

Article

Concurrent Prediction of Length of Stay, Mortality, and Total Charges in Patients with Acute Lymphoblastic Leukemia Using Continuous Machine Learning

Jiahui Ma ¹, Elizabeth Johnson ² , Bradley M. Whitaker ³ , Faraz Dadgostari ¹ , Hansjorg Schwertz ⁴
and Bernadette McCrory ^{1,5,*} 

¹ Department of Mechanical and Industrial Engineering, Norm Asbjornson College of Engineering, Montana State University, Norm Asbjornson Hall, Bozeman, MT 59717, USA; jiahuima@montana.edu (J.M.)

² Mark & Robyn Jones College of Nursing, Montana State University, Sherrick Hall 212, Bozeman, MT 59717, USA

³ Department of Electrical and Computer Engineering, Norm Asbjornson College of Engineering, Montana State University, Norm Asbjornson Hall, Bozeman, MT 59717, USA; bradley.whitaker1@montana.edu

⁴ Department of Occupational Medicine, Billings Clinic Bozeman, Bozeman, MT 59717, USA

⁵ BioReD Hub, Montana State University, Bozeman, MT 59717, USA

* Correspondence: bernadette.mccrory@montana.edu

Abstract

Acute lymphoblastic leukemia (ALL) presents significant clinical challenges due to its genetic complexity and high relapse rates. While outcomes like length of stay (LOS), mortality, and total charges (TCs) are critical quality indicators, most existing models rely on static data and separate outcome modeling. This study utilized the HCUP National Inpatient Sample (NIS) to develop a dynamic, concurrent prediction model for prolonged LOS and mortality (PLOSM), alongside a framework for TCs. By integrating temporally updated patient information, the concurrent approach outperformed single-outcome models. Within the first seven days of hospitalization, the model achieved accuracy and precision above 90%, with recall and F1-scores exceeding 80%. Key predictors of these outcomes included age, race, insurance type, financial indicators, and elective surgery status. Notably, both prolonged LOS and mortality were significant drivers of TCs. By bridging predictive modeling and real-time clinical data, this framework enables data-driven decision-making to optimize patient management, enhance safety, and mitigate the financial burden of ALL care.

Keywords: patient outcome concurrent prediction; machine learning; dynamic prediction; acute lymphoblastic leukemia; healthcare cost and utilization project



Academic Editor: Jiang Bian

Received: 15 January 2026

Revised: 5 March 2026

Accepted: 20 March 2026

Published: 24 March 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Blood cancer is the malignancy of hematopoietic stem cells, leading to the uncontrolled growth of immature stem cells to suppress the production of healthy white blood cells, red blood cells, or platelets [1,2]. Unlike solid tumors forming localized masses, the blood cancer can blast from the bone marrow and spread through the bloodstream, lymph system, and central nervous system [1]. Acute lymphoblastic leukemia (ALL) is a malignant proliferation of lymphoid cells that progresses rapidly and can occur both in children and adults, with a peak incidence between 1 year and 4 years [1,3]. Early symptoms of ALL, such as tiredness, fever, recurrent infections, bone discomfort, and loss of weight, can be

subtle, such that they are often mistaken for less serious conditions or side effects of other chronic illnesses or events [1,4].

1.1. Challenges for ALL Treatments

Acute leukemia presents significant challenges in diagnosis and treatment due to its genetic complexity, high relapse rates, drug resistance, and treatment-related toxicities, particularly in the central nervous system [5,6]. Advances in molecular research have improved risk stratification and targeted therapies, but overcoming resistance mechanisms and long-term complications remains a critical focus for future research [5,6]. These challenges place patients with ALL at significant disadvantages to timely diagnosis and early treatment, necessitating discussions surrounding clinical trial involvement much earlier in the disease stage [7]. While conventional chemotherapy yields high cure rates in pediatric ALL, adult patients, often presenting with higher-risk factors, comorbidities, and advanced age, experience significantly lower survival rates and clinical trials originally designed for pediatric patients due to the suboptimal efficacy of these treatments [5,8]. The integration of trial-specific patient safety and health considerations has been historically challenging in clinical care settings, such as inpatient hospital stays or routine outpatient visits. Treatments for ALL may cause known and unknown side effects or detrimental changes in body function, such as cardiovascular disease, secondary malignancies, endocrine complications (obesity, osteoporosis, infertility, and premature menopause), and alterations in cognition [9]. Meanwhile, there is an identified significant financial burden due to misalignment of treatment costs and insurance gaps, which may be amplified during active clinical trial participation if the organization is unfamiliar with clinical trial benefit submissions and patient rights when on a clinical trial [10]. Early financial situation screening, implementing financial navigation programs, and integrating discussions about costs into clinical practice are proposed as strategies to mitigate financial toxicity [10]. Through the Clinical Treatment Act, some standard of care exams, diagnostics, and imaging may be covered, provided the healthcare organization is competent in trial-specific billing and coding [11].

1.2. Data-Driven Methods to Enhance Outcomes of Patients with ALL in Inpatient Clinical Settings

Due to challenges in early identification, rapid progression, limited treatment, and high burdens, care coordination is vital for timely intervention. Unexpected trial-related hospitalizations can complicate treatment evaluation and increase risks for hospital-related complications (HACs) or in-hospital mortality [12]. Deepened understanding and contextualization of the anticipated progression of a patient with ALL can further enhance the potential for facilitating promising treatment options, timely enrollment in clinical trials, and improvement of patient outcomes.

Patient length of stay (LOS) and mortality are crucial metrics for evaluating the efficiency, quality, and cost-effectiveness of inpatient care [13–15]. In an inpatient setting, LOS, defined as the duration of an inpatient stay, and mortality, the occurrence of death during inpatient care, have been extensively studied [16,17]. Extended inpatient stays, known as prolonged length of stay (PLOS), can not only constrain resource allocation but also elevate the risk of HACs [18–20]. All-cause mortality, which is defined as the total number of deaths from any cause within a specific population over a given period, is the most undesired endpoint for patients in the hospital, making it a crucial metric for hospital assessment [21,22].

Patient total charges (TCs) include the charges for all billed services and procedures during an inpatient stay [23]. TCs for inpatient care encapsulate both the resource consumption, including labor, supplies, and medications, as well as the economic burden imposed on patients and their families during the care process [23]. TCs may vary if a

patient with ALL is in a clinical trial versus those who follow a standard of care pathway, as clinical trials require heightened monitoring through specimen collection, procedures, and imaging [12,24]. All three outcomes can help clinical facilities, such as hospitals, identify opportunities to improve patient safety and quality of care, reduce financial toxicity, and improve patient satisfaction with research-integrated clinical care, known as Clinical Research as a Care Option (CRAACO) [24].

Although patient outcome prediction has been widely studied, limited studies have been completed on predicting inpatient outcomes for patients with ALL [25–27]. LOS and mortality prediction for other diseases has been a subject of extensive research, with methodologies continually advancing to a noticeable shift towards utilizing more advanced machine learning models in recent years [17,28–40]. Currently, most prediction models on LOS and in-hospital mortality were performed separately, even though these models were completed in the same research [38,39]. The misleading results of using separate prediction models for associated outcomes, the lack of transparency of the data source, and poor interpretation of the model deteriorate the trust of healthcare professionals to adopt the model [28,39]. While a short length of stay is generally a positive indicator for hospital efficiency, some patients with acute conditions may deteriorate rapidly and pass away shortly after admission. Therefore, the need for concurrent prediction and usage of LOS and in-hospital mortality was identified as another gap in the current research [39]. TCs' prediction witnessed an improving performance with advanced machine learning models, which was also highly determined by LOS [41–44]. However, most models utilize data, such as baseline patient characteristics or post-discharge outcomes, to make static predictions [17,28–30,34,37,40]. Most patient prediction models for hospitalization were performed at admission; therefore, patient treatment information was overlooked in current models [17,28–40]. Dynamic prediction models have the potential to revolutionize clinical decision-making by adapting to evolving patient conditions, leading to more accurate and timely predictions [45–47].

For patient outcome prediction, classification models, such as logistic regression (LR), K-Nearest Neighbors (KNN), Random Forest (RF), Support Vector Machines (SVM), Gradient Boost (GB), and Neural Networks (NN), have been popularly applied [28,31,34,36,37]. LR has been reported to be most frequently used for classification prediction compared with other machine learning models, but ensemble models (such as RF and GB) have been reported to have better performance in predicting LOS and mortality [31,34,36,37,41,43,44,48]. RF and GB are selected to predict PLOS, mortality, and TCs. RF and GB are powerful ensemble methods that combine multiple decision trees to create robust and accurate predictive models [49,50]. RF is an ensemble machine learning method that utilizes bootstrap sampling and feature randomization to aggregate multiple decision trees to perform classification prediction [50]. GB is an ensemble machine learning method to optimize the model by sequentially adding decision trees through gradient descent to minimize the residuals from the previous model [49].

1.3. Purpose

There is a need to predict more clinically meaningful and reliable patient outcomes, enhance predictive accuracy, and enable timely interventions for patients with ALL. This research proposed the dynamic clinical outcome prediction model, which leverages a comprehensive set of patient-level factors identified through previous examination and literature review.

To develop this model, the following aims will be completed:

1. Adapt the model to dynamically predict patient outcomes using the most up-to-date available patient information.

2. Compare the performance of separate patient outcome predictions for LOS and mortality with both RF and GB in a concurrent prediction model that evaluates both outcomes simultaneously.
3. Implement a continuous prediction method to assess TCs by integrating the most accurate predicted results from LOS and mortality prediction.

2. Methods

This predictive analysis utilized machine learning algorithms to predict inpatient outcomes based on a retrospective database: Healthcare Cost and Utilization Project (HCUP), which is developed by the Agency for Healthcare Quality (AHRQ). The HCUP, hosting the largest collection of administrative and longitudinal hospital care data in the United States, was selected for building the inpatient LOS, mortality, and TCs prediction model [51].

The HCUP National (Nationwide) Inpatient Sample (NIS) is an all-payer inpatient healthcare database and comprises inpatient discharge billing data [51]. The NIS includes four discharge-level files: Core, Severity, Hospital, and Diagnosis and Procedure Groups:

- Core File: demographics, expected primary payer, TCs, discharge status, financial status, and International Classification of Diseases 10th Revision (ICD-10) coding for diagnoses and procedures [52].
- Severity File: illness severity and mortality risk for each discharge record, utilizing the All-Patient Refined Diagnosis-Related Group (APRDRG) system, which is assigned using software from 3M Health Information Systems [53].
- Hospital File: characteristics of each hospital, including hospital location, ownership, and size.
- Diagnosis and Procedure Groups File: patient's comorbidities presented at admission were defined by the Elixhauser Comorbidity Software (version 2021.1) Refined for ICD-10-CM in this file [54].

Information related to other outcomes and sampling information from NIS were not included in this study. The clinical trial designation code (ICD-10-CM Z00.6—Encounter for examination for normal comparison and control in clinical research program) was considered in this study to identify patients in clinical trial research, where clinical services are being reimbursed through trial benefits. Data retrieval and slicing were performed using IBM SPSS Statistics (version 29). All statistical analysis and machine learning prediction modeling were performed in Python (version 3.10.0).

2.1. Patient Outcomes Prediction

The 75th percentile of the distribution of the total length of hospital stay for the study population is widely applied in other studies and selected in this study to separate the normal and prolonged hospital stay [55,56]. A combination of levels in PLOS and mortality was used as new levels in the combined outcome, including mortality, normal, PLOS, and both (Figure 1). Individual predictions of PLOS and mortality were performed, whose prediction results were combined and then compared with the prediction of PLOSM. TCs were treated as a continuous outcome. The performance of predicting PLOS and mortality individually against a combined outcome model was evaluated; the better prediction outcomes were then utilized as input features for TC prediction (Figure 1).

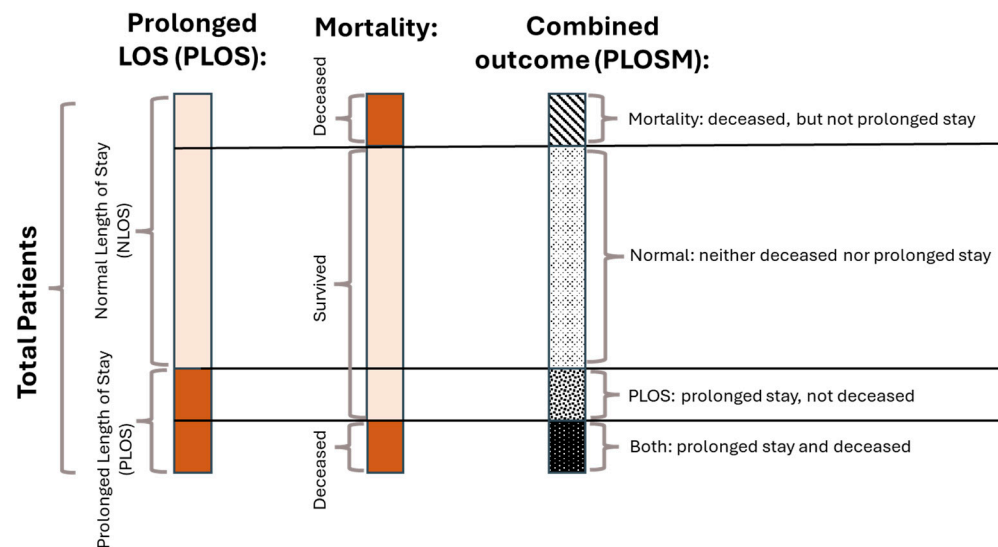


Figure 1. Patient grouping by individual and combined patient outcomes.

2.2. Data Slicing and Cleaning

NIS data were collected from the years 2019 to 2021. Patients with ALL diagnoses were identified using the ICD-10-CM diagnosis codes, including C9100 (ALL not received remission), C9101 (ALL in remission), and C9102 (ALL in relapse). A total of 20,371 discharge records were filtered from the 2019 and 2021 NIS databases (10,452 in 2019 and 9919 in 2021) (Figure 2). Missing values accounted for less than 3% of the total sample size, 180 ($n = 20,371$). These values were addressed using single-point imputation derived from the full study population: categorical features were imputed using the mode, while continuous variables were imputed using the median (Figure 2). ICD-10 diagnosis (DX) and procedure (PR) codes (up to 40 DX, 25 PR per record, with procedure day) were initially processed as strings. Low-frequency codes (<1%) were removed, and the remaining 253 DX and 29 PR codes were converted to binary features. A team review selected 51 ALL-relevant DX and 29 PR codes (Appendix A Table A1). The initial dataset of 20,371 unique patients was expanded to 23,713 records by restructuring the data into a long format (Figure 2). In this format, each record represents an individual procedure; consequently, patients who underwent multiple procedures appear as multiple distinct entries in the expanded dataset.

2.3. Statistical Analysis

Statistical analyses included descriptive statistics (mean/SD, frequency) and inferential analysis, including Pearson correlation and Analysis of Variance (ANOVA), to examine feature relationships. The significance level was set to 0.05 for all inferential analyses.

2.4. Feature Preparation and Selection

Categorical variables were converted to dummy variables (low-frequency categories excluded). Recursive feature elimination (RFE), combined with 5-fold cross-validation (CV), was used to select relevant features and reduce overfitting (Figure 3) [57–60].

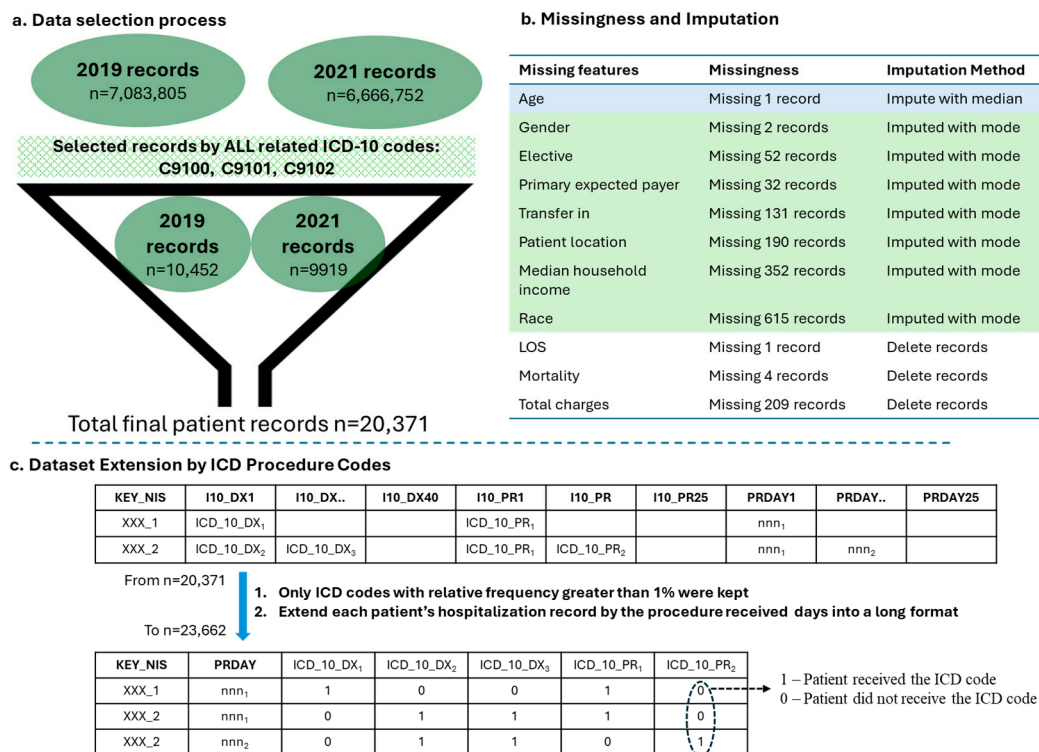


Figure 2. Data selection process, cleaning, and restructuring process. (a) Patient record selection process by ALL-related ICD-10 codes from NIS in 2019 and 2021. (b) Feature missingness and imputation methods summary. (c) Dataset extension methodology illustration to show how the dataset was extended into a long format by multiple ICD-10 procedure codes in each patient record. Note: Blue error represents the dataset was transformed from short format to the long format.

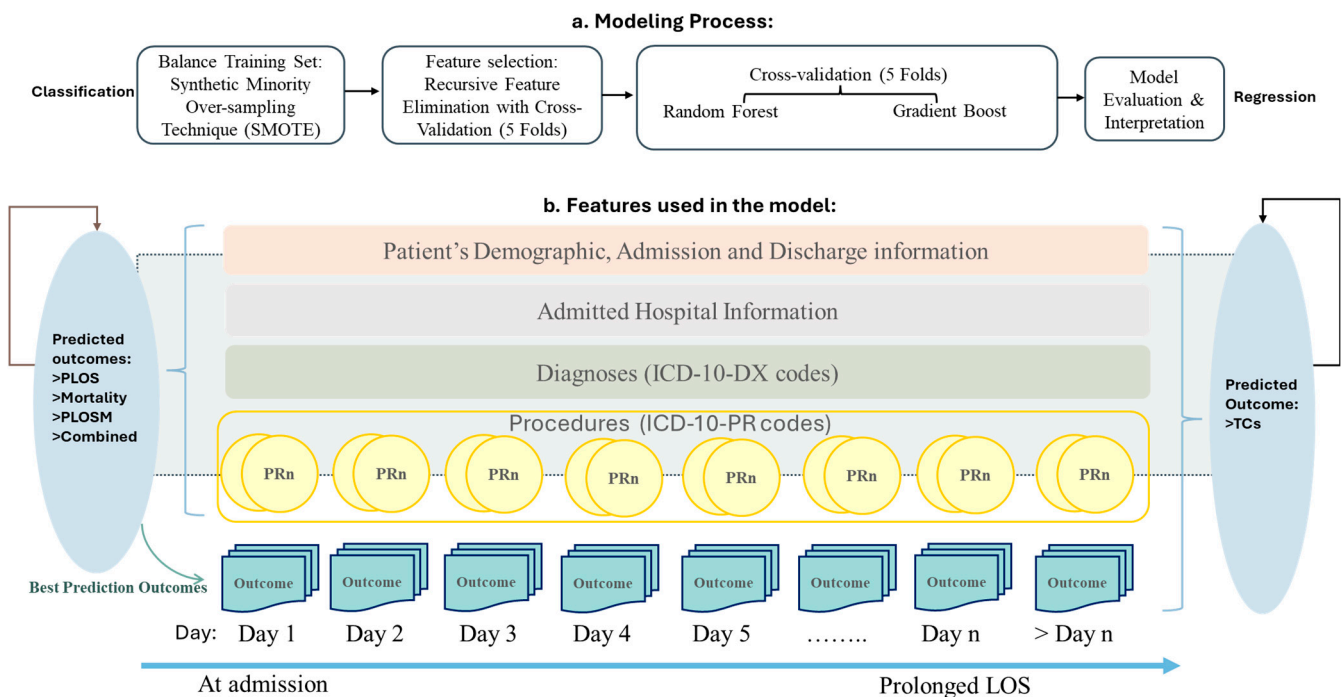


Figure 3. Patient outcome prediction process. (a) A machine learning technique flowchart to show how data balancing, feature selection, prediction model, and model evaluation and interpretation were performed. (b) Categorization of static and dynamic features utilized for prediction.

2.5. Prediction Modeling

Both RF and GB methods are designed for handling high-dimensional datasets and selected for this study (Figure 3) [36]. Five-fold cross-validation was applied to each model to enhance model performance and reduce over-fitting [59]. This study used a dynamic prediction process to continuously model patient outcomes from the first day of hospitalization up to the point defining prolonged LOS. For each prediction day, the feature set was reconstructed using only information available up to that day by masking procedures occurring afterward. The dataset was then randomly split at the hospitalization-record level into 80% training and 20% testing sets. SMOTE resampling, feature selection (RFE), and feature scaling were performed exclusively on the training set to prevent data leakage. Model performance was evaluated using 5-fold cross-validation within the training data and further validated on the independent test set. Models were trained independently for each prediction day. Two ensemble learning algorithms were evaluated, Random Forest ($n_estimators = 100$, $random_state = 42$) and Gradient Boosting ($n_estimators = 100$, $random_state = 42$), with other hyperparameters kept at default values in scikit-learn. For each prediction day, both models were trained independently and evaluated using 5-fold cross-validation within the training set. Final performance was assessed on an independent 20% held-out test set.

2.6. Model Evaluation and Interpretation

Model performance was assessed using multiple evaluation standards (Figure 3). Model accuracy, precision, recall, and F1-score were used to evaluate classification models for predicting PLOS and mortality [61]. Additionally, the area under the receiver operating characteristic (ROC) curve was computed to evaluate the model's ability to discriminate between classes [61].

- Precision = True Positives / (True Positives + False Positives)
- Recall = True Positives / (True Positives + False Negatives)
- F1-score = $2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$
- Accuracy = (True Positives + False Negatives) / (Total Number of Predictions)

Coefficient of determination (R^2) measures the proportion of variance in the dependent variable explained by the model [62]. Root mean square error (RMSE) is the square root of the average of the squared differences between predictions and actual observations [62]. Mean absolute error (MAE) calculates the average absolute difference between predicted and actual values [62]. All three metrics were used to evaluate TC prediction [62].

The feature importance method can quantify the contribution of each individual feature to help identify the most influential features for the prediction outcome [63]. Feature importance was assessed using Gini importance, a measure of how much each feature contributes to reducing impurity in the trees of the ensemble model [64]. A feature importance score was assigned to each feature, and the top 30 important features from each day's prediction were selected to interpret the prediction performance in this study [63].

3. Results

This study included 20,371 patients with ALL from the 2019 (10,452) and 2021 (9919) NIS databases. LOS (Mean: 8.9, SD: 13.3) and TCs (Mean: 152,358.2, SD: 334,148.1 \$) were skewed; mortality was imbalanced (19,887 deceased, 97.6%; 480 discharged, 2.4%) (Figure 4). PLOS was defined as >7 days (75th percentile after outlier removal), resulting in 13,198 prolonged stays (64.8%) and 7173 normal stays (35.2%) (Figure 4). PLOS and mortality were highly associated with TCs ($p < 0.001$, ANOVA); PLOS was also associated with mortality ($p < 0.001$, Chi-square). Age (Mean: 28.6, SD: 23.6) had a weak correlation with TCs ($r = 0.008$, $p > 0.05$). Patients with prolonged stay (Mean: 31.6, SD: 23.8) were

older than those with normal stay (Mean: 27.0, SD: 23.4); deceased patients (Mean: 45.7, SD: 23.5) were older than discharged patients (Mean: 28.2, SD: 24.1).

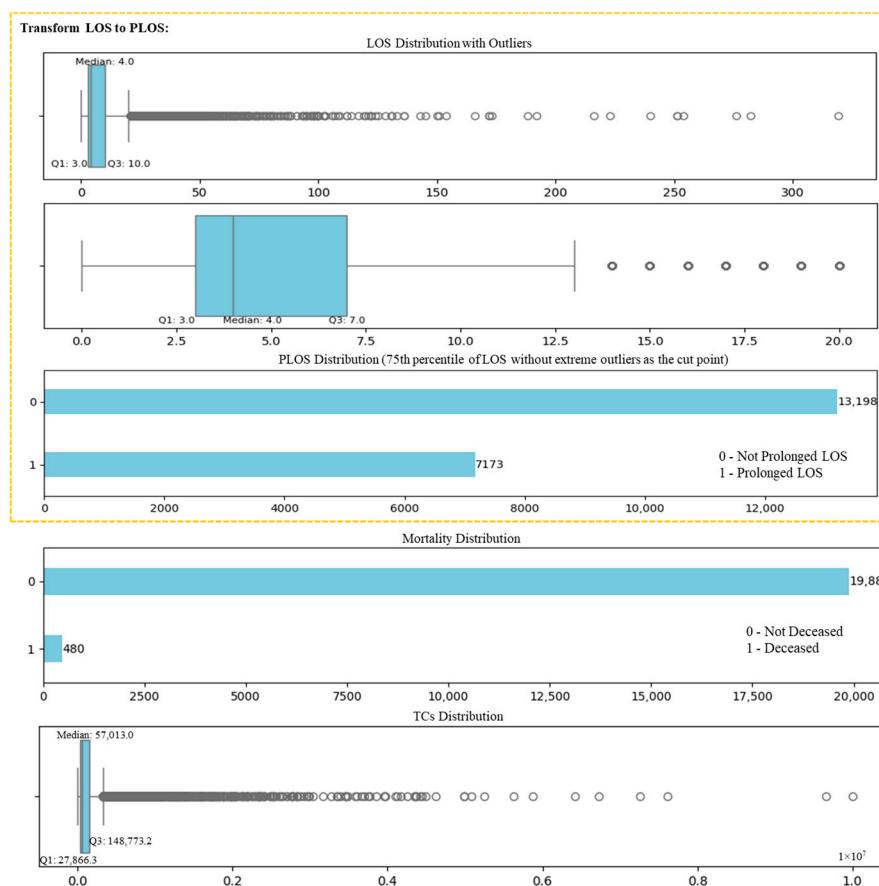


Figure 4. Distribution of Patient Outcomes.

Patient demographics, financial status, hospital, and medical information were summarized by outcome (Table 1). Low-frequency (<1%) binary features were removed, and potential multicollinearity features were excluded. In total, 136 features were used for modeling.

Table 1. Feature summaries ($n = 20,371$).

Variable	PLOS ¹ ($n = 7173$)	Mortality ¹ ($n = 480$)	TC (\$) ²
Minor (<18 years of age)	2868 (14.08)	88 (0.43)	139.62K (368.51K, 43.23K)
2019	3617 (17.76)	237 (1.16)	143.16K (320.64K, 52.41K)
2021	3556 (17.46)	243 (1.19)	162.16K (347.72K, 61.59K)
Demographic information			
Female	3140 (15.41)	204 (1.00)	153.73K (344.77K, 58.02K)
Race/Ethnicity			
White	3687 (18.10)	256 (1.26)	132.48K (286.21K, 51.99K)
Black	579 (2.84)	43 (0.21)	143.99K (301.82K, 52.59K)
Hispanic	2173 (10.67)	137 (0.67)	188.32K (418.74K, 65.19K)
Asian or Pacific Islander	334 (1.64)	19 (0.09)	160.08K (291.52K, 65.44K)
Native American	63 (0.31)	1 (0.01)	119.53K (201.62K, 57.86K)
Other	337 (1.65)	24 (0.12)	170.26K (345.47K, 67.39K)
Insurance and financial information			

Table 1. Cont.

Variable	PLOS ¹ (n = 7173)	Mortality ¹ (n = 480)	TC (\$) ²
Expected primary payer			
Medicare	1115 (5.47)	146 (0.72)	139.06K (255.48K, 66.88K)
Medicaid	2426 (11.91)	120 (0.59)	160.85K (365.29K, 55.91K)
Private insurance	3013 (14.79)	170 (0.83)	145.97K (311.71K, 54.25K)
Self-pay	256 (1.26)	18 (0.09)	162.33K (376.78K, 58.61K)
No charge	41 (0.20)	3 (0.01)	107.24K (194.46K, 51.68K)
Other	322 (1.58)	23 (0.11)	187.09K (454.80K, 57.25K)
Area ³			
Urban	4237 (20.80)	281 (1.38)	164.43K (343.39K, 63.82K)
Transitional	2028 (9.96)	135 (0.66)	138.75K (330.60K, 48.45K)
Rural	908 (4.46)	64 (0.31)	129.44K (296.36K, 47.23K)
Median household income			
0–25th percentile	2038 (10.00)	144 (0.71)	156.65K (372.86K, 55.23K)
26–50th percentile	1683 (8.26)	117 (0.57)	146.11K (316.77K, 54.87K)
51–75th percentile	1832 (8.99)	128 (0.63)	152.47K (350.77K, 56.75K)
76–100th percentile	1620 (7.95)	91 (0.45)	153.37K (279.23K, 60.98K)
Hospital information			
Location/teaching status of hospital			
Rural	37 (0.18)	5 (0.02)	130.77K (240.87K, 52.40K)
Urban. nonteaching	157 (0.77)	28 (0.14)	153.92K (349.08K, 56.60K)
Urban. teaching	6979 (34.26)	447 (2.19)	205.19K (338.87K, 57.69K)
Control/ownership of hospital			
Government. non-federal	1235 (6.06)	80 (0.39)	166.80K (368.99K, 65.20K)
Private. not-for-profit	5517 (27.08)	375 (1.84)	119.42K (252.96K, 45.44K)
Private. investor-owned	421 (2.07)	25 (0.12)	136.84K (355.77K, 90.74K)
Region of hospital			
Northeast	1231 (6.04)	71 (0.35)	166.80K (368.99K, 65.20K)
Midwest	1282 (6.04)	78 (0.38)	119.42K (252.96K, 45.44K)
South	2743 (13.47)	209 (1.03)	136.84K (296.84K, 50.47K)
West	1917 (9.41)	122 (0.60)	193.88K (409.97K, 74.67K)
Bed-size of hospital ⁴			
Small	857 (4.21)	53 (0.26)	119.54K (327.76K, 43.76K)
Medium	1345 (6.6)	109 (0.54)	160.82K (383.67K, 54.54K)
Large	4971 (24.40)	318 (1.56)	156.96K (318.96K, 61.38K)
Medical Information			
Admission on weekend	1262 (6.20)	98 (0.48)	172.63K (357.39K, 64.65K)
Elective surgery	1941 (9.53)	78 (0.38)	150.83K (368.14K, 50.77K)
Injury (incidence)	249 (1.22)	28 (0.14)	182.39K (347.40K, 70.38K)
Service Line (based on ICD-10)			
Surgical	862 (4.23)	83 (0.41)	434.47K (747.21K, 194.99K)
Medical	6258 (30.72)	394 (1.93)	133.37K (273.98K, 53.11K)
Transfer into the hospital			
Not transferred in	5912 (29.02)	358 (1.76)	143.15K (332.04K, 53.24K)
From an acute care hospital	1054 (5.17)	97 (0.48)	239.41K (353.58K, 137.21K)
From another health faculty	207 (1.02)	25 (0.12)	185.02K (278.09K, 75.65K)

Table 1. Cont.

Variable	PLOS ¹ (n = 7173)	Mortality ¹ (n = 480)	TC (\$) ²
Risk Mortality			
No class specified	2 (0.01)	1 (0.01)	1346.15K (2493.40K, 138.43K)
Minor likelihood of dying	1620 (7.95)	8 (0.04)	97.20K (156.19K, 41.76K)
Moderate likelihood of dying	2622 (12.87)	11 (0.05)	106.37K (179.28K, 49.11K)
Major likelihood of dying	1767 (8.67)	63 (0.31)	218.28K (346.93K, 98.24K)
Extreme likelihood of dying	1162 (5.70)	397 (1.95)	512.87K (863.67K, 243.41K)
Risk Severity			
No class specified	2 (0.01)	1 (0.01)	1346.15K (2493.40K, 138.43K)
Minor loss of function	88 (0.43)	0 (0)	82.09K (133.43K, 40.37K)
Moderate loss of function	1169 (5.74)	5 (0.02)	66.19K (105.04K, 35.50K)
Major loss of function	3256 (15.98)	34 (0.17)	133.22K (201.46K, 62.75K)
Extreme loss of function	2658 (13.05)	440 (2.16)	373.36K (638.22K, 173.41K)
Emergency department record	2906 (14.27)	229 (1.12)	147.17K (296.75K, 61.49K)
Major operating room	1341 (6.58)	94 (0.46)	396.17K (672.25K, 201.50K)
Enrolled in clinical trial	287 (1.41)	6 (0.03)	253.49K (583.56K, 109.19K)
Comorbidity_Acquired immune deficiency syndrome	20 (0.1)	2 (0.01)	272.57K (382.86K, 88.84K)
Comorbidity_Alcohol abuse	55 (0.27)	3 (0.01)	151.91K (263.87K, 75.48K)
Comorbidity_Dementia	33 (0.16)	9 (0.04)	116.04K (272.80K, 58.48K)
Comorbidity_Depression	761 (3.74)	44 (0.22)	185.82K (387.38K, 73.14K)
Comorbidity_Diabetes with chronic complications	834 (4.09)	67 (0.33)	194.11K (346.32K, 85.67K)
Comorbidity_Diabetes without chronic complications	391 (1.92)	28 (0.14)	135.58K (219.23K, 34.11K)
Comorbidity_Drug abuse	121 (0.59)	6 (0.03)	206.22K (374.59K, 37.25K)
Comorbidity_Hypertension. complicated	745 (3.66)	84 (0.41)	222.92K (466.18K, 88.50K)
Comorbidity_Hypertension. uncomplicated	1694 (8.32)	116 (0.57)	174.59K (312.60K, 72.87K)
Comorbidity_Chronic pulmonary disease	660 (3.24)	50 (0.25)	147.02K (305.98K, 60.90K)
Comorbidity_Obesity	910 (4.47)	76 (0.37)	203.48K (411.54K, 82.96K)
Comorbidity_Peripheral vascular disease	366 (1.80)	26 (0.13)	128.06K (321.99K, 48.75K)
Comorbidity_Hypothyroidism	478 (2.35)	30 (0.15)	153.07K (285.46K, 64.58K)
Comorbidity_Other thyroid disorders	79 (0.39)	3 (0.01)	192.54K (348.39K, 75.10K)

¹ Frequency (relative frequency %); ² Mean (Standard Deviation, Median); ³ Regroup of the six-category urban–rural classification scheme for the U.S. counties developed by the National Center for Health Statistics (NCHS). Urban: “Central/Fringe” counties of metro areas of ≥ 1 million population; Transitional: Counties in metro areas of 50,000–999,999 population; Rural: Micropolitan/Nonmetropolitan counties; ⁴ Bed-size categories are based on hospital beds and are specific to the hospital’s location and teaching status.

3.1. Model Performance

3.1.1. Individual Prediction of PLOS and Mortality

For the individual patient outcome prediction of PLOS and mortality, the mortality prediction models with higher AUC scores (RF AUC range: 0.9630–0.9956; GB AUC range:

0.9614–0.9709) outperformed the PLOS prediction (RF AUC range: 0.8913–0.9677; GB AUC range: 0.8641–0.8955) for all days' prediction (Figure 5).

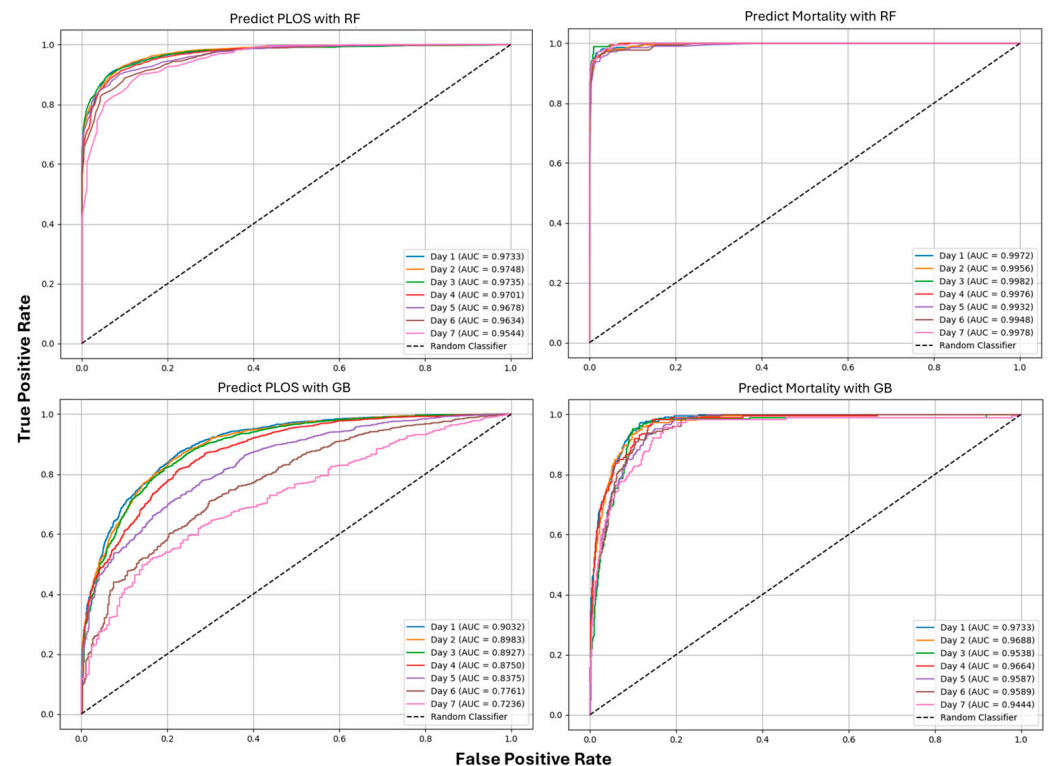


Figure 5. ROC Curve across Days for Individual PLOS and Mortality Prediction.

This superiority was also reflected in accuracy, precision, recall, and F1-scores (Appendix B Table A2). RF generally performed better than GB for both outcomes, exhibiting higher AUC values (Figure 5) and generally higher accuracy, precision, recall, and F1-scores (Appendix B Table A2).

3.1.2. Prediction of Combined Outcome (PLOSOM)

As detailed in Figure 1, there were four classes in the combined outcome (PLOSOM) of PLOS and mortality (normal: neither prolonged stay nor deceased; PLOS only: prolonged stay without deceased; mortality only: deceased without prolonged stay; and both: prolonged stay and deceased). ROC curves were generated for each class using both RF and GB models (Figure 6). Across all classes and predictions, RF models demonstrated superior performance compared to GB models, as evidenced by higher AUC scores (Figure 6).

Comparing performance metrics, PLOSOM prediction achieved higher accuracy, precision, recall, and F1-scores than PLOS prediction, but lower scores than mortality prediction (Appendix C Table A3). Importantly, the PLOSOM prediction models exhibited strong performance in predicting the minority classes: mortality only (RF AUC range: 0.9718–0.9978) and the combined PLOSOM outcome (RF AUC range: 0.9626–0.9987).

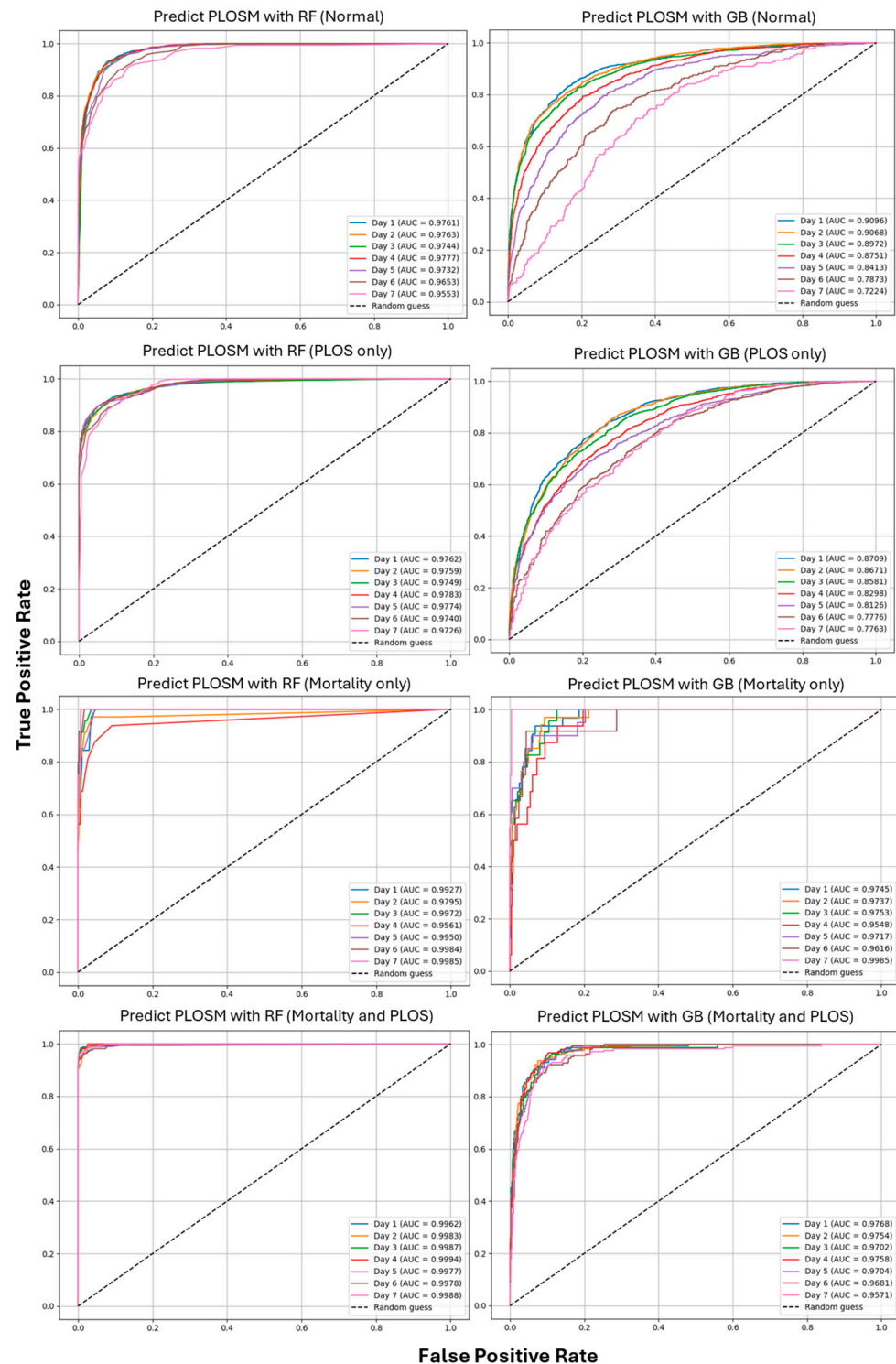


Figure 6. ROC Curve across Days for PLOSM Prediction in each class.

3.1.3. Comparing the Prediction of PLOSM with the Combined Outcome Prediction of PLOS and Mortality

A comparison was conducted between individual predictions of PLOS and mortality and the combined PLOSM prediction. To facilitate this comparison, the results of the individual PLOS and mortality predictions were combined, and performance metrics were recalculated specifically for the testing set (Appendix C Table A3). This combination method precluded the calculation of training set metrics (accuracy, precision, recall, and F1-score) (Appendix C Table A3). Comparing the testing set metrics (Figure 7) revealed that the RF

PLOSM prediction generally achieved higher accuracy, F1-scores, and recall. However, the combined PLOS and mortality prediction demonstrated superior performance on day 1 and after day 5 (Figure 7).

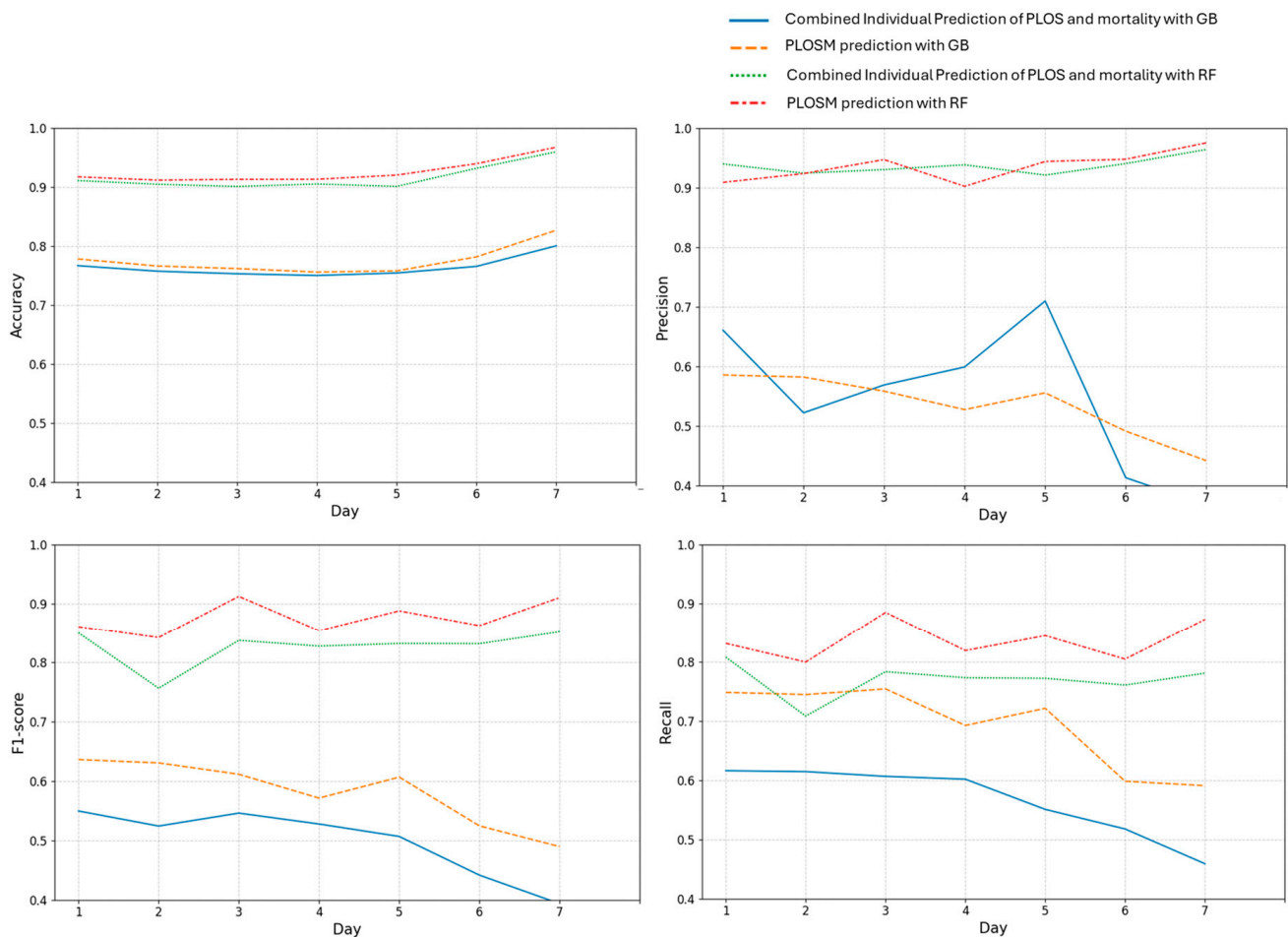


Figure 7. Model Performance Metrics for Combined and PLOSM Outcome.

3.1.4. Prediction of TCs

TCs' prediction with RF outperformed the prediction with GB for both training and testing sets across all evaluation metrics, including RMSE, MAE, and R^2 (Figure 8). Notably, RF's RMSE and MAE values decreased as the prediction day advanced, whereas GB's RMSE and MAE values exhibited an increasing trend (Figure 8). This resulted in RF consistently achieving lower RMSE and MAE values than GB, with the difference between the two models widening as the prediction day increased (Figure 8). While both models showed improved R^2 values with increasing prediction day, RF consistently achieved higher R^2 values than GB.

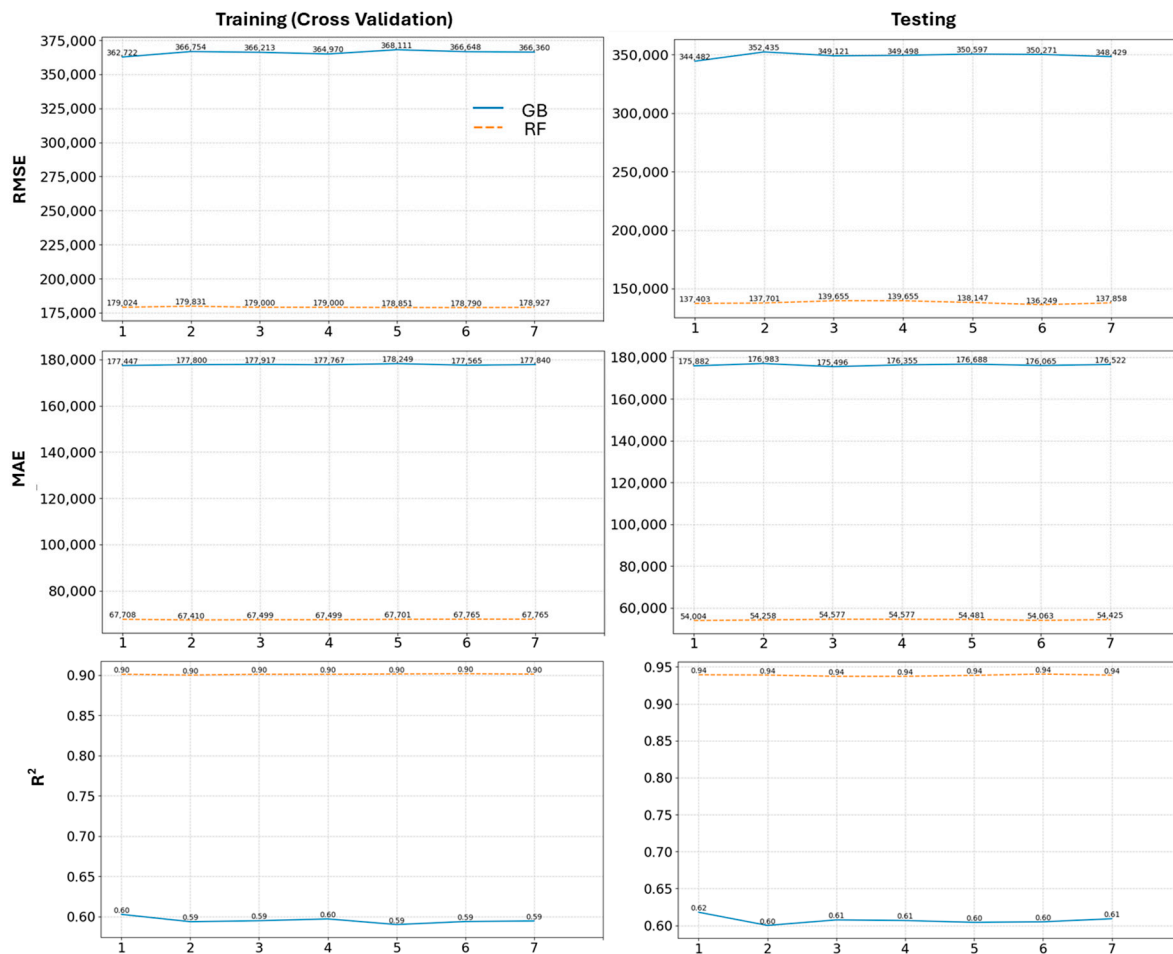


Figure 8. Model Performance Metrics for TCs. RMSE—Root mean square error is the square root of the average of the squared differences between predictions and actual observations; MAE—Mean absolute error is the average absolute difference between predicted and actual values; R^2 —Coefficient of determination is the proportion of variance in the dependent variable explained by the model.

3.2. Feature Interpretation

Early time series data proved critical for both PLOSM and TC prediction, with PLOSM models showing maximum feature selection within the first three days, and TC models extending feature relevance to the initial five days (Figure 9). The illness severity and mortality with higher risk scores were the most important features to predict the PLOSM with RF, while the prediction of PLOSM, especially for minority levels of PLOS only and both PLOS and mortality, was essential to predict the TCs for day 1 through day 7 (Figure 9). Other demographic and diagnosis factors, such as age, race, financial status, insurance choices, elective surgeries, major operating room usage, and diagnosis codes, were ranked into the top 10 important predictors for PLOSM and TC prediction (Figure 9). For PLOSM prediction, procedures of Introduction of Other Antineoplastic into Spinal Canal, Percutaneous Approach (3E0R305) and Introduction of Other Antineoplastic into Central Vein, Percutaneous Approach (3E04305) were more important than other procedure predictors, especially on early days at 1, 2, and 3 (Figure 9). However, no features were important to predict the TCs.



Figure 9. Selected Features Summary for PLOSM and TC Prediction. (a) Numbering of features selected for predicting PLOSM for each day. (b) Number of features selected for predicting TCs for each day. (c) Selected static features with importance values were shown for each day's PLOSM prediction. (d) Selected procedure features with importance values were shown for each day's PLOSM prediction. (e) Selected static features with importance values were shown for each day's TC prediction.

4. Discussion

Despite ALL treatment advancements, challenges persist in early diagnosis, treatment complexity, and financial burden [5,6,10,18]. Lack of reliable prediction models impedes timely decisions and resource allocation, impacting patient care. Current models often fail to integrate real-time data or combine LOS and mortality predictions, limiting personalized care [38]. This study addressed these limitations by using HCUP-NIS real-time data to develop dynamic prediction models for concurrent outcomes (PLOS, mortality, and TCs), aiming to improve clinical decisions, resource optimization, and patient safety in ALL.

4.1. Significance of Early Prediction in ALL Treatment

The study findings emphasized the importance of early prediction, particularly within the first three days of hospitalization, in forecasting patient outcomes (Figure 9). Different predictors were found to be significant for forecasting concurrent patient outcomes of PLOS and mortality at each stage, providing valuable insights into the evolving nature of risk factors over time (Figure 9). Notably, the RF model outperformed other methods in predicting concurrent patient outcomes (PLOSM) compared to individual outcome predictions. While RF achieved similar accuracy and precision values, it demonstrated higher AUC, F1, and recall scores (Figures 7–9). These findings indicate that the concurrent RF model is superior in predicting each patient's outcome class, particularly for minority cases, where patient outcomes are highly imbalanced (e.g., a greater proportion of normally discharged patients compared to those with prolonged stays or in-hospital mortality). However, the greater performance drop observed in the four-class PLOSM outcome compared to binary PLOS and mortality prediction, in Appendix C (Table A3), likely reflected increased classification complexity, class imbalance, and overlapping feature distributions among outcome categories. The multiclass task requires distinguishing between combinations of LOS and mortality states, which reduced effective sample size per class and increased boundary complexity, thereby increasing generalization error.

Since inpatient adverse events are less frequent than normal discharges and most prediction models focus on separately predicting LOS and mortality at admission, the ability to anticipate prolonged hospitalization and in-hospital mortality concurrently enables healthcare providers to personalize treatment plans, enhance patient safety, and allocate resources more efficiently [17,28–40]. These insights are particularly critical for patients with ALL, whose immunocompromised status makes them highly susceptible to hospital-related complications [5,6].

The continuous prediction methodology, which utilizes predicted PLOS and mortality as predictors for TC prediction, enhances the timely forecasting of secondary outcomes when critical features are unavailable at the time of early prediction (Figure 9). Traditional machine learning models for TC prediction rely on post-discharge LOS as a predictor, a variable not available in early-stage assessments [42,44]. This study demonstrates that integrating machine learning pipeline techniques allows for the early prediction of TCs, thereby offering a proactive approach to managing patient care.

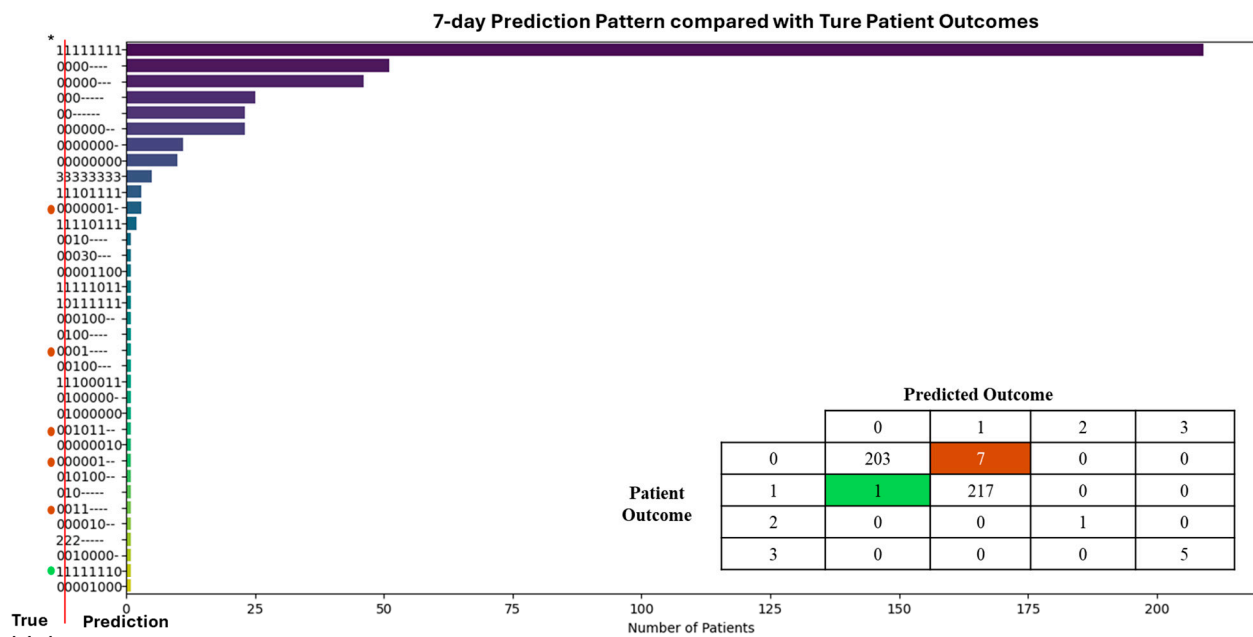
4.2. Clinical and Financial Implications

Multiple factors, including disease stage, APRDRG severity risk, APRDRG mortality risk, age, race, geographic location, household income, insurance type, hospital transfer status, medical service line, diagnosis, and treatment, played a significant role in predicting PLOS and mortality, aligning with the important factors from other prediction models [21,34]. Additionally, patient comorbidities (e.g., hypertension and depression), hospital resources and location, clinical trial participation, and predicted PLOSM were all key predictors for TCs (Figure 9). Notably, patients transferred from acute care hospitals significantly influenced PLOSM prediction, aligning with findings that pediatric patients registered in clinical trials exhibited similar risk factors [12]. These findings underscore the importance of identifying key predictors in inpatient care delivery systems to facilitate early interventions tailored to individual disease conditions, diagnoses, and hospital resources. Beyond clinical improvements, this study highlights the economic burden of ALL treatment. Early prediction of TCs provides valuable insights into financial toxicity, enabling timely interventions to mitigate economic strain on patients and families. Insurance type and patient income levels were found to influence both PLOSM and TC predictions (Figure 9).

Data from the Agency for Healthcare Research and Quality (2006–2008) indicated that patients with private insurance were associated with lower risk-adjusted mortality rates than those with other insurance types in U.S. hospitals [65]. Furthermore, patients with private insurance or higher income levels generally had faster access to healthcare services than those with lower income or alternative insurance coverage [66]. Clinical trial enrollment was another crucial predictor for TCs, with greater participation observed among patients with private insurance [67]. These insights reinforce the necessity of financial navigation programs and policy adjustments to enhance care quality and accessibility, aligning with the CRAACO framework [24].

The implementation of dynamic machine learning models enables real-time updates and continuous patient risk assessment, thereby enhancing clinical decision-making [45,47]. Despite the complexity of ALL treatment, modern first-line therapy of standard- and high-dose chemotherapy, hematopoietic stem-cell transplantation, and new targeted therapy improved the long-term outcome for patients with ALL [68]. Given the ever-evolving nature of ALL treatment, continuous updates to patient outcome prediction models are essential. By integrating the latest treatment strategies into prediction models, hospitals can account for new developments and refine patient risk assessments accordingly [45,47]. Daily patient outcome predictions offer a nuanced understanding of predictor effects, facilitating personalized treatment planning and more efficient hospital resource allocation [45,47].

To illustrate how the prediction model performed across days for each individual patient, the prediction results for a subset of records of patients with ALL enrolled in clinical trial research ($n = 434$) were summarized, as shown in Figure 10. Since the dataset was pivoted by the procedure received day (Figure 9), the subset was expanded to 774 records (22% from the testing set and 78% from the training set). The model performed well for most records, especially for two (deceased with short hospital stay) and three (prolonged hospital stay and deceased) patient outcomes. Some inaccurate predictions shown in the earlier days' predictions were able to be corrected in later days' predictions, such as the true patient outcome label was 1, and the day 1 through day 7 prediction pattern was 0111111. However, there were some patients who ended up with prolonged hospital stay prediction results when they were discharged normally from the hospital (e.g., 0001, 0011, 001011, 000001, and 0000001). This patient cohort presents a valuable opportunity for care teams to reassess disease progression and facilitate close monitoring of patients' health by family members or caregivers post-discharge. A ten-year-old patient, with a less severe risk and mortality score at admission, received an antineoplastic drug through the central vein on day 1 before achieving remission for ALL. The discrepancy between his hospital record and prediction results on day 7 (11111110) may show that there is a possibility of discharging the patient on day 7, since he only had an eight-day hospitalization. The retrospective nature of HCUP data, coupled with its collection methodology, may result in incomplete patient records, specifically regarding diagnoses and procedures. Critically, the absence of vital signs and other important clinical indicators limits the depth and accuracy of this prediction model. Therefore, this dynamic patient outcome prediction model is suggested to be integrated into hospital electronic health records (EHR) to make daily predictions on patient risk of adverse events (e.g., PLOS and mortality) and estimation of TCs to streamline patient monitoring, facilitate timely interventions, and reallocate the resources for each patient.



*The pattern formation: True Patient Outcome, Day 1 prediction, Day 2 prediction, Day 3 prediction, Day 4 prediction, Day 5 prediction, Day 6 prediction, Day 7 prediction;
 - represents no prediction of the day (The patient was either discharged alive or deceased in the hospital)
 0 represents normally discharged patients (no PLOS and not deceased)
 1 represents prolonged hospital stay (> 7 days) but not deceased
 2 represents deceased with short hospital stay (<= 7 days)
 3 represents prolonged hospital stay and deceased

Figure 10. Summary of Patient Outcome Prediction Patterns for Patients with ALL Enrolled in Clinical Trial Research. Note: The color of the dot is the pattern calculated in the colored cell of the table on the bottom right corner. The red line in the pattern formation was to separate the true label and the prediction label.

5. Limitations and Future Studies

This retrospective analysis of HCUP data is subject to limitations inherent in administrative databases, including potential inaccuracies in procedure coding and the absence of granular clinical data, such as vital signs, which are essential for understanding dynamic patient treatment. Local EHR integration is needed for hospital adoption, especially in rural settings. Generalizability is limited due to U.S. data. Imputation and data imbalance may affect accuracy. Future work should include external validation to re-evaluate the inaccurate prediction for certain patient cohorts and explore deep learning for improved predictions. Real-time EHR integration is crucial for clinical decision support. Potential covariate effects require combining prediction with causal analysis to ensure effective improvements in patient safety and financial outcomes [69].

6. Conclusions

This study provided an effective method to predict multiple patient outcomes in a dynamic and continuous way, combined with machine learning models. By anticipating potential outcomes, healthcare professionals can personalize treatment plans based on individual risk factors, improving the efficacy of care for each patient. Moreover, hospitals optimize resource allocation within healthcare facilities, ensuring that high-risk patients have access to necessary intensive care, specialized nursing, and other resources when they need them most. Timely predictions foster efficient communication among care teams, leading to better-informed decisions and improved patient care. Therefore, developing predictive models for patient outcomes like length of stay, mortality, and TCs is vital to

support informed decision-making, optimize clinical workflows, and ultimately improve patient safety and well-being in ALL treatments.

Author Contributions: E.J., B.M.W., F.D., H.S. and B.M. Conceptualization, J.M., E.J., B.M.W. and B.M.; methodology, J.M. and B.M.W.; software, J.M.; validation, J.M., E.J., B.M.W., F.D., H.S. and B.M.; formal analysis, J.M., E.J. and B.M.W.; investigation, J.M.; resources, E.J.; data curation, E.J.; writing—original draft preparation, J.M.; writing—review and editing, J.M., E.J., B.M.W., F.D., H.S. and B.M.; visualization, J.M.; supervision, E.J., B.M.W. and B.M.; project administration, B.M.; funding acquisition, E.J., B.M.W. and B.M. All authors have read and agreed to the published version of the manuscript.

Funding: Funding for this study was made possible with internal fund, library and the Office of Sponsored Programs at Montana State University.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable. The data was purchased from the HCUP database. All patient records were de-identified by HCUP.

Data Availability Statement: The data was purchased through the HEALTHCARE COST & UTILIZATION PROJECT (HCUP). The link to purchase is <https://hcup-us.ahrq.gov/db/nation/nis/nisdde.jsp> (accessed on 1 March 2025).

Conflicts of Interest: No conflicts of interest declared by any of the authors.

Appendix A

Table A1. ICD-10 Codes Description and Frequency.

ICD Codes	Description	Frequency	Relative Frequency
ICD-10-DX codes			
A00–A09	Intestinal infectious diseases	912	4.48%
A30–A49	Other bacterial diseases	2075	10.19%
B00–B09	Viral infections characterized by skin and mucous membrane lesions	451	2.21%
B25–B34	Other viral diseases	883	4.33%
B35–B49	Mycoses	1254	6.16%
B95–B97	Bacterial and viral infectious agents	2155	10.58%
C76–C80	Malignant neoplasms of ill-defined, other secondary and unspecified sites	223	1.09%
C81–C96	Malignant neoplasms of lymphoid, hematopoietic and related tissue	1049	5.15%
D60–D64	Aplastic and other anemias and other bone marrow failure syndromes	11,479	56.35%
D65–D69	Coagulation defects, purpura and other hemorrhagic conditions	4198	20.61%
D70–D77	Other disorders of blood and blood-forming organs	5261	25.83%
D80–D89	Certain disorders involving the immune mechanism	4596	22.56%
E40–E46	Malnutrition	2116	10.39%
E50–E64	Other nutritional deficiencies	872	4.28%
E70–E88	Metabolic disorders	8792	43.16%

Table A1. Cont.

ICD Codes	Description	Frequency	Relative Frequency
G60–G65	Polyneuropathies and other disorders of the peripheral nervous system	2415	11.86%
G89–G99	Other disorders of the nervous system	2630	12.91%
I10–I1A	Hypertensive diseases	5577	27.38%
I20–I25	Ischemic heart diseases	989	4.85%
I26–I28	Pulmonary heart disease and diseases of pulmonary circulation	385	1.89%
I30–I5A	Other forms of heart disease	2440	11.98%
I60–I69	Cerebrovascular diseases	510	2.50%
I70–I79	Diseases of arteries, arterioles and capillaries	246	1.21%
I80–I89	Diseases of veins, lymphatic vessels and lymph nodes, not elsewhere classified	1026	5.04%
I95–I99	Other and unspecified disorders of the circulatory system	951	4.67%
M50–M54	Other dorsopathies	888	4.36%
M86–M90	Other osteopathies	408	2.00%
N10–N16	Renal tubulo-interstitial diseases	287	1.41%
N17–N19	Acute kidney failure and chronic kidney disease	2580	12.67%
N25–N29	Other disorders of kidney and ureter	276	1.35%
N30–N39	Other diseases of the urinary system	887	4.35%
R00–R09	Symptoms and signs involving the circulatory and respiratory systems	3020	14.82%
R10–R19	Symptoms and signs involving the digestive system and abdomen	4485	22.02%
R20–R23	Symptoms and signs involving the skin and subcutaneous tissue	805	3.95%
R25–R29	Symptoms and signs involving the nervous and musculoskeletal systems	451	2.21%
R30–R39	Symptoms and signs involving the genitourinary system	734	3.60%
R40–R46	Symptoms and signs involving cognition, perception, emotional state and behavior	650	3.19%
R50–R69	General symptoms and signs	8310	40.79%
R70–R79	Abnormal findings on examination of blood, without diagnosis	3176	15.59%
R90–R94	Abnormal findings on diagnostic imaging and in function studies, without diagnosis	483	2.37%
T36–T50	Poisoning by, adverse effect of and underdosing of drugs, medicaments and biological substances	9922	48.71%
T80–T88	Complications of surgical and medical care, not elsewhere classified	1928	9.46%
U00–U49	Provisional assignment of new diseases of uncertain etiology or emergency use	369	1.81%

Table A1. Cont.

ICD Codes	Description	Frequency	Relative Frequency
Y83–Y84	Surgical and other medical procedures as the cause of abnormal reaction of the patient, or of later complication, without mention of misadventure at the time of the procedure	1204	5.91%
Y90–Y99	Supplementary factors related to causes of morbidity classified elsewhere	2249	11.04%
Z16–Z16	Resistance to antimicrobial drugs	251	1.23%
Z20–Z29	Persons with potential health hazards related to communicable diseases	5186	25.46%
Z40–Z53	Encounters for other specific health care	8233	40.42%
Z66	Do not resuscitate status	713	3.50%
Z69–Z76	Persons encountering health services in other circumstances	569	2.79%
Z77–Z99	Persons with potential health hazards related to family and personal history and certain conditions influencing health status	12,895	63.30%
ICD-10-PR codes			
30233N1	Transfusion of Nonautologous Red Blood Cells into Peripheral Vein, Percutaneous Approach	2460	12.08%
30233R1	Transfusion of Nonautologous Platelets into Peripheral Vein, Percutaneous Approach	1669	8.19%
30243N1	Transfusion of Nonautologous Red Blood Cells into Central Vein, Percutaneous Approach	914	4.49%
30243R1	Transfusion of Nonautologous Platelets into Central Vein, Percutaneous Approach	559	2.74%
3E03305	Introduction of Other Antineoplastic into Peripheral Vein, Percutaneous Approach	1385	6.80%
3E04305	Introduction of Other Antineoplastic into Central Vein, Percutaneous Approach	4256	20.89%
3E0430M	Introduction of Antineoplastic, Monoclonal Antibody, into Central Vein, Percutaneous Approach	229	1.12%
3E0436Z	Introduction of Nutritional Substance into Central Vein, Percutaneous Approach	319	1.57%
3E0G76Z	Introduction of Nutritional Substance into Upper GI, Via Natural or Artificial Opening	233	1.14%
3E0R305	Introduction of Other Antineoplastic into Spinal Canal, Percutaneous Approach	5688	27.92%
XW04351	Introduction of Blinatumomab Antineoplastic Immunotherapy into Central Vein, Percutaneous Approach, New Technology Group 1	459	2.25%
02H633Z	Insertion of Infusion Device into Right Atrium, Percutaneous Approach	418	2.05%
02HV33Z	Insertion of Infusion Device into Superior Vena Cava, Percutaneous Approach	3018	14.82%
02PYX3Z	Removal of Infusion Device from Great Vessel, External Approach	198	0.97%

Table A1. Cont.

ICD Codes	Description	Frequency	Relative Frequency
03HY32Z	Insertion of Monitoring Device into Upper Artery, Percutaneous Approach	205	1.01%
0JH60WZ	Insertion of Totally Implantable Vascular Access Device into Chest Subcutaneous Tissue and Fascia, Open Approach	657	3.23%
0JH63XZ	Insertion of Tunneled Vascular Access Device into Chest Subcutaneous Tissue and Fascia, Percutaneous Approach	357	1.75%
009U3ZX	Drainage of Spinal Canal, Percutaneous Approach, Diagnostic	2424	11.90%
009U3ZZ	Drainage of Spinal Canal, Percutaneous Approach	266	1.31%
5A09357	Assistance with Respiratory Ventilation, Less than 24 Consecutive Hours, Continuous Positive Airway Pressure	179	0.88%
5A1955Z	Respiratory Ventilation, Greater than 96 Consecutive Hours	268	1.32%
B01B1ZZ	Fluoroscopy of Spinal Cord using Low Osmolar Contrast	208	1.02%
B5181ZA	Fluoroscopy of Superior Vena Cava using Low Osmolar Contrast, Guidance	210	1.03%
B518ZZA	Fluoroscopy of Superior Vena Cava, Guidance	230	1.13%
B548ZZA	Ultrasonography of Superior Vena Cava, Guidance	441	2.16%
079T3ZX	Drainage of Bone Marrow, Percutaneous Approach, Diagnostic	286	1.40%
07DR3ZX	Extraction of Iliac Bone Marrow, Percutaneous Approach, Diagnostic	3187	15.64%
8E0ZXY6	Isolation	234	1.15%
0BH17EZ	Insertion of Endotracheal Airway into Trachea, Via Natural or Artificial Opening	360	1.77%

Appendix B

Table A2. Individual Prediction of PLOS and Mortality Evaluation.

Outcome	Model	Day	Training (Cross Validation)				Testing			
			Accuracy	F1-Score	Precision	Recall	Accuracy	F1-Score	Precision	Recall
Mortality	RF	1	0.9913	0.9911	0.9940	0.9886	0.9807	0.9773	0.9774	0.9807
Mortality	RF	2	0.9908	0.9906	0.9934	0.9881	0.9845	0.9830	0.9830	0.9845
Mortality	RF	3	0.9912	0.9910	0.9948	0.9876	0.9806	0.9775	0.9776	0.9806
Mortality	RF	4	0.9922	0.9920	0.9946	0.9898	0.9817	0.9784	0.9798	0.9817
Mortality	RF	5	0.9921	0.9919	0.9948	0.9894	0.9814	0.9795	0.9792	0.9814
Mortality	RF	6	0.9923	0.9922	0.9951	0.9895	0.9832	0.9810	0.9821	0.9832
Mortality	RF	7	0.9924	0.9922	0.9949	0.9898	0.9835	0.9808	0.9832	0.9835
Mortality	RF	>7	0.9946	0.9945	0.9958	0.9933	0.9869	0.9861	0.9866	0.9869
PLOS	RF	1	0.8672	0.8621	0.8735	0.8596	0.8324	0.8305	0.8298	0.8324
PLOS	RF	2	0.8674	0.8636	0.8743	0.8600	0.8345	0.8338	0.8335	0.8345
PLOS	RF	3	0.8641	0.8622	0.8645	0.8641	0.8581	0.8580	0.8580	0.8581
PLOS	RF	4	0.8632	0.8622	0.8615	0.8649	0.8606	0.8607	0.8608	0.8606
PLOS	RF	5	0.8675	0.8676	0.8648	0.8720	0.8558	0.8562	0.8572	0.8558
PLOS	RF	6	0.8662	0.8668	0.8623	0.8716	0.8626	0.8630	0.8644	0.8626
PLOS	RF	7	0.8630	0.8634	0.8604	0.8666	0.8750	0.8752	0.8758	0.8750
PLOS	RF	>7	0.8986	0.9011	0.8866	0.9180	0.9051	0.9048	0.9052	0.9051

Table A2. Cont.

Outcome	Model	Day	Training (Cross Validation)				Testing			
			Accuracy	F1-Score	Precision	Recall	Accuracy	F1-Score	Precision	Recall
Mortality	GB	1	0.9836	0.9836	0.9770	0.9904	0.9676	0.9694	0.9715	0.9676
Mortality	GB	2	0.9821	0.9821	0.9740	0.9907	0.9682	0.9718	0.9774	0.9682
Mortality	GB	3	0.9793	0.9795	0.9696	0.9899	0.9598	0.9663	0.9764	0.9598
Mortality	GB	4	0.9780	0.9782	0.9674	0.9894	0.9578	0.9638	0.9726	0.9578
Mortality	GB	5	0.9777	0.9779	0.9668	0.9894	0.9543	0.9612	0.9716	0.9543
Mortality	GB	6	0.9760	0.9761	0.9666	0.9861	0.9526	0.9596	0.9702	0.9526
Mortality	GB	7	0.9715	0.9716	0.9625	0.9813	0.9485	0.9572	0.9706	0.9485
Mortality	GB	>7	0.9579	0.9585	0.9421	0.9758	0.9305	0.9416	0.9606	0.9305
PLOS	GB	1	0.8331	0.8264	0.8372	0.8258	0.8063	0.8065	0.8068	0.8063
PLOS	GB	2	0.8200	0.8158	0.8214	0.8171	0.7894	0.7907	0.7928	0.7894
PLOS	GB	3	0.8098	0.8074	0.8077	0.8124	0.7840	0.7852	0.7876	0.7840
PLOS	GB	4	0.8067	0.8057	0.8027	0.8114	0.7851	0.7862	0.7884	0.7851
PLOS	GB	5	0.8021	0.8027	0.7972	0.8106	0.7847	0.7858	0.7889	0.7847
PLOS	GB	6	0.7956	0.7980	0.7875	0.8091	0.7865	0.7877	0.7927	0.7865
PLOS	GB	7	0.7816	0.7817	0.7811	0.7824	0.7882	0.7890	0.7923	0.7882
PLOS	GB	>7	0.8280	0.8321	0.8203	0.8467	0.8161	0.8158	0.8157	0.8161

Appendix C

Table A3. PLOSM Prediction and Combined Outcome Prediction Evaluation.

Outcome	Model	Day	Training (Cross Validation)				Testing			
			Accuracy	F1-Score	Precision	Recall	Accuracy	F1-Score	Precision	Recall
PLOSM	RF	1	0.9292	0.9287	0.9323	0.9292	0.8175	0.6689	0.7893	0.6088
PLOSM	RF	2	0.9305	0.9302	0.9329	0.9305	0.8194	0.7207	0.8381	0.6659
PLOSM	RF	3	0.9284	0.9283	0.9302	0.9284	0.8370	0.7260	0.8354	0.6668
PLOSM	RF	4	0.9275	0.9275	0.9285	0.9275	0.8303	0.6808	0.8036	0.6215
PLOSM	RF	5	0.9292	0.9291	0.9301	0.9292	0.8424	0.7235	0.8471	0.6671
PLOSM	RF	6	0.9285	0.9284	0.9288	0.9285	0.8387	0.7483	0.8725	0.6912
PLOSM	RF	7	0.9293	0.9292	0.9293	0.9293	0.8416	0.6585	0.8888	0.6042
PLOSM	RF	>7	0.9459	0.9459	0.9468	0.9459	0.8916	0.7887	0.9318	0.7469
PLOS + Mortality *	RF	1	-	-	-	-	0.8175	0.5971	0.7882	0.5425
PLOS + Mortality *	RF	2	-	-	-	-	0.8245	0.6944	0.7919	0.6430
PLOS + Mortality *	RF	3	-	-	-	-	0.8413	0.6535	0.7907	0.5939
PLOS + Mortality *	RF	4	-	-	-	-	0.8478	0.6651	0.8383	0.6014
PLOS + Mortality *	RF	5	-	-	-	-	0.8409	0.6947	0.8234	0.6411
PLOS + Mortality *	RF	6	-	-	-	-	0.8502	0.7278	0.8604	0.6633
PLOS + Mortality *	RF	7	-	-	-	-	0.8622	0.6728	0.8897	0.6155
PLOS + Mortality *	RF	>7	-	-	-	-	0.8947	0.7446	0.9290	0.7011
PLOSM	GB	1	0.8881	0.8858	0.8881	0.8892	0.7371	0.5176	0.5168	0.5576
PLOSM	GB	2	0.8697	0.8677	0.8697	0.8702	0.7514	0.5781	0.5730	0.6210
PLOSM	GB	3	0.8638	0.8618	0.8638	0.8631	0.7409	0.5297	0.5012	0.6258
PLOSM	GB	4	0.8548	0.8531	0.8548	0.8538	0.7521	0.5770	0.5415	0.6574
PLOSM	GB	5	0.8572	0.8552	0.8572	0.8556	0.7421	0.5358	0.5040	0.6249
PLOSM	GB	6	0.8544	0.8528	0.8544	0.8523	0.7394	0.5643	0.5522	0.6258
PLOSM	GB	7	0.8499	0.8482	0.8499	0.8474	0.7225	0.5714	0.5571	0.6542
PLOSM	GB	>7	0.8593	0.8580	0.8593	0.8586	0.7529	0.5478	0.5464	0.6095
PLOS + Mortality *	GB	1	-	-	-	-	0.7816	0.4911	0.5228	0.5145
PLOS + Mortality *	GB	2	-	-	-	-	0.7665	0.5258	0.5633	0.5799
PLOS + Mortality *	GB	3	-	-	-	-	0.7530	0.5360	0.5273	0.5946
PLOS + Mortality *	GB	4	-	-	-	-	0.7530	0.4789	0.4666	0.5408

Table A3. Cont.

Outcome	Model	Day	Training (Cross Validation)				Testing			
			Accuracy	F1-Score	Precision	Recall	Accuracy	F1-Score	Precision	Recall
PLOS + Mortality *	GB	5	-	-	-	-	0.7512	0.5029	0.5369	0.5590
PLOS + Mortality *	GB	6	-	-	-	-	0.7495	0.4994	0.5043	0.5514
PLOS + Mortality *	GB	7	-	-	-	-	0.7472	0.4972	0.5004	0.5687
PLOS + Mortality *	GB	>7	-	-	-	-	0.7529	0.5193	0.5316	0.5862

* This was the outcome of combining the individual prediction of PLOS and Mortality, which was compared with the PLOSD. The Training metrics can be referred to the Mortality's and PLOS's metrics.

References

- Meenaghan, T.; Dowling, M.; Kelly, M. Acute leukaemia: Making sense of a complex blood cancer. *Br. J. Nurs.* **2012**, *21*, 78–83. [\[CrossRef\]](#)
- Blumenreich, M.S. The White Blood Cell and Differential Count. In *Clinical Methods: The History, Physical, and Laboratory Examinations*; Walker, H.K., Hall, W.D., Hurst, J.W., Eds.; Butterworth Publishers: Boston, MA, USA, 1990.
- Malard, F.; Mohty, M. Acute lymphoblastic leukaemia. *Lancet* **2020**, *395*, 1146–1162. [\[CrossRef\]](#) [\[PubMed\]](#)
- Black, G.B.; Boswell, L.; Harris, J.; Whitaker, K.L. What causes delays in diagnosing blood cancers? A rapid review of the evidence. *Prim. Health Care Res. Dev.* **2023**, *24*, e26. [\[CrossRef\]](#)
- Kansagra, A.; Dahiya, S.; Litzow, M. Continuing challenges and current issues in acute lymphoblastic leukemia. *Leuk. Lymphoma* **2018**, *59*, 526–541. [\[CrossRef\]](#)
- Lonetti, A.; Iacobucci, I.; Masetti, R. Successes and Challenges for Diagnosis and Therapy of Acute Leukemia. *J. Oncol.* **2019**, *2019*, 3408318. [\[CrossRef\]](#)
- Fardin, M.F.; Akter, M. Review on blood cancer and their types. *J. Mol. Pharm. Regul. Aff.* **2021**, *3*, 18–26.
- Paul, S.; Kantarjian, H.; Jabbour, E.J. Adult Acute Lymphoblastic Leukemia. *Mayo Clin. Proc.* **2016**, *91*, 1645–1666. [\[CrossRef\]](#)
- Douvas, M.G.; Riegler, L.L. Meeting Challenges in the Long-Term Care of Children, Adolescents, and Young Adults with Acute Lymphoblastic Leukemia. *Curr. Hematol. Malign. Rep.* **2022**, *17*, 15–24. [\[CrossRef\]](#) [\[PubMed\]](#)
- Abrams, H.R.; Durbin, S.; Huang, C.X.; Johnson, S.F.; Nayak, R.K.; Zahner, G.J.; Peppercorn, J. Financial toxicity in cancer care: Origins, impact, and solutions. *Transl. Behav. Med.* **2021**, *11*, 2043–2054. [\[CrossRef\]](#)
- Lu, B.R. *Clinical Treatment Act*; Congress, T., Ed.; House—Energy and Commerce: Washington, DC, USA, 2019.
- Ma, J.; Johnson, E.A.; McCrory, B. Predicting risk factors for pediatric mortality in clinical trial research: A retrospective, cross-sectional study using a Healthcare Cost and Utilization Project database. *J. Clin. Transl. Sci.* **2023**, *7*, e211. [\[CrossRef\]](#)
- Newgard, C.D.; Fleischman, R.; Choo, E.; John Ma, O.; Hedges, J.R.; John McConnell, K. Validation of length of hospital stay as a surrogate measure for injury severity and resource use among injury survivors. *Acad. Emerg. Med.* **2010**, *17*, 142–150. [\[CrossRef\]](#)
- English, M.; Mwaniki, P.; Julius, T.; Chepkirui, M.; Gathara, D.; Ouma, P.O.; Cherutich, P.; Okiro, A.E.; Snow, R.W. Hospital Mortality—A neglected but rich source of information supporting the transition to higher quality health systems in low and middle income countries. *BMC Med.* **2018**, *16*, 32. [\[CrossRef\]](#)
- Nguyen, Q.; Wybrow, M.; Burstein, F.; Taylor, D.; Enticott, J. Understanding the impacts of health information systems on patient flow management: A systematic review across several decades of research. *PLoS ONE* **2022**, *17*, e0274493. [\[CrossRef\]](#)
- Huntley, D.A.; Cho, D.W.; Christman, J.; Csernansky, J.G. Predicting length of stay in an acute psychiatric hospital. *Psychiatr. Serv.* **1998**, *49*, 1049–1053. [\[CrossRef\]](#) [\[PubMed\]](#)
- Zhao, P.; Liu, C.; Zhang, C.; Hou, Y.; Zhang, X.; Zhao, J.; Sun, G.; Zhou, J. Using Machine Learning to Predict the In-Hospital Mortality in Women with ST-Segment Elevation Myocardial Infarction. *Rev. Cardiovasc. Med.* **2023**, *24*, 126. [\[CrossRef\]](#)
- Nghiem, S.; Afoakwah, C.; Scuffham, P.; Byrnes, J. Benchmarking hospital safety and identifying determinants of hospital-acquired complication: The case of Queensland cardiac linkage longitudinal cohort. *Infect. Prev. Pract.* **2022**, *4*, 100198. [\[CrossRef\]](#)
- Rosman, M.; Rachminov, O.; Segal, O.; Segal, G. Prolonged patients' In-Hospital Waiting Period after discharge eligibility is associated with increased risk of infection, morbidity and mortality: A retrospective cohort analysis. *BMC Health Serv Res* **2015**, *15*, 246. [\[CrossRef\]](#)
- Doctoroff, L.; Herzig, S.J. Predicting Patients at Risk for Prolonged Hospital Stays. *Med. Care* **2020**, *58*, 778–784. [\[CrossRef\]](#) [\[PubMed\]](#)
- Novitski, P.; Cohen, C.M.; Karasik, A.; Shalev, V.; Hodik, G.; Moskovitch, R. All-Cause Mortality Prediction in T2D Patients. In *Proceedings of the 18th International Conference on Artificial Intelligence in Medicine (AIME)*; Electr Network; University of Minnesota: Minneapolis, MN, USA, 2020.

22. Parshuram, C.S.; Dryden-Palmer, K.; Farrell, C.; Gottesman, R.; Gray, M.; Hutchison, J.S.; Helfaer, M.; Hunt, E.A.; Joffe, A.R.; Lacroix, J.; et al. Effect of a Pediatric Early Warning System on All-Cause Mortality in Hospitalized Pediatric Patients: The EPOCH Randomized Clinical Trial. *JAMA* **2018**, *319*, 1002–1012. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Hamavid, H.; Birger, M.; Bulchis, A.G.; Lomsadze, L.; Joseph, J.; Baral, R.; Bui, A.L.; Horst, C.; Johnson, E.; Dieleman, J.L. Assessing the Complex and Evolving Relationship between Charges and Payments in US Hospitals: 1996–2012. *PLoS ONE* **2016**, *11*, e0157912. [\[CrossRef\]](#)
24. Van de Beek, H. Clinical Research as a Care Option: Optimizing Approaches. *Appl. Clin. Trials* **2019**, *28*, 32–35.
25. Kashef, A.; Khatibi, T.; Mehrvar, A. Treatment outcome classification of pediatric Acute Lymphoblastic Leukemia patients with clinical and medical data using machine learning: A case study at MAHAK hospital. *Inform. Med. Unlocked* **2020**, *20*, 100399. [\[CrossRef\]](#)
26. Thakkar, S.G.; Fu, A.Z.; Sweetenham, J.W.; Mciver, Z.A.; Mohan, S.R.; Ramsingh, G.; Advani, A.S.; Sobecks, R.; Rybicki, L.; Kalaycio, M.; et al. Survival and predictors of outcome in patients with acute leukemia admitted to the intensive care unit. *Cancer* **2008**, *112*, 2233–2240. [\[CrossRef\]](#)
27. Cleaver, A.L.; Beesley, A.H.; Firth, M.J.; Sturges, N.C.; O’Leary, R.A.; Hunger, S.P.; Baker, D.L.; Kees, U.R. Gene-based outcome prediction in multiple cohorts of pediatric T-cell acute lymphoblastic leukemia: A Children’s Oncology Group study. *Mol. Cancer* **2010**, *9*, 105. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Stone, K.; Zwiggelaar, R.; Jones, P.; Mac Parthaláin, N. A systematic review of the prediction of hospital length of stay: Towards a unified framework. *PLoS Digit. Health* **2022**, *1*, e0000017. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Gokhale, S.; Taylor, D.; Gill, J.; Hu, Y.; Zeps, N.; Lequertier, V.; Prado, L.; Teede, H.; Enticott, J. Hospital length of stay prediction tools for all hospital admissions and general medicine populations: Systematic review and meta-analysis. *Front. Med.* **2023**, *10*, 1192969. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Medeiros, N.B.; Fogliatto, F.S.; Rocha, M.K.; Tortorella, G.L. Forecasting the length-of-stay of pediatric patients in hospitals: A scoping review. *BMC Health Serv. Res.* **2021**, *21*, 938. [\[CrossRef\]](#)
31. Lequertier, V.; Wang, T.; Fondrevelle, J.; Augusto, V.; Duclos, A. Hospital Length of Stay Prediction Methods: A Systematic Review. *Med. Care* **2021**, *59*, 929–938. [\[CrossRef\]](#)
32. Farimani, R.M.; Karim, H.; Atashi, A.; Tohidinezhad, F.; Bahaadini, K.; Abu-Hanna, A.; Eslami, S. Models to predict length of stay in the emergency department: A systematic literature review and appraisal. *BMC Emerg. Med.* **2024**, *24*, 54. [\[CrossRef\]](#)
33. Gokhale, S.; Taylor, D.; Gill, J.; Hu, Y.; Zeps, N.; Lequertier, V.; Teede, H.; Enticott, J. Hospital length of stay prediction for general surgery and total knee arthroplasty admissions: Systematic review and meta-analysis of published prediction models. *Digit. Health* **2023**, *9*, 20552076231177497. [\[CrossRef\]](#)
34. Bacchi, S.; Tan, Y.; Oakden-Rayner, L.; Jannes, J.; Kleinig, T.; Koblar, S. Machine learning in the prediction of medical inpatient length of stay. *Intern. Med. J.* **2022**, *52*, 176–185. [\[CrossRef\]](#)
35. Peres, I.T.; Hamacher, S.; Oliveira, F.L.C.; Thomé, A.M.T.; Bozza, F.A. What factors predict length of stay in the intensive care unit? Systematic review and meta-analysis. *J. Crit. Care* **2020**, *60*, 183–194. [\[CrossRef\]](#)
36. Naemi, A.; Schmidt, T.; Mansourvar, M.; Naghavi-Behzad, M.; Ebrahimi, A.; Wiil, U.K. Machine learning techniques for mortality prediction in emergency departments: A systematic review. *BMJ Open* **2021**, *11*, e052663. [\[CrossRef\]](#)
37. Mpanya, D.; Celik, T.; Klug, E.; Ntsinjana, H. Predicting mortality and hospitalization in heart failure using machine learning: A systematic literature review. *IJC Heart Vasc.* **2021**, *34*, 100773. [\[CrossRef\]](#)
38. Clark, D.E.; Ryan, L.M. Concurrent Prediction of Hospital Mortality and Length of Stay from Risk Factors on Admission. *Health Serv. Res.* **2002**, *37*, 631–645. [\[CrossRef\]](#)
39. Awad, A.; Bader, M.; McNicholas, J. Patient length of stay and mortality prediction: A survey. *Health Serv. Manag. Res.* **2017**, *30*, 105–120. [\[CrossRef\]](#)
40. Du, X.; Wang, H.; Wang, S.; He, Y.; Zheng, J.; Zhang, H.; Hao, Z.; Chen, Y.; Xu, Z.; Lu, Z. Machine Learning Model for Predicting Risk of In-Hospital Mortality after Surgery in Congenital Heart Disease Patients. *Rev. Cardiovasc. Med.* **2022**, *23*, 376. [\[CrossRef\]](#) [\[PubMed\]](#)
41. Gowd, A.K.; Agarwalla, A.; Beck, E.C.; Rosas, S.; Waterman, B.R.; Romeo, A.A.; Liu, J.N. Prediction of total healthcare cost following total shoulder arthroplasty utilizing machine learning. *J. Shoulder Elb. Surg.* **2022**, *31*, 2449–2456. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Rao, A.R.; Raunak, J.; Mrityunjai, S.; Rahul, G. Predictive interpretable analytics models for forecasting healthcare costs using open healthcare data. *Healthc. Anal.* **2024**, *6*, 100351. [\[CrossRef\]](#)
43. Lu, Y.; Lavoie-Gagne, O.; Forlenza, E.M.; Pareek, A.; Kunze, K.N.; Forsythe, B.; Levy, B.A.; Krych, A.J. Duration of Care and Operative Time Are the Primary Drivers of Total Charges After Ambulatory Hip Arthroscopy: A Machine Learning Analysis. *Arthrosc. J. Arthrosc. Relat. Surg.* **2022**, *38*, 2204–2216.e3. [\[CrossRef\]](#)
44. Muhlestein, W.E.; Akagi, D.S.; McManus, A.R.; Chambless, L.B. Machine learning ensemble models predict total charges and drivers of cost for transsphenoidal surgery for pituitary tumor. *J. Neurosurg.* **2019**, *131*, 507–516. [\[CrossRef\]](#)

45. Placido, D.; Thorsen-Meyer, H.C.; Kaas-Hansen, B.S.; Reguant, R.; Brunak, S. Development of a dynamic prediction model for unplanned ICU admission and mortality in hospitalized patients. *PLoS Digit. Health* **2023**, *2*, e0000116.
46. Gupta, A.; Liu, T.; Crick, C. Utilizing time series data embedded in electronic health records to develop continuous mortality risk prediction models using hidden Markov models: A sepsis case study. *Stat. Methods Med. Res.* **2020**, *29*, 3409–3423.
47. Jenkins, D.A.; Martin, G.P.; Sperrin, M.; Riley, R.D.; Debray, T.P.; Collins, G.S.; Peek, N. Continual updating and monitoring of clinical prediction models: Time for dynamic prediction systems? *Diagn. Progn. Res.* **2021**, *5*, 1. [[CrossRef](#)] [[PubMed](#)]
48. Karnuta, J.M.; Churchill, J.L.; Haeberle, H.S.; Nwachukwu, B.U.; Taylor, S.A.; Ricchetti, E.T.; Ramkumar, P.N. The value of artificial neural networks for predicting length of stay, discharge disposition, and inpatient costs after anatomic and reverse shoulder arthroplasty. *J. Shoulder Elb. Surg.* **2020**, *29*, 2385–2394. [[CrossRef](#)]
49. Natekin, A.; Knoll, A. Gradient boosting machines, a tutorial. *Front. Neurobot.* **2013**, *7*, 21. [[CrossRef](#)]
50. Rigatti, S.J. Random Forest. *J. Insur. Med.* **2017**, *47*, 31–39. [[CrossRef](#)]
51. Stulberg, J.J.; Haut, E.R. Practical Guide to Surgical Data Sets: Healthcare Cost and Utilization Project National Inpatient Sample (NIS). *JAMA Surg.* **2018**, *153*, 586–587. [[CrossRef](#)]
52. ICD-10-CM/PCS MS-DRG; v37.0; Definitions Manual; Centers for Medicare & Medicaid Services: Baltimore, MD, USA, 2019.
53. Averill, R.F.; Goldfield, N.; Hughes, J.S.; Bonazelli, J.; McCullough, E.C.; Steinbeck, B.A.; Mullin, R.; Tang, A.M. *All Patient Refined Diagnosis Related Groups (APR-DRGs) Version 20.0*; M.H.I. Systems: Tokyo, Japan, 2003.
54. *Elixhauser Comorbidity Software Refined for ICD-10-CM DIAGNOSES*; v2024.1; Agency for Healthcare Research and Quality (AHRQ): Rockville, MD, USA, 2023.
55. Lee, S.Y.; Lee, S.H.; Tan, J.H.H.; Foo, H.S.L.; Phan, P.H.; Kow, A.W.C.; Lwin, S.; Seah, P.M.Y.; Mordiffi, S.Z. Factors associated with prolonged length of stay for elective hepatobiliary and neurosurgery patients: A retrospective medical record review. *BMC Health Serv. Res.* **2018**, *18*, 5. [[CrossRef](#)] [[PubMed](#)]
56. Fetene, D.; Tekalegn, Y.; Abdela, J.; Aynalem, A.; Bekele, G.; Molla, E. Prolonged Length of Hospital Stay and Associated Factors Among Patients Admitted at a Surgical Ward in Selected Public Hospitals Arsi Zone, Oromia, Ethiopia, 2022. *medRxiv* **2022**. [[CrossRef](#)]
57. Murdoch, W.J.; Singh, C.; Kumbier, K.; Abbasi-Asl, R.; Yu, B. Interpretable machine learning: Definitions, methods, and applications. *arXiv* **2019**, arXiv:1901.04592. [[CrossRef](#)]
58. Darst, B.F.; Malecki, K.C.; Engelman, C.D. Using recursive feature elimination in random forest to account for correlated variables in high dimensional data. *BMC Genet.* **2018**, *19*, 65. [[CrossRef](#)]
59. Berrar, D. Cross-Validation. In *Reference Module in Life Sciences Encyclopedia of Bioinformatics and Computational Biology*; Ranganathan, S., Gribskov, M., Nakai, K., Christian Schönbach, C., Eds.; Elsevier: Amsterdam, The Netherlands, 2019; Volume 1, pp. 542–545. [[CrossRef](#)]
60. Awad, M.; Fraihat, S. Recursive Feature Elimination with Cross-Validation with Decision Tree: Feature Selection Method for Machine Learning-Based Intrusion Detection Systems. *J. Sens. Actuator Netw.* **2023**, *12*, 67. [[CrossRef](#)]
61. Hossin, M.; Sulaiman, M.N. A review on evaluation metrics for data classification evaluations. *Int. J. Data Min. Knowl. Manag. Process* **2015**, *5*, 1.
62. Chicco, D.; Warrens, M.J.; Jurman, G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Comput. Sci.* **2021**, *7*, e623. [[CrossRef](#)]
63. Musolf, A.M.; Holzinger, E.R.; Malley, J.D.; Bailey-Wilson, J.E. What makes a good prediction? Feature importance and beginning to open the black box of machine learning in genetics. *Human Genet.* **2022**, *141*, 1515–1528. [[CrossRef](#)]
64. Pudjihartono, N.; Fadason, T.; Kempa-Liehr, A.W.; O'Sullivan, J.M. A Review of Feature Selection Methods for Machine Learning-Based Disease Risk Prediction. *Front. Bioinform.* **2022**, *2*, 927312. [[CrossRef](#)]
65. Spencer, C.S.; Gaskin, D.J.; Roberts, E.T. The Quality Of Care Delivered To Patients Within The Same Hospital Varies By Insurance Type. *Health Aff.* **2013**, *32*, 1731–1739.
66. Roll, K.; Stargardt, T.; Schreyögg, J. Effect of Type of Insurance and Income on Waiting Time for Outpatient Care. *Geneva Pap. Risk Insur.-Issues Pract.* **2012**, *37*, 609–632.
67. Obeng-Gyasi, S.; O'Neill, A.; Zhao, F.; Kircher, S.M.; Lava, T.R.; Wagner, L.I.; Miller, K.D.; Sparano, J.D.; Sledge, G.W.; Carlos, R.C. Impact of insurance and neighborhood socioeconomic status on clinical outcomes in therapeutic clinical trials for breast cancer. *Cancer Med.* **2021**, *10*, 45–52. [[PubMed](#)]
68. Bassan, R.; Hoelzer, D. Modern Therapy of Acute Lymphoblastic Leukemia. *J. Clin. Oncol.* **2011**, *29*, 532–543. [[CrossRef](#)] [[PubMed](#)]
69. Prosperi, M.; Guo, Y.; Sperrin, M.; Koopman, J.S.; Min, J.S.; He, X.; Rich, S.; Wang, M.; Buchan, I.E.; Bian, J. Causal inference and counterfactual prediction in machine learning for actionable healthcare. *Nat. Mach. Intell.* **2020**, *2*, 369–375. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.