Med* 1988 Oct 01;109(7):582-589. [doi: [10.7326/0003-4819-109-7-582](https://doi.org/10.7326/0003-4819-109-7-582)] [Medline: [3421565](https://pubmed.ncbi.nlm.nih.gov/3421565/)]
7. Newman-Toker DE, Moy E, Valente E, Coffey R, Hines AL. Missed diagnosis of stroke in the emergency department: a cross-sectional analysis of a large population-based sample. *Diagnosis (Berl)* 2014 Jun;1(2):155-166 [FREE Full text] [doi: [10.1515/dx-2013-0038](https://doi.org/10.1515/dx-2013-0038)] [Medline: [28344918](https://pubmed.ncbi.nlm.nih.gov/28344918/)]
8. Merino JG, Luby M, Benson RT, Davis LA, Hsia AW, Latour LL, et al. Predictors of acute stroke mimics in 8187 patients referred to a stroke service. *J Stroke Cerebrovasc Dis* 2013 Nov;22(8):e397-e403 [FREE Full text] [doi: [10.1016/j.jstrokecerebrovasdis.2013.04.018](https://doi.org/10.1016/j.jstrokecerebrovasdis.2013.04.018)] [Medline: [23680681](https://pubmed.ncbi.nlm.nih.gov/23680681/)]
9. Gioia LC, Zewude RT, Kate MP, Liss K, Rowe BH, Buck B, et al. Prehospital systolic blood pressure is higher in acute stroke compared with stroke mimics. *Neurology* 2016 Jun 07;86(23):2146-2153 [FREE Full text] [doi: [10.1212/WNL.0000000000002747](https://doi.org/10.1212/WNL.0000000000002747)] [Medline: [27194383](https://pubmed.ncbi.nlm.nih.gov/27194383/)]
10. Messé SR, Khatri P, Reeves MJ, Smith EE, Saver JL, Bhatt DL, et al. Why are acute ischemic stroke patients not receiving IV tPA? Results from a national registry. *Neurology* 2016 Oct 11;87(15):1565-1574 [FREE Full text] [doi: [10.1212/WNL.0000000000003198](https://doi.org/10.1212/WNL.0000000000003198)] [Medline: [27629092](https://pubmed.ncbi.nlm.nih.gov/27629092/)]
11. Bayona H, Ropero B, Salazar AJ, Pérez JC, Granja MF, Martínez CF, et al. Comprehensive telestroke network to optimize health care delivery for cerebrovascular diseases: algorithm development. *J Med Internet Res* 2020 Jul 27;22(7):e18058 [FREE Full text] [doi: [10.2196/18058](https://doi.org/10.2196/18058)] [Medline: [32716302](https://pubmed.ncbi.nlm.nih.gov/32716302/)]
12. Berg LM, Källberg AS, Göransson KE, Östergren J, Florin J, Ehrenberg A. Interruptions in emergency department work: an observational and interview study. *BMJ Qual Saf* 2013 Aug;22(8):656-663. [doi: [10.1136/bmjqs-2013-001967](https://doi.org/10.1136/bmjqs-2013-001967)] [Medline: [23584208](https://pubmed.ncbi.nlm.nih.gov/23584208/)]
13. Hamden K, Jeanmonod D, Gualtieri D, Jeanmonod R. Comparison of resident and mid-level provider productivity in a high-acuity emergency department setting. *Emerg Med J* 2014 Mar;31(3):216-219. [doi: [10.1136/emered-2012-201904](https://doi.org/10.1136/emered-2012-201904)] [Medline: [23360904](https://pubmed.ncbi.nlm.nih.gov/23360904/)]
14. Schnapp BH, Sun JE, Kim JL, Strayer RJ, Shah KH. Cognitive error in an academic emergency department. *Diagnosis (Berl)* 2018 Sep 25;5(3):135-142 [FREE Full text] [doi: [10.1515/dx-2018-0011](https://doi.org/10.1515/dx-2018-0011)] [Medline: [30016277](https://pubmed.ncbi.nlm.nih.gov/30016277/)]
15. Stiell A, Forster AJ, Stiell IG, van Walraven C. Prevalence of information gaps in the emergency department and the effect on patient outcomes. *CMAJ* 2003 Nov 11;169(10):1023-1028 [FREE Full text] [Medline: [14609971](https://pubmed.ncbi.nlm.nih.gov/14609971/)]
16. Abedi V, Khan A, Chaudhary D, Misra D, Avula V, Mathrawala D, et al. Using artificial intelligence for improving stroke diagnosis in emergency departments: a practical framework. *Ther Adv Neurol Disord* 2020 Aug 25;13:1756286420938962 [FREE Full text] [doi: [10.1177/1756286420938962](https://doi.org/10.1177/1756286420938962)] [Medline: [32922515](https://pubmed.ncbi.nlm.nih.gov/32922515/)]
17. Letham B, Rudin C, McCormick TH, Madigan D. Interpretable classifiers using rules and Bayesian analysis: building a better stroke prediction model. *Ann Appl Stat* 2015 Sep 1;9(3):1350-1371. [doi: [10.1214/15-aos848](https://doi.org/10.1214/15-aos848)]
18. Lee JW, Lim HS, Kim DW, Shin SA, Kim J, Yoo B, et al. The development and implementation of stroke risk prediction model in National Health Insurance Service's personal health record. *Comput Methods Programs Biomed* 2018 Jan;153:253-257 [FREE Full text] [doi: [10.1016/j.cmpb.2017.10.007](https://doi.org/10.1016/j.cmpb.2017.10.007)] [Medline: [29157457](https://pubmed.ncbi.nlm.nih.gov/29157457/)]
19. Teoh D. Towards stroke prediction using electronic health records. *BMC Med Inform Decis Mak* 2018 Dec 04;18(1):127 [FREE Full text] [doi: [10.1186/s12911-018-0702-y](https://doi.org/10.1186/s12911-018-0702-y)] [Medline: [30509279](https://pubmed.ncbi.nlm.nih.gov/30509279/)]

20. Sung SF, Hsieh CY, Kao Yang YH, Lin HJ, Chen CH, Chen YW, et al. Developing a stroke severity index based on administrative data was feasible using data mining techniques. *J Clin Epidemiol* 2015 Nov;68(11):1292-1300. [doi: [10.1016/j.jclinepi.2015.01.009](https://doi.org/10.1016/j.jclinepi.2015.01.009)] [Medline: [25700940](https://pubmed.ncbi.nlm.nih.gov/25700940/)]
21. Sung SF, Chen SC, Hsieh CY, Li CY, Lai EC, Hu YH. A comparison of stroke severity proxy measures for claims data research: a population-based cohort study. *Pharmacoepidemiol Drug Saf* 2016 Apr;25(4):438-443. [doi: [10.1002/pds.3944](https://doi.org/10.1002/pds.3944)] [Medline: [26696591](https://pubmed.ncbi.nlm.nih.gov/26696591/)]
22. Asadi H, Dowling R, Yan B, Mitchell P. Machine learning for outcome prediction of acute ischemic stroke post intra-arterial therapy. *PLoS One* 2014 Feb 10;9(2):e88225 [FREE Full text] [doi: [10.1371/journal.pone.0088225](https://doi.org/10.1371/journal.pone.0088225)] [Medline: [24520356](https://pubmed.ncbi.nlm.nih.gov/24520356/)]
23. Heo J, Yoon JG, Park H, Kim YD, Nam HS, Heo JH. Machine learning-based model for prediction of outcomes in acute stroke. *Stroke* 2019 May;50(5):1263-1265. [doi: [10.1161/STROKEAHA.118.024293](https://doi.org/10.1161/STROKEAHA.118.024293)] [Medline: [30890116](https://pubmed.ncbi.nlm.nih.gov/30890116/)]
24. Garg R, Oh E, Naidech A, Kording K, Prabhakaran S. Automating ischemic stroke subtype classification using machine learning and natural language processing. *J Stroke Cerebrovasc Dis* 2019 Jul;28(7):2045-2051. [doi: [10.1016/j.jstrokecerebrovasdis.2019.02.004](https://doi.org/10.1016/j.jstrokecerebrovasdis.2019.02.004)] [Medline: [31103549](https://pubmed.ncbi.nlm.nih.gov/31103549/)]
25. Ni Y, Alwell K, Moomaw CJ, Woo D, Adeoye O, Flaherty ML, et al. Towards phenotyping stroke: leveraging data from a large-scale epidemiological study to detect stroke diagnosis. *PLoS One* 2018 Feb 14;13(2):e0192586 [FREE Full text] [doi: [10.1371/journal.pone.0192586](https://doi.org/10.1371/journal.pone.0192586)] [Medline: [29444182](https://pubmed.ncbi.nlm.nih.gov/29444182/)]
26. Zhao Y, Fu S, Bielinski SJ, Decker PA, Chamberlain AM, Roger VL, et al. Natural language processing and machine learning for identifying incident stroke from electronic health records: algorithm development and validation. *J Med Internet Res* 2021 Mar 08;23(3):e22951 [FREE Full text] [doi: [10.2196/22951](https://doi.org/10.2196/22951)] [Medline: [33683212](https://pubmed.ncbi.nlm.nih.gov/33683212/)]
27. Mouridsen K, Thurner P, Zaharchuk G. Artificial intelligence applications in stroke. *Stroke* 2020 Aug;51(8):2573-2579. [doi: [10.1161/STROKEAHA.119.027479](https://doi.org/10.1161/STROKEAHA.119.027479)] [Medline: [32693750](https://pubmed.ncbi.nlm.nih.gov/32693750/)]
28. Shen J, Zhang CJ, Jiang B, Chen J, Song J, Liu Z, et al. Artificial intelligence versus clinicians in disease diagnosis: systematic review. *JMIR Med Inform* 2019 Aug 16;7(3):e10010 [FREE Full text] [doi: [10.2196/10010](https://doi.org/10.2196/10010)] [Medline: [31420959](https://pubmed.ncbi.nlm.nih.gov/31420959/)]
29. Kabra R, Robbie H, Connor SE. Diagnostic yield and impact of MRI for acute ischaemic stroke in patients presenting with dizziness and vertigo. *Clin Radiol* 2015 Jul;70(7):736-742. [doi: [10.1016/j.crad.2015.01.016](https://doi.org/10.1016/j.crad.2015.01.016)] [Medline: [25956665](https://pubmed.ncbi.nlm.nih.gov/25956665/)]
30. Chaturvedi S, Ofner S, Baye F, Myers LJ, Phipps M, Sico JJ, et al. Have clinicians adopted the use of brain MRI for patients with TIA and minor stroke? *Neurology* 2017 Jan 17;88(3):237-244 [FREE Full text] [doi: [10.1212/WNL.0000000000003503](https://doi.org/10.1212/WNL.0000000000003503)] [Medline: [27927939](https://pubmed.ncbi.nlm.nih.gov/27927939/)]
31. Chan KS, Roberts E, McCleary R, Buttorff C, Gaskin DJ. Community characteristics and mortality: the relative strength of association of different community characteristics. *Am J Public Health* 2014 Sep;104(9):1751-1758. [doi: [10.2105/AJPH.2014.301944](https://doi.org/10.2105/AJPH.2014.301944)] [Medline: [25033152](https://pubmed.ncbi.nlm.nih.gov/25033152/)]
32. Hill PL, Weston SJ, Jackson JJ. Connecting social environment variables to the onset of major specific health outcomes. *Psychol Health* 2014;29(7):753-767. [doi: [10.1080/08870446.2014.884221](https://doi.org/10.1080/08870446.2014.884221)] [Medline: [24438045](https://pubmed.ncbi.nlm.nih.gov/24438045/)]
33. Closing the gap in a generation: health equity through action on the social determinants of health. CSDH final report. World Health Organization, Commission on Social Determinants of Health. Geneva, Switzerland: World Health Organization; 2008. URL: <https://apps.who.int/iris/bitstream/handle/10665/43943/97892?sequence=1> [accessed 2022-06-26]
34. Social determinants of health. Promedia. URL: <https://www.promedica.org/social-determinants-of-health/> [accessed 2022-06-26]
35. County health rankings and roadmaps. The University of Wisconsin Population Health Institute. 2022. URL: <https://www.countyhealthrankings.org/> [accessed 2022-06-26]
36. Judd SE, Gutiérrez OM, Newby PK, Howard G, Howard VJ, Locher JL, et al. Dietary patterns are associated with incident stroke and contribute to excess risk of stroke in Black Americans. *Stroke* 2013 Dec;44(12):3305-3311 [FREE Full text] [doi: [10.1161/STROKEAHA.113.002636](https://doi.org/10.1161/STROKEAHA.113.002636)] [Medline: [24159061](https://pubmed.ncbi.nlm.nih.gov/24159061/)]
37. Avendano M, Glymour MM. Stroke disparities in older Americans: is wealth a more powerful indicator of risk than income and education? *Stroke* 2008 May;39(5):1533-1540 [FREE Full text] [doi: [10.1161/STROKEAHA.107.490383](https://doi.org/10.1161/STROKEAHA.107.490383)] [Medline: [18436891](https://pubmed.ncbi.nlm.nih.gov/18436891/)]
38. Glymour MM, Avendaño M, Haas S, Berkman LF. Lifecourse social conditions and racial disparities in incidence of first stroke. *Ann Epidemiol* 2008 Dec;18(12):904-912 [FREE Full text] [doi: [10.1016/j.annepidem.2008.09.010](https://doi.org/10.1016/j.annepidem.2008.09.010)] [Medline: [19041589](https://pubmed.ncbi.nlm.nih.gov/19041589/)]
39. Salaycik KJ, Kelly-Hayes M, Beiser A, Nguyen AH, Brady SM, Kase CS, et al. Depressive symptoms and risk of stroke: the Framingham study. *Stroke* 2007 Jan;38(1):16-21. [doi: [10.1161/01.STR.0000251695.39877.ca](https://doi.org/10.1161/01.STR.0000251695.39877.ca)] [Medline: [17138952](https://pubmed.ncbi.nlm.nih.gov/17138952/)]
40. Chen M, Tan X, Padman R. Social determinants of health in electronic health records and their impact on analysis and risk prediction: a systematic review. *J Am Med Inform Assoc* 2020 Nov 01;27(11):1764-1773 [FREE Full text] [doi: [10.1093/jamia/ocaa143](https://doi.org/10.1093/jamia/ocaa143)] [Medline: [33202021](https://pubmed.ncbi.nlm.nih.gov/33202021/)]
41. Freij M, Dullabh P, Hovey L, Leonard J, Card A, Dhopeswarkar R. Incorporating social determinants of health in Electronic Health Records: a qualitative study of perspectives on current practices among top vendors. NORC at the University of Chicago. 2018 Nov 13. URL: <https://aspe.hhs.gov/sites/default/files/private/pdf/259901/NORCSDH.pdf> [accessed 2022-06-26]
42. HIPAA. Healthcare Cost and Utilization Project User Support. URL: [https://www.hcup-us.ahrq.gov/DUA/dua\\_508/DUA508version.jsp#hipaa](https://www.hcup-us.ahrq.gov/DUA/dua_508/DUA508version.jsp#hipaa) [accessed 2023-01-23]

43. American Community Survey (ACS). United States Census Bureau. URL: <https://www.census.gov/programs-surveys/acs> [accessed 2022-01-27]
44. Clinical Classifications Software (CCS) for ICD-9-CM fact sheet. Healthcare Cost and Utilization Project (HCUP). Rockville, MD, USA: Agency for Healthcare Research and Quality; 2012. URL: <https://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccsfactsheet.jsp> [accessed 2022-01-27]
45. Epocrates, an athenahealth service. Epocrates. 2019. URL: <https://www.epocrates.com/> [accessed 2022-06-26]
46. Lewis MW, Khodneva Y, Safford MM. Access to care, income, education and coronary heart disease outcomes in the reasons for geographic and racial differences in stroke study. *Circulation* 2016 Mar;133(suppl\_1):A43. [doi: [10.1161/circ.133.suppl\\_1.43](https://doi.org/10.1161/circ.133.suppl_1.43)]
47. Min YI, Anugu P, Butler KR, Hartley TA, Mwasongwe S, Norwood AF, et al. Cardiovascular disease burden and socioeconomic correlates: findings from the Jackson Heart Study. *J Am Heart Assoc* 2017 Aug 04;6(8):e004416 [FREE Full text] [doi: [10.1161/JAHA.116.004416](https://doi.org/10.1161/JAHA.116.004416)] [Medline: [28778943](https://pubmed.ncbi.nlm.nih.gov/28778943/)]
48. Sabbah W, Tsakos G, Chandola T, Sheiham A, Watt RG. Social gradients in oral and general health. *J Dent Res* 2007 Oct;86(10):992-996. [doi: [10.1177/154405910708601014](https://doi.org/10.1177/154405910708601014)] [Medline: [17890677](https://pubmed.ncbi.nlm.nih.gov/17890677/)]
49. Fretz A, Schneider AL, McEvoy JW, Hoogeveen R, Ballantyne CM, Coresh J, et al. The association of socioeconomic status with subclinical myocardial damage, incident cardiovascular events, and mortality in the ARIC study. *Am J Epidemiol* 2016 Mar 01;183(5):452-461 [FREE Full text] [doi: [10.1093/aje/kwv253](https://doi.org/10.1093/aje/kwv253)] [Medline: [26861239](https://pubmed.ncbi.nlm.nih.gov/26861239/)]
50. González MA, Rodríguez Artalejo F, Calero JR. Relationship between socioeconomic status and ischaemic heart disease in cohort and case-control studies: 1960-1993. *Int J Epidemiol* 1998 Jun;27(3):350-358. [doi: [10.1093/ije/27.3.350](https://doi.org/10.1093/ije/27.3.350)] [Medline: [9698119](https://pubmed.ncbi.nlm.nih.gov/9698119/)]
51. Patel N, Arora S, Lahewala S, Shah H, Patel P, Bambhroliya C, et al. Congestive heart failure: predictors of readmission and short-term mortality - a national perspective. *Circulation* 2016;134(suppl\_1):A15225. [doi: [10.1161/circ.134.suppl\\_1.15225](https://doi.org/10.1161/circ.134.suppl_1.15225)]
52. Agyemang P, Ayers C, Barrington D, Suresh V, McClurkin M, Powell-Wiley TM. Food insecurity and inadequate health insurance coverage are barriers to ideal cardiovascular health for U.S. Adults: data from the National Health and Nutrition Examination Survey (NHANES). *Circulation* 2013;128(suppl\_22):A14135 [FREE Full text] [doi: [10.1161/circ.128.suppl\\_22.A14135](https://doi.org/10.1161/circ.128.suppl_22.A14135)]
53. Bai X, Padman R, Ramsey J, Spirtes P. Tabu search-enhanced graphical models for classification in high dimensions. *INFORMS J Comput* 2008 Aug;20(3):423-437. [doi: [10.1287/ijoc.1070.0255](https://doi.org/10.1287/ijoc.1070.0255)]
54. He H, Bai Y, Garcia EA, Li S. ADASYN: adaptive synthetic sampling approach for imbalanced learning. In: Proceedings of the 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence). 2008 Presented at: IJCNN '08; June 1-8, 2008; Hong Kong, China p. 1322-1328. [doi: [10.1109/ijcnn.2008.4633969](https://doi.org/10.1109/ijcnn.2008.4633969)]
55. Sáez JA, Luengo J, Stefanowski J, Herrera F. Managing borderline and noisy examples in imbalanced classification by combining SMOTE with ensemble filtering. In: Proceedings of the 15th International Conference on Intelligent Data Engineering and Automated Learning. 2014 Presented at: IDEAL '14; September 10-12, 2014; Salamanca, Spain p. 61-68. [doi: [10.1007/978-3-319-10840-7\\_8](https://doi.org/10.1007/978-3-319-10840-7_8)]
56. Wang BX, Japkowicz N. Imbalanced data set learning with synthetic samples. In: Proceedings of the 2004 IRIS Machine Learning Workshop. 2004 Presented at: IRIS '14; June 9, 2004; Ottawa, Canada.
57. Hosmer Jr DW, Lemeshow S, Sturdivant RX. Applied Logistic Regression. 3rd edition. Hoboken, NJ, USA: John Wiley & Sons; 2013.
58. Ho TK. Random decision forests. In: Proceedings of 3rd International Conference on Document Analysis and Recognition. 1995 Presented at: ICDAR '95; August 14-16, 1995; Montreal, Canada p. 278-282. [doi: [10.1109/icdar.1995.598994](https://doi.org/10.1109/icdar.1995.598994)]
59. Ho TK. The random subspace method for constructing decision forests. *IEEE Trans Pattern Anal Machine Intell* 1998 Aug;20(8):832-844 [FREE Full text] [doi: [10.1109/34.709601](https://doi.org/10.1109/34.709601)]
60. Friedman JH. Greedy function approximation: a gradient boosting machine. *Ann Statist* 2001 Oct 1;29(5):1189-1232. [doi: [10.1214/aos/1013203451](https://doi.org/10.1214/aos/1013203451)]
61. Nguyen TT, van den Wijngaard IR, Bosch J, van Belle E, van Zwet EW, Dofferhoff-Vermeulen T, et al. Comparison of prehospital scales for predicting large anterior vessel occlusion in the ambulance setting. *JAMA Neurol* 2021 Feb 01;78(2):157-164 [FREE Full text] [doi: [10.1001/jamaneurol.2020.4418](https://doi.org/10.1001/jamaneurol.2020.4418)] [Medline: [33252631](https://pubmed.ncbi.nlm.nih.gov/33252631/)]
62. Duvekot MH, Venema E, Rozeman AD, Moudrouts W, Vermeij FH, Biekart M, PRESTO investigators. Comparison of eight prehospital stroke scales to detect intracranial large-vessel occlusion in suspected stroke (PRESTO): a prospective observational study. *Lancet Neurol* 2021 Mar;20(3):213-221. [doi: [10.1016/S1474-4422\(20\)30439-7](https://doi.org/10.1016/S1474-4422(20)30439-7)] [Medline: [33422191](https://pubmed.ncbi.nlm.nih.gov/33422191/)]
63. Zhang Y, Zhou Y, Zhang D, Song W. A stroke risk detection: improving hybrid feature selection method. *J Med Internet Res* 2019 Apr 02;21(4):e12437 [FREE Full text] [doi: [10.2196/12437](https://doi.org/10.2196/12437)] [Medline: [30938684](https://pubmed.ncbi.nlm.nih.gov/30938684/)]
64. Lloyd T, Deeny SR, Steventon A. Weekend admissions may be associated with poorer recording of long-term comorbidities: a prospective study of emergency admissions using administrative data. *BMC Health Serv Res* 2018 Nov 16;18(1):863 [FREE Full text] [doi: [10.1186/s12913-018-3668-7](https://doi.org/10.1186/s12913-018-3668-7)] [Medline: [30445942](https://pubmed.ncbi.nlm.nih.gov/30445942/)]

65. McKinney JS, Deng Y, Kasner SE, Kostis JB, Myocardial Infarction Data Acquisition System (MIDAS 15) Study Group. Comprehensive stroke centers overcome the weekend versus weekday gap in stroke treatment and mortality. *Stroke* 2011 Sep;42(9):2403-2409. [doi: [10.1161/STROKEAHA.110.612317](https://doi.org/10.1161/STROKEAHA.110.612317)] [Medline: [21868723](https://pubmed.ncbi.nlm.nih.gov/21868723/)]
66. Mohammed MA, Sidhu KS, Rudge G, Stevens AJ. Weekend admission to hospital has a higher risk of death in the elective setting than in the emergency setting: a retrospective database study of national health service hospitals in England. *BMC Health Serv Res* 2012 Apr 02;12:87 [FREE Full text] [doi: [10.1186/1472-6963-12-87](https://doi.org/10.1186/1472-6963-12-87)] [Medline: [22471933](https://pubmed.ncbi.nlm.nih.gov/22471933/)]
67. Han L, Sutton M, Clough S, Warner R, Doran T. Impact of out-of-hours admission on patient mortality: longitudinal analysis in a tertiary acute hospital. *BMJ Qual Saf* 2018 Jun;27(6):445-454 [FREE Full text] [doi: [10.1136/bmjqs-2017-006784](https://doi.org/10.1136/bmjqs-2017-006784)] [Medline: [28971886](https://pubmed.ncbi.nlm.nih.gov/28971886/)]
68. Anderson D, Gao GG, Golden B. Life is all about timing: an examination of differences in treatment quality for trauma patients based on hospital arrival time. *Prod Oper Manag* 2014 Dec;23(12):2178-2190. [doi: [10.1111/poms.12236](https://doi.org/10.1111/poms.12236)]
69. Howard G, Kissela BM, Kleindorfer DO, McClure LA, Soliman EZ, Judd SE, et al. Differences in the role of black race and stroke risk factors for first vs. recurrent stroke. *Neurology* 2016 Feb 16;86(7):637-642 [FREE Full text] [doi: [10.1212/WNL.0000000000002376](https://doi.org/10.1212/WNL.0000000000002376)] [Medline: [26791153](https://pubmed.ncbi.nlm.nih.gov/26791153/)]
70. Jiménez MC, Manson JE, Cook NR, Kawachi I, Wassertheil-Smoller S, Haring B, et al. Racial variation in stroke risk among women by stroke risk factors. *Stroke* 2019 Apr;50(4):797-804 [FREE Full text] [doi: [10.1161/STROKEAHA.117.017759](https://doi.org/10.1161/STROKEAHA.117.017759)] [Medline: [30869565](https://pubmed.ncbi.nlm.nih.gov/30869565/)]
71. Casagrande VA, Diamond IT. Ablation study of the superior colliculus in the tree shrew (*Tupaia glis*). *J Comp Neurol* 1974 Jul;156(2):207-237. [doi: [10.1002/cne.901560206](https://doi.org/10.1002/cne.901560206)] [Medline: [4424699](https://pubmed.ncbi.nlm.nih.gov/4424699/)]
72. Bengtson E, Roth D. Understanding the value of features for coreference resolution. In: Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing. 2008 Oct Presented at: EMNLP '08; October 25-27, 2008; Honolulu, HI, USA p. 294-303. [doi: [10.3115/1613715.1613756](https://doi.org/10.3115/1613715.1613756)]
73. Hessel M, Modayil J, Van Hasselt H, Schaul T, Ostrovski G, Dabney W, et al. Rainbow: combining improvements in deep reinforcement learning. *Proc AAAI Conf Artif Intell* 2018 Apr 29;32(1):3215-3222. [doi: [10.1609/aaai.v32i1.11796](https://doi.org/10.1609/aaai.v32i1.11796)]
74. Lundberg SM, Erion GG, Lee SI. Consistent individualized feature attribution for tree ensembles. arXiv 2018.
75. Lundberg S, Lee SI. A unified approach to interpreting model predictions. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. 2017 Dec Presented at: NeurIPS '17; December 4-9, 2017; Long Beach, CA, USA p. 4768-4777.
76. Lundberg SM, Nair B, Vavilala MS, Horibe M, Eisses MJ, Adams T, et al. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. *Nat Biomed Eng* 2018 Oct;2(10):749-760 [FREE Full text] [doi: [10.1038/s41551-018-0304-0](https://doi.org/10.1038/s41551-018-0304-0)] [Medline: [31001455](https://pubmed.ncbi.nlm.nih.gov/31001455/)]
77. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. From local explanations to global understanding with explainable AI for trees. *Nat Mach Intell* 2020 Jan;2(1):56-67 [FREE Full text] [doi: [10.1038/s42256-019-0138-9](https://doi.org/10.1038/s42256-019-0138-9)] [Medline: [32607472](https://pubmed.ncbi.nlm.nih.gov/32607472/)]
78. Brandler ES, Sharma M, Sinert RH, Levine SR. Prehospital stroke scales in urban environments: a systematic review. *Neurology* 2014 Jun 17;82(24):2241-2249 [FREE Full text] [doi: [10.1212/WNL.0000000000000523](https://doi.org/10.1212/WNL.0000000000000523)] [Medline: [24850487](https://pubmed.ncbi.nlm.nih.gov/24850487/)]
79. Alijanpour S, Mostafazdeh-Bora M, Ahmadi Ahangar A. Different stroke scales; which scale or scales should be used? *Caspian J Intern Med* 2021;12(1):1-21 [FREE Full text] [doi: [10.22088/cjim.12.1.1](https://doi.org/10.22088/cjim.12.1.1)] [Medline: [33680393](https://pubmed.ncbi.nlm.nih.gov/33680393/)]
80. Gong X, Chen Z, Shi F, Zhang M, Xu C, Zhang R, et al. Conveniently-grasped field assessment stroke triage (CG-FAST): a modified scale to detect large vessel occlusion stroke. *Front Neurol* 2019 Apr 17;10:390 [FREE Full text] [doi: [10.3389/fneur.2019.00390](https://doi.org/10.3389/fneur.2019.00390)] [Medline: [31057480](https://pubmed.ncbi.nlm.nih.gov/31057480/)]
81. Antipova D, Eadie L, Macaden A, Wilson P. Diagnostic accuracy of clinical tools for assessment of acute stroke: a systematic review. *BMC Emerg Med* 2019 Sep 04;19(1):49 [FREE Full text] [doi: [10.1186/s12873-019-0262-1](https://doi.org/10.1186/s12873-019-0262-1)] [Medline: [31484499](https://pubmed.ncbi.nlm.nih.gov/31484499/)]
82. Vaid A, Somani S, Russak AJ, De Freitas JK, Chaudhry FF, Paranjpe I, et al. Machine learning to predict mortality and critical events in a cohort of patients with COVID-19 in New York city: model development and validation. *J Med Internet Res* 2020 Nov 06;22(11):e24018 [FREE Full text] [doi: [10.2196/24018](https://doi.org/10.2196/24018)] [Medline: [33027032](https://pubmed.ncbi.nlm.nih.gov/33027032/)]
83. Zanutto BS, Beck da Silva Etges AP, Dal Bosco A, Cortes EG, Ruschel R, De Souza AC, et al. Stroke outcome measurements from Electronic Medical Records: cross-sectional study on the effectiveness of neural and nonneural classifiers. *JMIR Med Inform* 2021 Nov 01;9(11):e29120 [FREE Full text] [doi: [10.2196/29120](https://doi.org/10.2196/29120)] [Medline: [34723829](https://pubmed.ncbi.nlm.nih.gov/34723829/)]
84. Goyal N, Tsvigoulis G, Male S, Metter EJ, Iftikhar S, Kerro A, et al. FABS: an intuitive tool for screening of stroke mimics in the emergency department. *Stroke* 2016 Sep;47(9):2216-2220. [doi: [10.1161/STROKEAHA.116.013842](https://doi.org/10.1161/STROKEAHA.116.013842)] [Medline: [27491733](https://pubmed.ncbi.nlm.nih.gov/27491733/)]
85. Vest JR, Ben-Assuli O. Contribution of different data sources to the prediction of emergency department revisits in a safety-net population. In: Proceedings of the 2018 International Conference on Information Systems. 2018 Presented at: ICIS '18; December 13-16, 2018; San Francisco, CA, USA p. 3.
86. Kasthurirathne SN, Vest JR, Menachemi N, Halverson PK, Grannis SJ. Assessing the capacity of social determinants of health data to augment predictive models identifying patients in need of wraparound social services. *J Am Med Inform Assoc* 2018 Jan 01;25(1):47-53 [FREE Full text] [doi: [10.1093/jamia/ocx130](https://doi.org/10.1093/jamia/ocx130)] [Medline: [29177457](https://pubmed.ncbi.nlm.nih.gov/29177457/)]



87. Tsao CW, Aday AW, Almarzooq ZI, Alonso A, Beaton AZ, Bittencourt MS, et al. Heart disease and stroke statistics-2022 update: a report from the American Heart Association. *Circulation* 2022 Feb 22;145(8):e153-e639 [FREE Full text] [doi: [10.1161/CIR.0000000000001052](https://doi.org/10.1161/CIR.0000000000001052)] [Medline: [35078371](https://pubmed.ncbi.nlm.nih.gov/35078371/)]
88. Ben-Assuli O, Padman R. Trajectories of repeated readmissions of chronic disease patients: risk stratification, profiling, and prediction. *MIS Q* 2020 Jan 01;44(1):201-226. [doi: [10.25300/misq/2020/15101](https://doi.org/10.25300/misq/2020/15101)]

## Abbreviations

**ACS:** American Community Survey  
**AUC:** area under the receiver operating characteristic curve  
**CT:** computed tomography  
**ED:** emergency department  
**EHR:** electronic health record  
**GBM:** gradient boosting machine  
**ML:** machine learning  
**RF:** random forest  
**SDoH:** social determinants of health  
**SHAP:** Shapley Additive Explanations

*Edited by G Eysenbach; submitted 14.02.22; peer-reviewed by S Somanchi, M Ayavaci, N Marotta; comments to author 06.04.22; revised version received 17.07.22; accepted 18.12.22; published 30.01.23*

*Please cite as:*

*Chen M, Tan X, Padman R*

*A Machine Learning Approach to Support Urgent Stroke Triage Using Administrative Data and Social Determinants of Health at Hospital Presentation: Retrospective Study*

*J Med Internet Res* 2023;25:e36477

URL: <https://www.jmir.org/2023/1/e36477>

doi: [10.2196/36477](https://doi.org/10.2196/36477)

PMID:

©Min Chen, Xuan Tan, Rema Padman. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 30.01.2023. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.