emergency care and medicine

MDPI

Article

Development and Validation of a Machine Learning Model That Predicts Short Inpatient Stays Among Urgent Admissions

Yan Gao [1,*], Sunku Srivatsava [2], Hong Choon Oh [1], Siang Hiong Goh [2] and Hoon Chin Steven Lim [2]

[1] Health Services Research, Changi General Hospital, Singapore 529889, Singapore; oh.hong.choon@singhealth.com.sg
[2] Emergency Department, Changi General Hospital, Singapore 529889, Singapore; sunku.srivatsava@mohh.com.sg (S.S.); goh.siang.hiong@singhealth.com.sg (S.H.G.); lim.hoon.chin@singhealth.com.sg (H.C.S.L.)
* Correspondence: gao.yan@singhealth.com.sg

Abstract: Background/Objectives: This study aimed to explore the feasibility of predicting short stays among urgent admissions to an acute hospital in Singapore. With an increase in the average length of stay (LOS) in hospitals in recent years, accurately predicting short stays could enable hospitals to better manage inpatient demand and reduce emergency department (ED) overcrowding. Methods: This was a retrospective study of urgent admissions to Changi General Hospital, Singapore, from 1 January 2016 to 30 June 2022. To identify potential short stayers, a total of 25 features comprising demographic characteristics, admission and clinical characteristics, and healthcare utilization history were analyzed for each admitted patient at the point when the ED physician decided to admit the patient. The dataset was further split into a development dataset and an external validation dataset based on the year of admission. A CatBoost classifier was trained using 75% of the development dataset. Apart from reporting the model's prediction accuracy, we conducted various analyses and simulations to study the effects of the features crucial to the prediction output. Results: The prediction accuracy of the model was evaluated on both the development test dataset (25%) and the external validation dataset. On the former, the area under the receiver operating characteristic (AUROC) and the area under the precision-recall curve (AUPRC) were 0.803 (95% CI: 0.799, 0.808) and 0.755 (95% CI: 0.749, 0.762), respectively, with the precision = 0.700 (95% CI: 0.694, 0.707) and recall = 0.692 (95% CI: 0.685, 0.699). On the external validation dataset, the performance was similar. The ED diagnosis and whether the admission required a surgical procedure were the most important features for making the prediction. Conclusions: The LOS prediction model could help providers to identify short stayers early in the course of their inpatient journeys so they could make interventions to better manage the overall utilization of hospital beds.

Keywords: inpatient admissions; length of stay; machine learning

check for updates

Academic Editor: George Notas

Received: 7 January 2025
Revised: 11 February 2025
Accepted: 20 February 2025
Published: 25 February 2025

Citation: Gao, Y.; Srivatsava, S.; Oh, H.C.; Goh, S.H.; Lim, H.C.S. Development and Validation of a Machine Learning Model That Predicts Short Inpatient Stays Among Urgent Admissions. *Emerg. Care Med.* **2025**, *2*, 11. https://doi.org/10.3390/ecm2010011

Copyright: © 2025 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) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

As Singapore moves away from the COVID-19 crisis, healthcare demands continue to increase and the average length of stay (LOS) in public hospitals has increased significantly, from 6.1 days to 7 days. This has led to an increase in local hospital bed utilization by 15%, which has in turn contributed to longer waiting times for inpatient admission [1]. Compared to other developed countries, such as the United States, the acute hospital bed occupancy rate in Singapore is higher. This is likely the result of multiple factors, such as a

lower capacity per 1000 of the population in intermediate or long-term care facilities that offer rehabilitation and other non-acute services [2].

Unceasingly, hospitals have aimed to cut down on unnecessarily long LOSs because prolonged stays are known to be associated with increased mortality, hospital-acquired infection, higher costs, and higher utilization of healthcare resources. A better patient experience might also be achieved by improving the inpatient flow, starting at the point of admission until the time of hospital discharge. Shortening the LOS would help to decongest emergency departments (EDs) by promoting the egress of admitted patients from the ED, reducing its overcrowding while at the same time enhancing patient safety [3,4].

Whenever decisions are made to admit ED patients to inpatient wards, the accurate identification of short stayers would enable doctors to consider alternative care plans that could potentially preclude admissions. For those patients who still need to be admitted to the hospital, interventions, such as right-siting to short-stay observation medicine units or the initiation of early discharge planning, could help to keep the LOS short without affecting the quality of care. Evidence-based transitional care activities could also be prioritized for such patients to improve outcomes.

Recently, machine learning (ML) technology has demonstrated the capability to predict the LOS of general admissions [5–7]. In Singapore, the National University Health System (NUHS) has begun to harness the power of ML to predict the LOS. By analyzing both the patients' histories and doctors' notes in real time, accurate predictions of patients' LOSs may be made to flag those who are expected to stay for more than two weeks [8].

In a study by Ng and Tan [9], a predictive model was developed to identify frequent admitters prior to discharge. This technology enabled the adoption of individual-centric discharge processes by hospitals so that the latter could manage frequent admissions, like using a "high risk" case management program to curb their consumption of hospital resources.

Changi General Hospital (CGH) is a tertiary hospital that receives over 130,000 ED visits and has 40,000 admissions annually. The hospital is an academic medical center belonging to a regional healthcare system in Singapore, which serves an estimated population base of 1.5 million. Over the years, CGH has managed short stayers to alleviate the bed demand pressure by sending eligible short stayers to short-stay wards in the ED for timely care and discharge. While this has been performed manually by doctors based on specific selection criteria, we hope to evaluate whether a machine learning model could support the hospital in identifying a bigger pool of short stayers who, if the potential intervention was early enough, could reduce inpatient bed demand. To our knowledge, this has not been investigated locally or regionally.

This paper is organized as follows: Section 2 describes the materials and methods used in this study. Section 3 discusses the main findings, including the classification performance and feature importance analysis. Section 4 presents the conclusions and limitations of the current study.

2. Materials and Methods

The design and reporting of this study followed the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis + Artificial Intelligence (TRIPOD + AI) [10].

2.1. Data Source

We conducted a retrospective, single-center study of CGH to derive a short-stayer predictive model using electronic health records (EHRs) of patients in the ED and inpatient wards. This study was reviewed by Singapore Health Services' Centralised Institutional

Review Board, which determined that this study was a model evaluation project that did not require further ethical deliberation.

2.2. Patient Selection

All urgent admissions through the CGH ED between 1 January 2016 and 30 June 2022 were employed in development and validation of a new short-stayer predictive model. In total, 164,941 urgent admissions during the period from 1 January 2016 to 31 December 2019 were extracted as the development dataset. A set of 105,365 urgent admissions from 1 January 2020 to 30 June 2022 were employed separately as the external validation dataset, in accordance with the TRIPOD statement [11]. For the development dataset, we further split it into 75% (n = 123,705) training and 25% (n = 41,236) testing datasets using random sampling. All the data in this study were extracted from the hospital's Electronic Health Intelligence System before they were de-identified by a trusted third party and made available to the study team.

2.3. Outcomes

The LOS was calculated using the difference between the discharge date and the admission date. Short stayers were defined as those with LOSs less than or equal to 3 days, and not discharged due to death, cancellation, or against medical advice. All the remaining admissions formed the negative class.

2.4. Variables

A total of 25 feature variables (listed in Tables 1 and 2) were identified in this study, and they include the demographics and historical healthcare utilization patterns of patients at their respective time of ED admission. These features have been shown to be useful for the purpose of short-stayer predictions [12–15], or assessed to be clinically relevant by our ED doctors.

Table 1. Demographic characteristics of the study population.

Characteristics		Development Dataset	External Validation Dataset
LOS in days, mean (SD)		6.6 (10.4)	6.8 (10.3)
Age in years, mean (SD)		63.5 (19.3)	64.2 (19.0)
Gender, n (%)			
	Female	72,363 (43.9)	44,994 (42.7)
	Male	92,578 (56.1)	60,371 (57.3)
Race, n (%)			
	Chinese	96,875 (58.7)	62,710 (59.5)
	Indian	15,102 (9.2)	9996 (9.5)
	Malay	34,805 (21.1)	20,488 (19.4)
	Other	18,159 (11.0)	12,171 (11.6)
Singapore Resident, n (%)			
	No	58,378 (35.4)	41,454 (39.3)
	Yes	106,563 (64.6)	63,911 (60.7)
Residential Type *, n (%)			
	1–2 Room	18,238 (11.9)	11,544 (11.7)
	3 Room	37,095 (24.1)	24,830 (25.2)
	4 Room	48,338 (31.5)	30,659 (31.1)
	5 Room or Executive	33,112 (21.5)	19,706 (20.0)
	Condominium	9666 (6.3)	6968 (7.1)
	Landed	7230 (4.7)	4798 (4.9)

Table 1. *Cont.*

Characteristics		Development Dataset	External Validation Dataset
Rental Flat Residents, n (%)			
	No	147,696 (96.1)	94,759 (96.2)
	Yes	5983 (3.9)	3746 (3.8)
Admitting Bed Class [#], n (%)			
	A	12,564 (7.6)	4492 (4.3)
	B1	9933 (6.0)	7688 (7.3)
	B2	45,090 (27.3)	20,705 (19.7)
	C	97,354 (59.0)	72,480 (68.8)

* 1–5 rooms refer to public housing flat types, while condominium and landed are private properties. [#] Abbreviations: A = air-conditioned single room; B1 = air-conditioned 4-bedded room; B2 = natural-ventilated 4–6 bedded room; C = natural-ventilated 8-bedded open ward. Patient's choice of class of beds determines amount of government subsidies (if any) after accounting for patient's eligibilities. Note: There were 18,122 records with missing information on residential type and rental flat. Missing values were excluded when calculating frequencies and percentages.

Table 2. Admission, clinical characteristics, and healthcare utilization history of study population.

Characteristics		Development Dataset	External Validation Dataset
Admission Time of Day, n (%)			6.8 (10.3)
	0800 to 1229 h.	17,844 (10.8)	11,063 (10.5)
	1230 to 1659 h.	47,870 (29.0)	31,245 (29.7)
	1700 to 0759 h.	99,227 (60.2)	63,057 (59.8)
Admission Day of Week, n (%)			60,371 (57.3)
	Monday	26,341 (16.0)	16,718 (15.9)
	Tuesday	24,594 (14.9)	15,661 (14.9)
	Wednesday	24,505 (14.9)	15,645 (14.8)
	Thursday	23,650 (14.3)	15,462 (14.7)
	Friday	23,401 (14.2)	15,251 (14.5)
	Saturday	21,954 (13.3)	13,730 (13.0)
	Sunday	20,496 (12.4)	12,898 (12.2)
ED Visit Time of Day *, n (%)			63,911 (60.7)
	0800 to 1229 h.	38,664 (23.4)	24,890 (23.6)
	1230 to 1659 h.	50,488 (30.6)	33,711 (32.0)
	1700 to 0759 h.	75,789 (45.9)	46,764 (44.4)
ED Referral Type, n (%)			30,659 (31.1)
	Referred	44,094 (26.7)	30,013 (28.5)
	No referral	120,847 (73.3)	75,352 (71.5)
ED Conveyance Type, n (%)			4798 (4.9)
	No Ambulance	94,966 (57.6)	58,091 (55.1)
	Private Ambulance	8717 (5.3)	5081 (4.8)
	SCDF[+] Ambulance	61,258 (37.1)	42,193 (40.0)
Admission Bed Prioritization Category [^], n (%)			
	0	13,434 (8.2)	9023 (8.6)
	1	3691 (2.3)	705 (0.7)
	2	8009 (4.9)	1973 (1.9)
	3	27,323 (16.7)	18,179 (17.3)
	4	45,684 (27.9)	36,578 (34.7)
	5	65,829 (40.1)	38,842 (36.9)
Required Surgical Procedure, n (%)			
	No	125,176 (75.9)	81,481 (77.3)
	Yes	39,765 (24.1)	23,884 (22.7)

Table 2. *Cont.*

Characteristics		Development Dataset	External Validation Dataset
Admit Specialty, n (%)			
	CVM	19,509 (11.8)	11,719 (11.1)
	END	1987 (1.2)	875 (0.8)
	ENT	659 (0.4)	302 (0.3)
	GAS	8128 (4.9)	4602 (4.4)
	GER11	10,255 (6.2)	6273 (6.0)
	MED	66,144 (40.1)	46,621 (44.2)
	OTO	16,327 (9.9)	8081 (7.7)
	REN	2564 (1.6)	3888 (3.7)
	RES	12,405 (7.5)	5696 (5.4)
	SUR	20,682 (12.5)	12,527 (11.9)
	URO	2458 (1.5)	1444 (1.4)
	Others	3823 (2.3)	3337 (3.2)
ED Diagnosis, n (%)			
	Pneumonia, unspecified	9292 (5.6)	5513 (5.2)
	Sepsis, unspecified	6014 (3.6)	2555 (2.4)
	Fluid overload	4363 (2.6)	3178 (3.0)
	Stroke, not specified as hemorrhage or infarction	4271 (2.6)	2757 (2.6)
	Urinary tract infection, site not specified	3845 (2.3)	2123 (2.0)
	Dizziness and giddiness	3698 (2.2)	2357 (2.2)
	Chest pain, unspecified	3657 (2.2)	3390 (3.2)
	Cellulitis, unspecified	3069 (1.9)	2145 (2.0)
	Tendency to fall, not elsewhere classified	3057 (1.9)	2886 (2.7)
	Other and unspecified abdominal pain	2677 (1.6)	2313 (2.2)
	Other	120,998 (73.4)	76,148 (72.3)
Mean Number of Urgent Admissions in Preceding 12 months (SD)		1.0 (2.2)	1.0 (2.2)
Mean Number of Elective Admissions in Preceding 12 months (SD)		0.1 (0.3)	0.1 (0.3)
Mean Number of ED Visits in Preceding 12 months (SD)		0.7 (3.5)	0.5 (3.7)
Mean Number of Outpatient Specialist Visits in Preceding 12 months (SD)		3.5 (5.5)	3.3 (5.0)
Mean Number of Outpatient Specialist Types Visited in Preceding 12 months (SD)		1.6 (2.0)	1.5 (2.0)
Mean Percentage of No-shows at Outpatient Specialist Visits in Preceding 12 months (SD) *		0.3 (0.3)	0.3 (0.3)

Table 2. *Cont.*

Characteristics	Development Dataset	External Validation Dataset
Mean LOS of Previous Admission (SD) ^	7.1 (10.7)	6.9 (10.0)
Mean Number of Days Lapsed Since the Previous Admission (SD) ^	238.2 (301.1)	406.0 (526.4)
Admitted to Nursing Home or Community Hospital in Preceding 12 months, n (%)		
No	154,068 (93.4)	97,983 (93.0)
Yes	10,873 (6.6)	7382 (7.0)

* 115,392 admission episodes did not have history of outpatient specialist visit in preceding 12 months. ^ 119,925 admission episodes did not have history of admission in preceding 12 months. ~1036 admission episodes did not have records for admission bed prioritization category. Acuity level (zero to five) was based on patient's clinical condition as assessed by ED physician, with acuity level being higher with smaller number. Abbreviations: CVM = cardiovascular medicine; END = endocrinology; ENT = ear nose and throat; GAS = gastroenterology; GER = geriatric medicine; MED = general medicine, OTO = orthopedic; PSY = psychiatric medicine; REN = renal medicine; RES = respiratory; RMD = rehabilitation; SUR = general surgery. + Abbreviation: SCDF = Singapore Civil Defence Force.

2.5. Missing Data

There were missing data in the following feature variables: residential type, rental flat, and admission bed prioritization category. See Tables 1 and 2 for the number of missing records. We did not exclude records due to missing values, but treated missingness as an attribute.

2.6. Development

In this study, we employed CatBoost [16] classifier to develop the predictive model. CatBoost is a member of the family of Gradient-Boosted Decision Trees (GBDT) machine learning ensemble techniques. The latter are based on iterative optimization of regression trees in decision trees so that loss function in each iteration is minimized before final prediction result is derived. Essentially, CatBoost is well suited to machine learning tasks like those in our study, which involve much categorical, heterogeneous data, since it is capable of overcoming the gradient bias and avoids overfitting effectively [17]. In addition, CatBoost handles missing values intrinsically and thus we did not exclude records due to missing values.

While CatBoost is known for good performance using default parameters, we still tuned some of the important parameters using 3-fold cross-validation.

2.7. Analysis

Predictive model performance was evaluated by examining the area under the receiver operator curve (AUROC) and area under the precision-recall curve (AUPRC), with associated 95% confidence interval (CI) on the testing and external validation datasets.

We also used CatBoost's feature importance ranking method to determine the relative importance of features in the model. The individual importance value for each of the input features was estimated by calculating how much, on average, the prediction changed if the feature's value changed [16]. In our effort to develop an interpretable model, we used a post hoc interpretability method that employed Shapley Additive explanations (SHAP) [18] to examine the relative importance of all features when making a prediction.

All analyses were performed using Python 3.11 on a Windows 11 workstation with i7 Intel processor (Intel Corporation, Santa Clara, CA, USA) and 32 GB memory. A summary of our study dataset setup and methods used is shown in Figure 1.

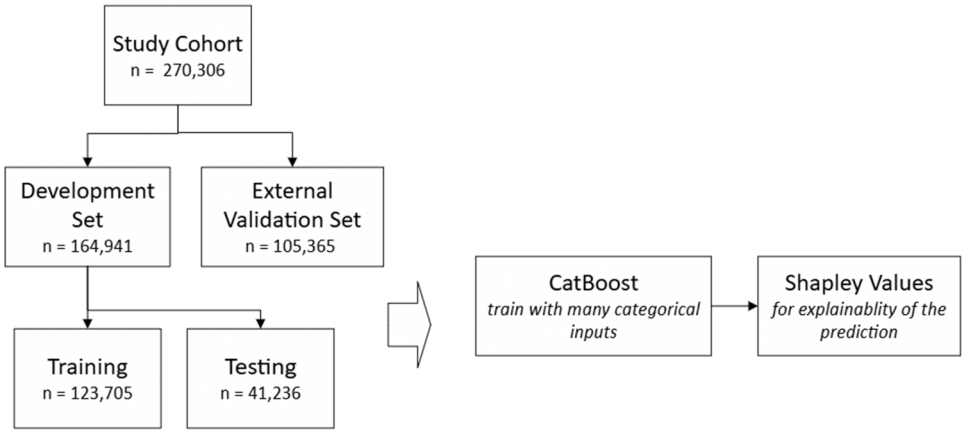


Figure 1. Study dataset and methods.

3. Results

In the development and external validation datasets, the proportion of urgent admission patients who were short stayers was 45.1% and 43.7%, respectively. Table 1 shows the demographic characteristics of the study population in these two groups. Table 2, in turn, summarizes the admission and clinical characteristics as well as the healthcare utilization history of the study population in the two datasets. To this end, it is important to note that the healthcare utilization history records were limited to episodes which involved the CGH. Only top 10 diagnoses (out of 2478 possible ones), in terms of frequency, are listed.

A CatBoost model was trained and fine-tuned using 3-fold cross-validation. The hyperopt [19] package was used to tune l2_leaf_reg and learning_rate used in CatBoost. The final parameters used were iterations = 500, loss_fuction = 'Logloss', learning_rate = 0.16196, and l2_leaf_reg = 2.

When the trained model was employed on the development test dataset, the area under the receiver operating characteristic (AUROC) and the area under precision-recall curve (AUPRC) were 0.803 (95% CI: 0.799, 0.808) and 0.755 (95% CI: 0.749, 0.762), respectively, with precision = 0.700 (95% CI: 0.694, 0.707) and recall = 0.692 (95% CI: 0.685, 0.699) at a default cut-off of 0.5. Overall, the model performance remained similar when applied to the external validation dataset. Table 3 shows a summary of the model's performance on both the development test and external validation datasets.

Table 3. AUROC, AUPRC, precision, and recall (95% CI) based on development test and external validation datasets.

Dataset	AUROC	AUPRC	Precision	Recall
Development test	0.803 (0.799, 0.808)	0.755 (0.749, 0.762)	0.700 (0.694, 0.707)	0.692 (0.685, 0.699)
External validation	0.792 (0.790, 0.795)	0.733 (0.729, 0.738)	0.670 (0.666, 0.674)	0.698 (0.694, 0.702)

The precision and recall of the model can be adjusted based on operational needs. When limited resources are available for clinical and operational interventions, we can use a threshold that has a higher precision with a reduced recall. Table 4 shows the trade-off between precision and recall, with a threshold of 0.5 offering a relatively good performance on both fronts.

We also used CatBoost's feature importance ranking method to determine the relative importance of the features in the model. The key features which predicted short stayers included the ED diagnosis, required surgical procedure, age, admit specialty, patient acuity category, LOS of the previous admission, admission day of the week, ED conveyance type,

and number of days lapsed since the previous admission. A full list of all the features, ranked by their feature importance, is shown in Figure 2, with the ED diagnosis and required surgical procedure being ranked the top two.

Table 4. Trade-off between precision and recall when the operating point (threshold) is changed.

Threshold	Precision	Recall
0.5	0.700	0.692
0.6	0.749	0.549
0.7	0.805	0.383
0.8	0.863	0.199
0.9	0.929	0.042

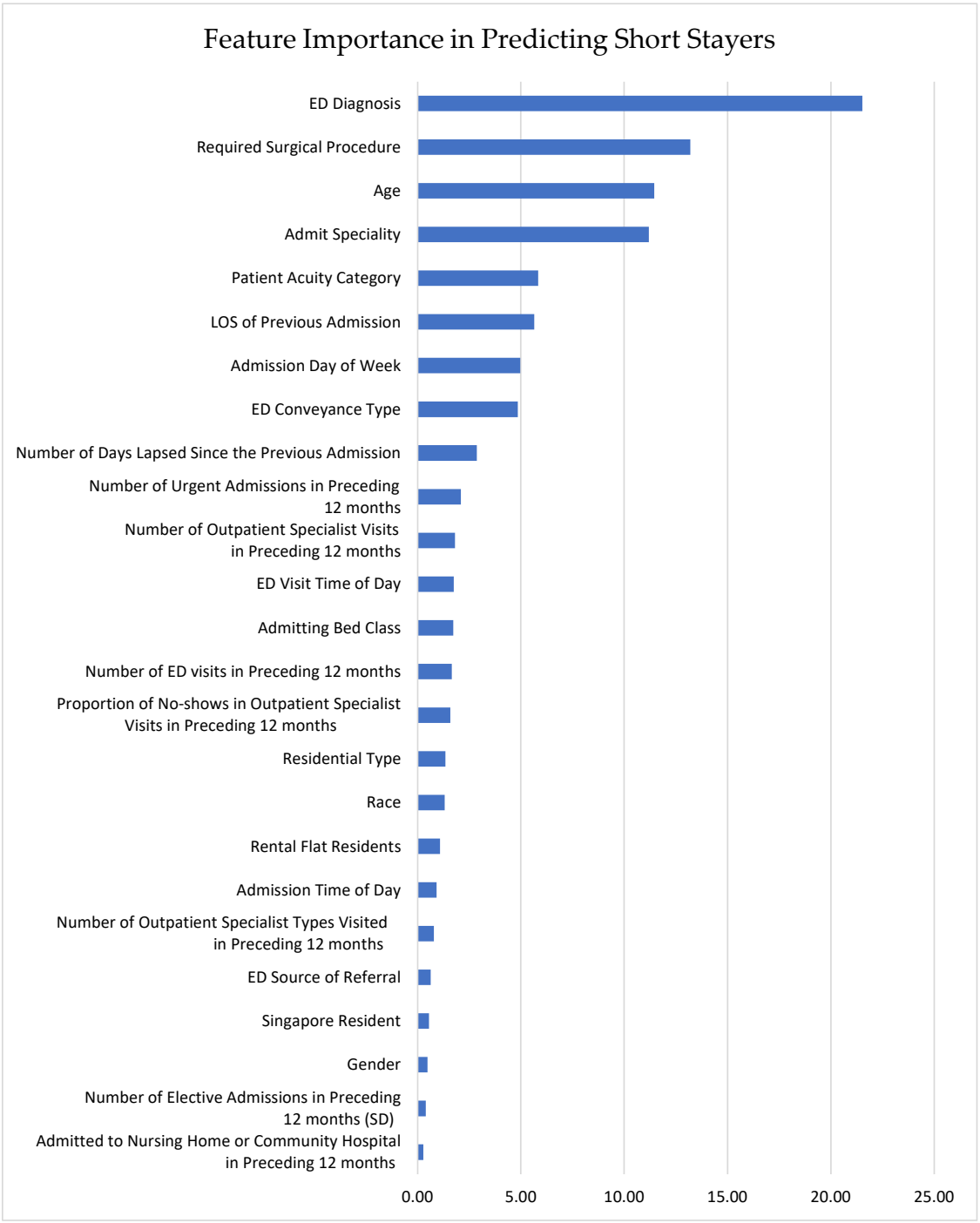


Figure 2. Feature importance of all the features used in the model.

We used Shapley Additive explanations (SHAP) [18] to examine the impact of the ED diagnosis type on the model's predictions. The SHAP value measures how important a diagnosis is towards the prediction of a short stayer. The higher the value, the greater the likelihood that the diagnosis contributes towards a short-stayer prediction. Figure 3 shows the SHAP values for the top 10 most frequent diagnoses, the top 10 diagnoses that predicted a short stayer, and the top 10 diagnoses that predicted against a short stayer. Among the 10 most frequent diagnoses, chest pain and dizziness lead to short stays, while the remaining 8 all contribute to longer stays. It is also evident that trauma injuries can be good predictors of short stayers or non-short stayers, depending on the nature of the injuries.

It is also important to note that SHAP values can be employed to interpret the predictive model's classification of a patient. For example, Figure 4 is a waterfall plot of the SHAP values, which shows how each feature of a real patient (from the development test set) contributed to the prediction of them being a non-short stayer. The patient was diagnosed with pneumonia in the ED, was admitted on a Thursday, was brought in by an SCDF ambulance, and had not turned up for all of his outpatient specialist visits. These all contributed towards a longer stay and outweighed other features which pointed to a short stay, such as the admit specialty and no required surgical procedure. This case was a correct prediction (true negative). Thus, the visualization offered by the waterfall plot, as illustrated in Figure 4, could potentially help ED doctors to understand better the rationale behind the model's classification of a patient before the doctors decide if there is any need to change the classification after adding in their clinical judgement.

Our model requires a doctor's judgement or projection at the point of admitting a patient as to whether that will need a surgical procedure during their inpatient stay. To assess the potential impact of doctors' projection accuracies on the model's performance, we simulated three scenarios where there was a 10%, 20%, and 30% error by the doctors when indicating the surgical procedure required by a patient at the point of admission, using the development test dataset. For each of these scenarios, we randomly generated errors in the aforementioned indicators so that the overall error percentage matched the level of that scenario, and we repeated this process for over 500 runs to obtain the corresponding mean AUROC and AUPRC with a 95% confidence interval. Table 5 shows the decline in the model's performance on the development test dataset as the error percentage increased, as well as when this field was completely removed from the model. It is observed that when the percentage of mistakes is 10%, the inclusion of the field is still better than that without the field. This indicates that including the field would improve the predictive ability of the model to identify short stayers.

Table 5. Predictive model performance in simulation study.

Percentage Error in Projection of Surgical Procedure Required at the Point of Admission	AUROC	AUPRC
10%	0.786 (0.784, 0.787)	0.742 (0.741, 0.744)
20%	0.769 (0.767, 0.771)	0.729 (0.727, 0.731)
30%	0.754 (0.752, 0.757)	0.716 (0.713, 0.719)
Absence of surgery requirement in the predictive model	0.781 (0.776, 0.785)	0.734 (0.727, 0.741)

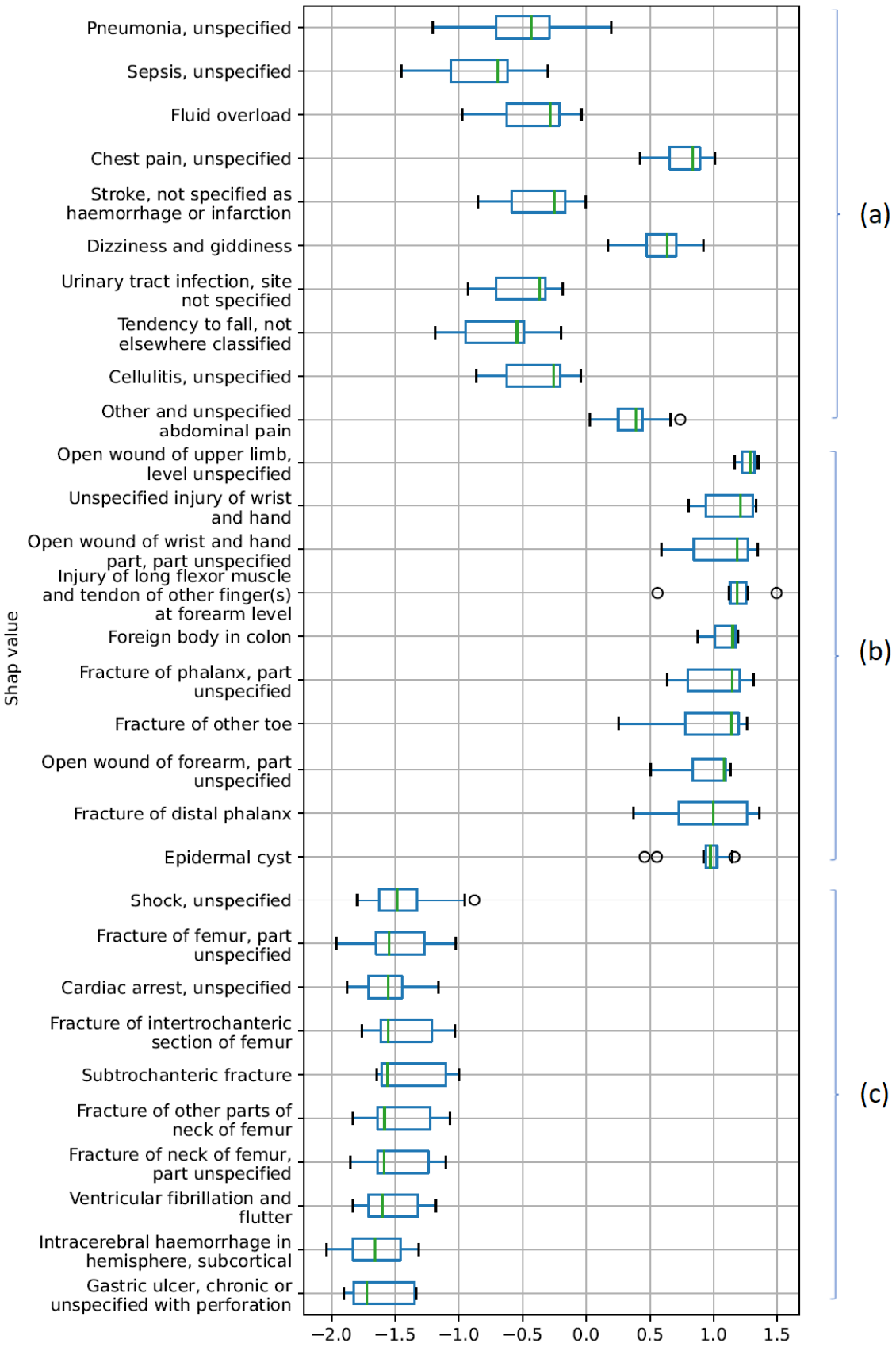


Figure 3. SHAP values of (**a**) top 10 frequent diagnoses, (**b**) top 10 diagnoses predicting short stayers, and (**c**) top 10 diagnoses predicting against short stayers.

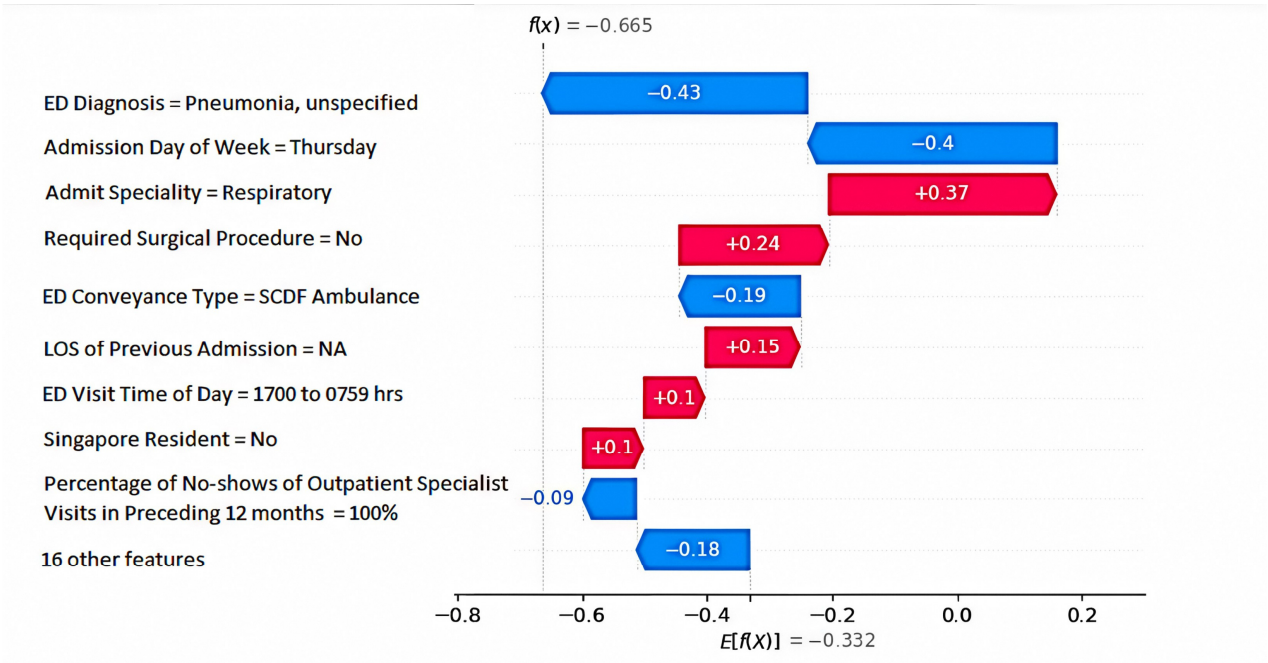


Figure 4. Waterfall chart showing SHAP values for a real patient predicted to be a non-short stayer. Red: positive effect towards a short stayer. Blue: negative effect towards a short stayer.

4. Discussion

ML models have been developed to support clinical management and to improve workflows in hospitals. Some of these models employ biomarkers and other data points to predict hospital admissions [20,21]. ML and artificial neural networks have shown promise in the field of predicting hospital LOS for patients with specific diagnoses, such as COVID-19 [22], dengue fever, syncope [23], and chronic obstructive pulmonary disease (COPD) [24], and for surgical conditions, such as post-colorectal cancer surgery [25] and pediatric burn patients [26]. However, the development of ML models for identifying patients at risk of being short stayers in an acute hospital is more challenging as it is influenced by multiple diverse and dynamic patient factors.

In our study, we developed an ML-based short-stayer predictive model, using the CatBoost Classifier and with good performance, to help identify inpatient short stayers following urgent admissions from the ED. Once potential short-stayer patients are identified using our model, we foresee multiple advantages that may benefit both patients and healthcare systems. They include, but are not limited to, improved patient workflows and better resource allocation and informed decision making by patients and families. Patients who are identified as short stayers could be initiated on early discharge planning, reducing bed waiting time among newly admitted patients. Moreover, informed discussions with these patients and their family members would facilitate discharge planning and improve overall satisfaction. This approach streamlines the process, better addresses the needs of patients and families, and helps to coordinate the placement of a patient. A study by Alharbi et al. found that early discharge planning and facilitation of the patient discharge process significantly reduces the average length of an inpatient stay and hospital costs, and improved patient outcomes [27].

Wards with identified short stayers may also trigger other actions earlier in order to smoothen the discharge process, such as more regular reviews of patients' progress through daily multi-disciplinary interactions when needed; planning for transitional care post-discharge [28]; and more proactive, earlier alerting of social care practitioners for patients with poor social determinants of health. Quality improvement projects which

have addressed these ideas and facilitated the flow of discharged patients to the "discharge lounge" or discharge hospitality center have resulted in reduced LOSs and improved the timeliness of departure from inpatient beds, respectively [29–31].

To this end, it is important to highlight that our model's threshold may be adjusted to meet the operational needs of identifying short stayers. For example, in situations where there is a need to better allocate scarce resources, such as limited fast-track outpatient clinic appointment slots or investigation slots (force-in outpatient MRI slots), the threshold of the model may be increased. This would increase the model's precision (positive predictive value) so that these limited resources could be spent on the "true" short stayers, leading to better admission avoidance and inpatient discharge strategies. Conversely, in situations where resources are not limited, for example, the sharing of digital information concerning "hospital discharge instructions to patients", the threshold of the model may be decreased in order to increase recall (sensitivity), leading to most short stayers receiving this service message with their early discharge instructions. Even though, as the precision of the model is reduced, a greater number of non-short stayers would also receive these messages, there would be no harm from receiving such information earlier.

Overall, our model is able to fairly accurately identify patients whose admission is likely to be less than three days, as evident by the AUROC of 0.792 and AUPRC of 0.733 when tested on the external validation group. The model was able to aggregate structured data from general features, which were available at the point of admission. Nevertheless, the model has four key limitations that need to be addressed before it can be deployed in any hospital. First, the model needs to be validated by other hospital sites to further prove its generalizability. Second, our current model is not as interpretable as other reported predictive models based on scoring [32–34]. To improve interpretability, we used SHAP values to help healthcare practitioners to better understand the reasons behind the patient classifications made by our model. Third, we have not accounted for real-time diagnostic test results and unstructured data in our model. The addition of unstructured data has been found to improve predictions [35].

The final key limitation of our model is its reliance on ED doctors to predict whether a newly admitted patient requires a surgical procedure or not. If this feature is omitted from our model, the AUPRC drops marginally from 0.755 to 0.734 based on the retrospective test sets. We also found that when ED doctors can indicate the need for a surgical procedure with no more than 10% error, the aforementioned feature remains useful for the identification of short stayers.

5. Conclusions

We have developed a machine learning, predictive model for short stayers with good performance that helps to identify urgent admissions that may then receive interventions to reduce hospital bed demand. It required only the data that were readily obtained from existing databases at the point of admission. The precision of the model may be adjusted in response to the bed situation during a surge to aid decisions on the use of limited resources. Such machine learning models offer great potential in itself. However, combining it with other machine learning models to predict mortality and readmission risks may create a more comprehensive clinical decision system.

Author Contributions: Conceptualization, Y.G., H.C.O., S.H.G. and H.C.S.L.; methodology, Y.G. and H.C.O.; software, Y.G.; validation, Y.G., S.S., H.C.O., S.H.G. and H.C.S.L.; formal analysis, Y.G.; investigation, Y.G.; resources, H.C.O., S.H.G. and H.C.S.L.; data curation, Y.G.; writing—original draft preparation, Y.G. and H.C.S.L.; writing—review and editing, Y.G., S.S., H.C.O., S.H.G. and H.C.S.L.; visualization, Y.G.; supervision, H.C.O.; project administration, H.C.O. and H.C.S.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Ethical review and approval were waived for this study due to the deidentified nature of the data.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this work are available upon reasonable request from the corresponding author, Y.G., and are subject to terms and conditions of Changi General Hospital.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

LOS	Length of Stay
SHAP	SHapley Additive xPlanations
CGH	Changi General Hospital
ED	Emergency Department

References

1. Singapore MOH. Waiting Times for Admissions at Public Hospitals. Available online: https://www.moh.gov.sg/news-highlights/details/waiting-times-for-admissions-at-public-hospitals (accessed on 1 January 2025).
2. Zhou, K.; Vidyarthi, A.; Matchar, D.; Cheung, Y.B.; Lam, S.W.; Ong, M. The relationship between workload and length of stay in Singapore. *Health Policy* **2018**, *122*, 769–774. [CrossRef] [PubMed]
3. Kulstad, E.B.; Sikka, R.; Sweis, R.T.; Kelley, K.M.; Rzechula, K.H. ED overcrowding is associated with an increased frequency of medication errors. *Am. J. Emerg. Med.* **2010**, *28*, 304–309. [CrossRef] [PubMed]
4. Trzeciak, S.; Rivers, E.P. Emergency department overcrowding in the United States: An emerging threat to patient safety and public health. *Emerg. Med. J.* **2003**, *20*, 402–405. [CrossRef] [PubMed]
5. Bacchi, S.; Gluck, S.; Tan, Y.; Chim, I.; Cheng, J.; Gilbert, T.; Menon, D.K.; Jannes, J.; Kleinig, T.; Koblar, S. Prediction of general medical admission length of stay with natural language processing and deep learning: A pilot study. *Intern. Emerg. Med.* **2020**, *15*, 989–995. [CrossRef]
6. Jaotombo, F.; Pauly, V.; Fond, G.; Orleans, V.; Auquier, P.; Ghattas, B.; Boyer, L. Machine-learning prediction for hospital length of stay using a French medico-administrative database. *J. Mark. Access Health Policy* **2022**, *11*, 2149318. [CrossRef]
7. Chrusciel, J.; Girardon, F.; Roquette, L.; Laplanche, D.; Duclos, A.; Sanchez, S. The prediction of hospital length of stay using unstructured data. *BMC Med. Inform. Decis. Mak.* **2021**, *21*, 351. [CrossRef]
8. NUHS+. Shorter Hospital Waiting Times with Artificial Intelligence. Available online: https://nuhsplus.edu.sg/article/shorter-hospital-waiting-times-with-artificial-intelligence# (accessed on 1 January 2025).
9. Ng, R.; Tan, K.B. Implementing an Individual-Centric Discharge Process across Singapore Public Hospitals. *Int. J. Environ. Res. Public Health* **2021**, *18*, 8700. [CrossRef]
10. Collins, G.S.; Moons, K.G.M.; Dhiman, P.; Riley, R.D.; Beam, A.L.; Van Calster, B.; Ghassemi, M.; Liu, X.; Reitsma, J.B.; van Smeden, M.; et al. TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ* **2024**, *385*, e078378. [CrossRef]
11. Collins, G.S.; Reitsma, J.B.; Altman, D.G.; Moons, K.G. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) the TRIPOD statement. *Circulation* **2015**, *131*, 211–219. [CrossRef]
12. Turgeman, L.; May, J.H.; Sciulli, R. Insights from a machine learning model for predicting the hospital Length of Stay (LOS) at the time of admission. *Expert Syst. Appl.* **2017**, *78*, 376–385. [CrossRef]
13. Earnest, A.; Chen, M.I.C.; Seow, E. Exploring if day and time of admission is associated with average length of stay among inpatients from a tertiary hospital in Singapore: An analytic study based on routine admission data. *BMC Health Serv. Res.* **2006**, *6*, 6. [CrossRef] [PubMed]
14. Hachesu, P.R.; Ahmadi, M.; Alizadeh, S.; Sadoughi, F. Use of data mining techniques to determine and predict length of stay of cardiac patients. *Healthc. Inform. Res.* **2013**, *19*, 121–129. [CrossRef] [PubMed]
15. Gentimis, T.; Alnaser, A.J.; Durante, A.; Cook, K.; Steele, R. Predicting Hospital Length of Stay Using Neural Networks. *Int. J. Big Data Intell.* **2019**, *6*, 1. [CrossRef]

16. Prokhorenkova, L.; Gusev, G.; Vorobev, A.; Dorogush, A.V.; Gulin, A. CatBoost: Unbiased boosting with categorical features. *Adv. Neural Inf. Process. Syst.* **2018**, 6639–6649.
17. Zhang, C.; Chen, X.; Wang, S.; Hu, J.; Wang, C.; Liu, X. Using CatBoost algorithm to identify middle-aged and elderly depression, national health and nutrition examination survey 2011–2018. *Psychiatry Res.* **2021**, *306*, 114261. [CrossRef]
18. Lundberg, S.M.; Lee, S.-I. A unified approach to interpreting model predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; Curran Associates Inc.: Red Hook, NY, USA, 2017; pp. 4768–4777.
19. Bergstra, J.; Yamins, D.; Cox, D.D. Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. In Proceedings of the 30th International Conference on International Conference on Machine Learning—Volume 28, Atlanta, GA, USA, 17–19 June 2013; pp. I–115–I–123.
20. Patel, D.; Cheetirala, S.N.; Raut, G.; Tamegue, J.; Kia, A.; Glicksberg, B.; Freeman, R.; Levin, M.A.; Timsina, P.; Klang, E. Predicting Adult Hospital Admission from Emergency Department Using Machine Learning: An Inclusive Gradient Boosting Model. *J. Clin. Med.* **2022**, *11*, 6888. [CrossRef]
21. Feretzakis, G.; Karlis, G.; Loupelis, E.; Kalles, D.; Chatzikyriakou, R.; Trakas, N.; Karakou, E.; Sakagianni, A.; Tzelves, L.; Petropoulou, S.; et al. Using Machine Learning Techniques to Predict Hospital Admission at the Emergency Department. *J. Crit. Care Med.* **2022**, *8*, 107–116. [CrossRef]
22. Alabbad, D.A.; Almuhaideb, A.M.; Alsunaidi, S.J.; Alqudaihi, K.S.; Alamoudi, F.A.; Alhobaishi, M.K.; Alaqeel, N.A.; Alshahrani, M.S. Machine learning model for predicting the length of stay in the intensive care unit for Covid-19 patients in the eastern province of Saudi Arabia. *Inform. Med. Unlocked* **2022**, *30*, 100937. [CrossRef]
23. Shahid Ansari, M.; Jain, D.; Harikumar, H.; Rana, S.; Gupta, S.; Budhiraja, S.; Venkatesh, S. Identification of predictors and model for predicting prolonged length of stay in dengue patients. *Health Care Manag. Sci.* **2021**, *24*, 786–798. [CrossRef]
24. Lee, S.; Reddy Mudireddy, A.; Kumar Pasupula, D.; Adhaduk, M.; Barsotti, E.J.; Sonka, M.; Statz, G.M.; Bullis, T.; Johnston, S.L.; Evans, A.Z.; et al. Novel Machine Learning Approach to Predict and Personalize Length of Stay for Patients Admitted with Syncope from the Emergency Department. *J. Pers. Med.* **2022**, *13*, 7. [CrossRef]
25. Dogu, E.; Albayrak, Y.E.; Tuncay, E. Length of hospital stay prediction with an integrated approach of statistical-based fuzzy cognitive maps and artificial neural networks. *Med. Biol. Eng. Comput.* **2021**, *59*, 483–496. [CrossRef] [PubMed]
26. Masum, S.; Hopgood, A.; Stefan, S.; Flashman, K.; Khan, J. Data analytics and artificial intelligence in predicting length of stay, readmission, and mortality: A population-based study of surgical management of colorectal cancer. *Discov. Oncol.* **2022**, *13*, 11. [CrossRef] [PubMed]
27. Alharbi, S.; Alasmari, A.; Hanafy, E.; Ellawindy, A. Reduction of hospital bed cost for inpatient overstay through optimisation of patient flow. *BMJ Open Qual.* **2023**, *12*, e002142. [CrossRef] [PubMed]
28. Ang, Y.H.; Ginting, M.L.; Wong, C.H.; Tew, C.W.; Liu, C.; Sivapragasam, N.R.; Matchar, D.B. From Hospital to Home: Impact of Transitional Care on Cost, Hospitalisation and Mortality. *Ann. Acad. Med. Singap.* **2019**, *48*, 333–337. [CrossRef]
29. Emes, M.; Smith, S.; Ward, S.; Smith, A. Improving the patient discharge process: Implementing actions derived from a soft systems methodology study. *Health Syst.* **2018**, *8*, 117–133. [CrossRef]
30. Hernandez, N.; John, D.; Mitchell, J. A reimagined discharge lounge as a way to an efficient discharge process. *BMJ Qual. Improv. Rep.* **2014**, *3*, u204930-w202080. [CrossRef]
31. Hammer, C.; DePrez, B.; White, J.; Lewis, L.; Straughen, S.; Buchheit, R. Enhancing Hospital-Wide Patient Flow to Reduce Emergency Department Crowding and Boarding. *J. Emerg. Nurs.* **2022**, *48*, 603–609. [CrossRef]
32. Tan, H.S.; Liu, N.; Tan, C.W.; Sia, A.T.H.; Sng, B.L. Developing the BreakThrough Pain Risk Score: An interpretable machine-learning-based risk score to predict breakthrough pain with labour epidural analgesia. *Can. J. Anesth. J. Can. D'anesthésie* **2022**, *69*, 1315–1317. [CrossRef]
33. Tan, T.F.; Dai, P.; Zhang, X.; Jin, L.; Poh, S.; Hong, D.; Lim, J.; Lim, G.; Teo, Z.L.; Liu, N.; et al. Explainable artificial intelligence in ophthalmology. *Curr. Opin. Ophthalmol.* **2023**, *34*, 422–430. [CrossRef]
34. Yu, J.Y.; Feng, X.; Liu, N.; Yoon, S.Y.; Ong, M.E.H.; Ng, Y.Y.; Cha, W.C. An External Validation Study of the Score for Emergency Risk Prediction (SERP), an Interpretable Machine Learning-Based Triage Score for the Emergency Department. *Sci. Rep.* **2022**, *12*, 17466. [CrossRef]
35. Jiang, L.Y.; Liu, X.C.; Nejatian, N.P.; Nasir-Moin, M.; Wang, D.; Abidin, A.; Eaton, K.; Riina, H.A.; Laufer, I.; Punjabi, P.; et al. Health system-scale language models are all-purpose prediction engines. *Nature* **2023**, *619*, 357–362. [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.