Original Article

Healthc Inform Res. 2025 January;31(1):23-36.
https://doi.org/10.4258/hir.2025.31.1.23
pISSN 2093-3681 • eISSN 2093-369X

HIR
Healthcare Informatics Research

Triage Data-Driven Prediction Models for Hospital Admission of Emergency Department Patients: A Systematic Review

Hyun A Shin[1], **Hyeonji Kang**[1], **Mona Choi**[2]

[1]College of Nursing, Yonsei University, Seoul, Korea
[2]College of Nursing, Mo-Im Kim Nursing Research Institute, Yonsei University, Seoul, Korea

Objectives: Emergency department (ED) overcrowding significantly impacts healthcare efficiency, safety, and resource management. Predictive models that utilize triage information can streamline the admission process. This review evaluates existing hospital admission prediction models that have been developed or validated using triage data for adult ED patients. **Methods:** A systematic search of PubMed, Embase, CINAHL, Web of Science, and the Cochrane Library was conducted. Studies were selected if they developed or validated predictive models for hospital admission using triage data from adult ED patients. Data extraction adhered to the CHARMS (Checklist for Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies), and the risk of bias was evaluated using PROBAST (Prediction model Risk of Bias Assessment Tool). **Results:** Twenty studies met the inclusion criteria, employing logistic regression and machine learning techniques. Logistic regression was noted for its traditional use and clinical interpretability, whereas machine learning provided enhanced flexibility and potential for better predictive accuracy. Common predictors included patient demographics, triage category, vital signs, and mode of arrival. The area under the curve values for model performance ranged from 0.80 to 0.89, demonstrating strong discriminatory ability. However, external validation was limited, and there was variability in outcome definitions and model generalizability. **Conclusions:** Predictive models based on triage data show promise in supporting ED operations by facilitating early predictions of hospital admissions, which could help decrease boarding times and enhance patient flow. Further research is necessary to validate these models in various settings to confirm their applicability and reliability.

Keywords: Admission, Hospitalization, Prognosis, Emergencies, Triage

Submitted: October 9, 2024
Revised: December 13, 2024
Accepted: December 14, 2024

Corresponding Author
Mona Choi
College of Nursing, Mo-Im Kim Nursing Research Institute, Yonsei University, 50-1, Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea. Tel: +82-22283341, E-mail: monachoi@yuhs.ac (https://orcid.org/0000-0003-4694-0359)

I. Introduction

Emergency departments (EDs) play a critical role in delivering timely and appropriate emergency medical services to patients with acute illnesses of varying severity, often under unpredictable conditions. However, ED overcrowding is an increasing global problem, resulting in extended waiting times, inefficient resource utilization, and compromised patient safety [1]. To tackle these issues, healthcare systems need effective resource allocation strategies and informatics-driven solutions that facilitate prompt clinical decision-making. Specifically, predictive models that utilize real-time data and advanced algorithms can greatly improve the operational efficiency of EDs by accurately determining which patients need hospital admission [2,3]. In the current system, the admission process starts only after a patient's ED visit is complete, causing prolonged waiting times and unnecessary overcrowding as patients wait for administrative staff to process their admission and assign hospital beds. By implementing early prediction models at the time of patient arrival, the administrative process could be initiated simultaneously with the ED visit, thus reducing waiting times and preventing further ED crowding if the patient requires admission [4]. This approach would streamline patient flow and improve overall ED efficiency [2].

Predictive models have been developed and implemented to address these challenges, with the goal of forecasting which ED patients will require hospital admission [5,6]. By accurately predicting the need for admission early in the triage process, these models can reduce boarding times, improve resource allocation, and improve the overall efficiency of the ED [4].

Triage information, which includes age, sex, vital signs, and mode of arrival, is collected during the initial patient assessment in the ED. These data are crucial for quickly evaluating patient conditions and can be used to develop predictive models for hospital admission [7-9]. However, the systematic development of predictive models based solely on triage data is limited. This limitation is due to the reliance on detailed clinical data, such as laboratory or imaging results, which are not available at triage. Additionally, variability in triage systems and practices across different institutions hinders the generalizability of these models [10-12]. Previous reviews have highlighted usability challenges in clinical implementation [10] and the potential of machine learning (ML) techniques to improve prediction accuracy [12], laying the groundwork for further research into models that leverage triage-only data. Models based on triage data are broadly applicable across EDs, aiding in the evaluation of patient conditions and informing admission decisions. Advances in artificial intelligence and ML techniques have further improved predictive accuracy, enabling the development of sophisticated triage-based models [13].

This systematic review aims to identify and assess studies that have developed or validated hospital admission prediction models for adult ED patients using triage data. It specifically targets adult patients, acknowledging the unique clinical characteristics and care pathways that differ from pediatric populations. By examining the features and limitations of these models, this review intends to offer insights into enhancing ED resource management and the quality of patient care, as well as propose directions for future research.

II. Methods

1. Study Design

This study is a systematic review of studies on hospital admission prediction models using ED triage data and was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [14].

2. Eligibility Criteria

This study included all studies that developed or validated predictive models for the hospital admission of ED patients. It included data collected during the triage stage and utilized both retrospective and prospective study designs.

Inclusion criteria:

1) Participants: Adult patients aged 16 years or older presenting to the ED. Studies involving mixed-age populations were also included, provided their findings were stratified by age groups or were relevant to adults.
2) Intervention: Predictive models developed using data collected during the triage stage.
3) Outcomes: Hospital admission, defined as including both general ward and intensive care unit admission.
4) Only peer-reviewed journal articles were included.

Exclusion criteria:

1) Studies that focused exclusively on pediatric populations or specific diseases or symptoms.
2) Studies that were published in languages other than English or Korean.

3. Data Sources and Search Strategy

The literature search was conducted on October 20, 2023, utilizing five databases: PubMed, Embase, CINAHL, Web of Science, and the Cochrane Library, with no restrictions on publication dates. Additionally, manual checks of the references from the retrieved studies were performed to identify additional studies that met the inclusion criteria.

Search terms were constructed using MeSH terms, and queries were adapted to the specific features of each database. The search terms were combined using the operators OR and AND. The primary MeSH terms employed were “emergencies,” “triage,” and “prognosis.” The final search strategy was formulated as follows: (emergency OR emergencies OR emergence OR emergent OR emergencies[MeSH Terms]) AND (triage OR triages OR triaging OR triaged OR

https://doi.org/10.4258/hir.2025.31.1.23

triage[MeSH Terms]) AND (prognosis OR prediction OR predictive OR predicting OR predict OR prognosis[MeSH Terms]) AND (model OR modeling OR tool).

4. Study Selection

The references retrieved were organized using reference management software (EndNote 20.6; Clarivate, Philadelphia, PA, USA). After reviewing the titles, duplicates were removed. Titles and abstracts were further screened, and studies that were clearly irrelevant to the research question were excluded. The study selection process was conducted independently by two researchers based on the core research question, and inclusion and exclusion criteria. In cases of disagreement, a third researcher facilitated consensus to finalize the selection of studies.

5. Data Extraction

Data extraction from the selected studies was independently conducted by two researchers and subsequently verified. In cases of disagreement, a consensus was reached through discussion with a third researcher. The data extraction form was based on the Checklist for Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies (CHARMS) [15]. The extracted data encompassed authors, publication year, country, setting, type of prediction modeling, data source, study design, study period, population, sample size, outcome variable, candidate predictors, important variables, handling of missing data, algorithms used, validation methods, and the performance of the final model.

6. Risk of Bias and Applicability Assessment

The risk of bias and applicability of each study were independently assessed by two researchers using the Prediction model Risk of Bias Assessment Tool (PROBAST) checklist [16]. Any disagreements were resolved through consultation with a third researcher. PROBAST systematically evaluates the risk of bias (ROB) by examining four critical domains—participants, predictors, outcomes, and analysis—through 20 targeted questions that identify methodological biases. Applicability is determined by evaluating how well the study's population, predictors, and outcomes align with the research question, thus assessing the relevance and generalizability of the study findings.

III. Results

1. Study Selection

A comprehensive search across five databases identified a total of 3,690 records. After removing duplicates, 2,219 unique records were left. These underwent a screening process based on their titles and abstracts, which led to the exclusion of 1,819 records that did not relate to the research question. The abstracts of the remaining 400 records were further reviewed, and 76 studies that met the inclusion criteria were selected. Two independent reviewers then conducted a full-text assessment of these studies, resulting in the exclusion of 56 studies for various reasons, including the use of factor analysis instead of predictive modeling, a focus on non-adult populations, or the absence of relevant prediction outcomes. As a result, 20 studies were included in the final systematic review. The search process, based on the PRISMA 2020 flow diagram, is illustrated in Figure 1.

2. Risk of Bias and Applicability

Using the PROBAST tool, most studies were assessed as having a low risk of bias in participant selection and outcome measurement. However, they encountered challenges with missing data and predictor selection. As shown in Table 1, although most studies conducted internal validation (e.g., cross-validation), only three studies utilized external validation methods [5,6,17], underscoring the need for enhanced generalizability of the models.

3. Study Characteristics

The characteristics of the studies included in this review are summarized in Table 2. This systematic review included 20 studies [2-6,8-9,17-29], with research settings distributed across various countries and regions, including the USA, Australia, the UK, Spain, Singapore, Taiwan, and Austria. The methodologies predominantly involved retrospective analyses, although several studies also featured prospective validation [5,6,17]. While most studies focused exclusively on model development, a few integrated both development and validation processes [17,18]. The data sources were diverse, ranging from ED databases and hospital information systems to national health surveys, such as the National Hospital Ambulatory Medical Care Survey. These studies were conducted in both single-center and multicenter environments, with some specifically employing electronic health intelligence systems (eHINTS) or extensive datasets like the Medical Information Mart for Intensive Care – Emergency Department (MIMIC-IV-ED) [8,19]. The duration of the

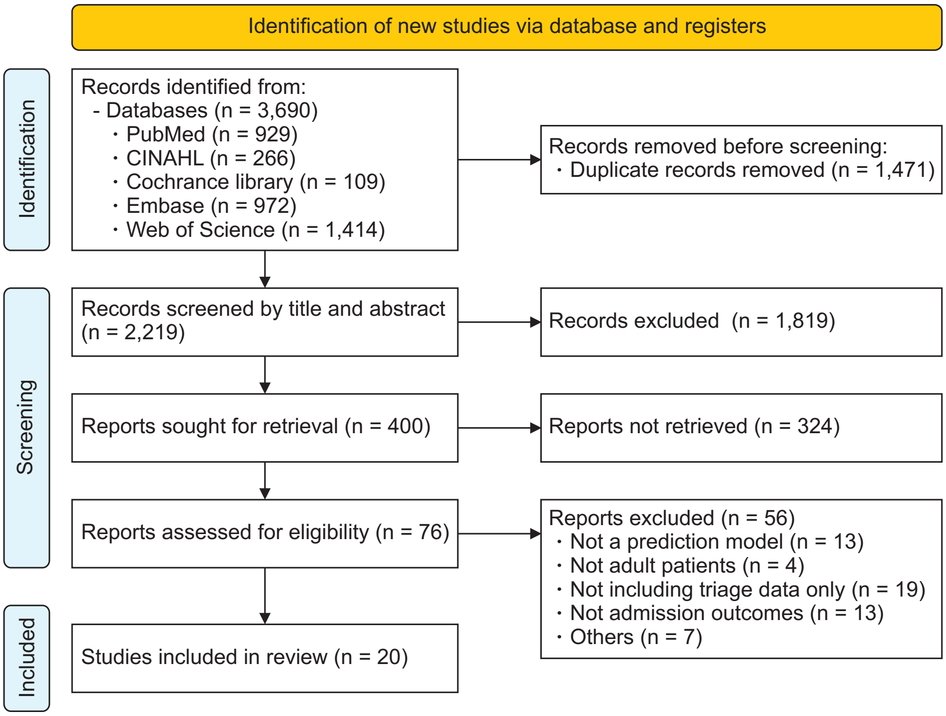

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses flowchart of study selection process.

studies varied from 1 to 9 years, tailored to the particular design and scope of each study. The primary focus was on adult patients, though some studies included participants of all ages. Sample sizes varied widely, from 894 to over 3 million patients or events, reflecting the diverse scopes and settings of the studies.

Table 3 summarizes the primary outcome assessed across the studies, which was hospital admission, encompassing both ward and intensive care unit (ICU) admissions. Other outcomes evaluated included mortality, critical outcomes, length of stay, and readmission [5,19-22]. Admission rates varied widely, ranging from 11.0% [2] to 47.3% [19], depending on the definition of admission, study setting, and patient population.

The studies reviewed incorporated a variety of candidate predictors, including patient demographics (such as age, sex/gender, and ethnicity), triage details (such as triage category, vital signs, chief complaints, and mode of arrival), medical history (including previous admissions and comorbidities), and administrative information (such as date and time of attendance, shift time, and insurance status). Key predictive variables that were often identified included age, triage category, mode of arrival (for example, ambulance), and vital signs like body temperature, heart rate, and respiratory rate.

Regarding data handling, several studies addressed the issue of missing data either by excluding incomplete cases or by employing imputation techniques. For instance, Xie et al. [19] utilized median imputation, whereas Cameron et al. [23] applied a combination of exclusion, removal, and imputation strategies. However, despite being retrospective in nature, four studies failed to provide details on how they managed missing data [9,18,24,25].

Various ML algorithms were utilized in the studies reviewed, with logistic regression being the predominant method. More recent research has incorporated advanced algorithms, including gradient boosting machines (GBM), random forest, and neural networks, as well as more complex models like long short-term memory [2,19,21,22,26].

Validation methods varied across the studies. Many employed cross-validation techniques, such as bootstrap cross-validation [3,20,23] and k-fold cross-validation [2,4,26]. Several studies also conducted external validation [5,6,17], which enhances the generalizability of their findings.

4. Model Performance of the Final Models

The study compared the performance of various predictive models designed to forecast hospital admissions using triage data from the ED, as detailed in Table 4. The evaluation concentrated on three key aspects: discrimination, calibration, and classification. Each aspect was crucial for assessing the primary outcome, which was hospital admission.

Discrimination, primarily assessed through the area under the receiver operating characteristic curve (AUC), indicated how well each model could differentiate between patients who required hospital admission and those who did not. The GBM model developed by Cusido et al. [2] exhibited

Table 1. PROBAST results

Study, year	ROB				Applicability			Overall	
	Partici-pants	Predictors	Outcome	Analysis	Partici-pants	Predictors	Outcome	ROB	Applica-bility
Cameron et al. [23], 2015	+	+	+	–	+	–	–	–	–
Cusido et al. [2], 2022	+	+	+	–	+	–	+	–	–
Dinh et al. [3], 2016	+	+	+	+	+	+	–	+	–
Ebker-White et al. [6], 2018	+	+	+	+	+	+	+	+	+
Ebker-White et al. [17], 2018	+	+	+	+	+	+	–	+	–
Graham et al. [27], 2018	+	+	+	–	–	+	+	–	–
Handly et al. [24], 2015	+	+	+	–	+	+	+	–	+
Jones et al. [5], 2019	+	+	+	+	+	–	+	+	–
Lee et al. [28], 2021	+	+	+	–	–	+	+	–	–
Levin et al. [20], 2018	+	+	+	–	–	+	–	–	–
Parker et al. [8], 2019	+	+	+	–	+	+	+	–	+
Peck et al. [25], 2012	?	+	+	–	–	+	+	–	–
Peck et al. [18], 2013	?	–	+	–	?	+	+	–	?
Raita et al. [21], 2019	+	+	+	–	+	+	–	–	–
Rendell et al. [26], 2019	+	+	+	–	+	+	+	–	+
Sun et al. [9], 2011	+	+	+	–	+	+	+	–	+
Tschoellitsch et al. [22], 2023	+	+	+	–	+	+	+	–	+
Xie et al. [19], 2022	+	+	+	+	+	+	+	+	+
Zhang et al. [4], 2017	+	+	+	+	+	+	–	+	–
Zlotnik et al. [29], 2016	+	+	+	+	+	+	+	+	+

PROBAST: Prediction model Risk of Bias Assessment Tool, ROB: risk of bias.
+ indicates low ROB/low concern regarding applicability; – indicates high ROB/high concern regarding applicability; and ? indicates unclear ROB/unclear concern regarding applicability.

the highest discrimination ability, achieving an AUC of 0.891 (95% CI, 0.890–0.892). This was closely followed by the Glasgow Admission Prediction Score (GAPS) model by Cameron et al. [23], which recorded an AUC of 0.877. The Sydney Triage to Admission Risk Tool (START), introduced by Dinh et al. [3], demonstrated an AUC of 0.820. Although slightly lower than the previous models, it still showed robust discrimination performance.

In terms of calibration, the Hosmer-Lemeshow goodness of fit test was commonly used to evaluate the alignment between predicted probabilities and actual outcomes. The GAPS model [23], with a p-value of 0.524, demonstrated strong calibration, indicating that its predictions closely matched observed admissions. In contrast, the START model [3] exhibited poor calibration ($p<0.001$), suggesting that despite reasonable discrimination, there was a significant mismatch between its predictions and the actual outcomes.

When assessing classification performance, the accuracy metric exhibited significant variations across the models. Cusido et al. [2] achieved the highest accuracy (89.8%), while Cameron et al. [23] followed with 80.3%. The positive predictive value and negative predictive value were evaluated in seven studies [3,6,8,9,21,22,28], indicating the models' effectiveness in accurately predicting true positives and true negatives. Additional metrics, such as the F1-score [22] and net reclassification improvement [21,24], provided further insights into the classification capabilities and enhancements of these models.

IV. Discussion

This systematic review identified a total of 20 studies that focused on developing predictive models for hospital admissions in ED settings. Of these, 16 models were newly developed, including START [3] and GAPS [23]. Additionally, four studies were dedicated to expanding or externally

Table 2. Characteristics of the included studies

Study, year	Country	Setting	Type of prediction modelling	Source of data	Study design	Study period	Population	Sample size (patients or events)
Cameron et al. [23], 2015	UK	Multicenter	Development only	6 units in 3 hospitals	Retrospective	Mar 21, 2010–Mar 20, 2012	Aged ≥16 yr	322,846 (191,653)
Cusido et al. [2], 2022	Spain	Multicenter	Development only	60 EDs	Retrospective	Jan 1, 2018–Dec 31, 2018	All ages	3,189,204 (1,805,096)
Dinh et al. [3], 2016	Australia	Multicenter	Development only	EDDC Registry	Retrospective	2013–2014	Aged ≥16 yr	1,721,294
Ebker-White et al. [6], 2018	Australia	Multicenter	Validation only	2 EDs	Prospective	Nov 2016–Jun 2017	Aged ≥16 yr	894
Ebker-White et al. [17], 2018	Australia	Multicenter	Development & validation	2 EDs	Prospective	Nov 2016–Jun 2017	Aged ≥16 yr	894
Graham et al. [27], 2018	UK	Multicenter	Development only	2 EDs	Retrospective	During the 2015	All ages	107,545
Handly et al. [24], 2015	USA	Single-center	Development only	HIS	Retrospective	Jul 1, 2007–Dec 31, 2010	Aged ≥18 yr	159,200
Jones et al. [5], 2019	UK	Multicenter	Validation only	2 EDs	Prospective	Feb 2016–May 2016	Aged ≥16 yr	1,420
Lee et al. [28], 2021	Taiwan	Single-center	Development only	Tertiary hospital	Retrospective	Jan 1, 2015–Dec 31, 2019	All ages only triage level 3	282,971
Levin et al. [20], 2018	USA	Multicenter	Development only	2EDs Urban ED & Community ED	Retrospective	Aug 2014–Oct 2015 (urban) Jun 2013–Oct 2015 (community)	Aged ≥18 yr	172,726
Parker et al. [8], 2019	Singapore	Single-center	Development only	eHINTS of SGH	Retrospective	Jan 1, 2005–Dec 31, 2014	Aged ≥21 yr	1,232,016

Continued on the next page.

https://doi.org/10.4258/hir.2025.31.1.23

Table 2. Continued

Study, year	Country	Setting	Type of prediction modelling	Source of data	Study design	Study period	Population	Sample size (patients or events)
Peck et al. [25], 2012	USA	Single-center	Development only	13 bed ED	Retrospective, Prospective	Jan 1, 2010–Nov 26, 2010	All ages	6,961 (Retrosp) 767 (Prosp)
Peck et al. [18], 2013	USA	Multicenter	Development & validation	4 hospitals	Retrospective, Prospective	Varies by hospital	All ages	28,865 (Retrosp) 910 (Prosp)
Raita et al. [21], 2019	USA	Multicenter	Development only	NHAMCS ED data	Retrospective	2007–2015	Aged ≥18 yr	135,470
Rendell et al. [26], 2019	Australia	Multicenter	Development only (update)	State wide ED data	Retrospective	2013–2014	Aged ≥16 yr	1,721,294
Sun et al. [9], 2011	Singapore	Single-center	Development only	EDWeb	Retrospective	Jan 2007–Dec 2008	All ages	317,581 (207,069)
Tschoellitsch et al. [22], 2023	Austria	Single-center	Development only	HIS	Retrospective	Dec 1, 2015–Aug 31, 2020	Aged ≥18 yr	77,477 (58,323)
Xie et al. [19], 2022	USA	Multicenter	Development only	MIMIC-IV-ED database	Retrospective	2011–2019	Aged ≥18 yr	441,437
Zhang et al. [4], 2017	USA	Multicenter	Development only	NHAMCS ED data	Retrospective	2012–2013	All ages	47,200
Zlotnik et al. [29], 2016	Spain	Single-center	Development only	HIS	Retrospective	2011–2012	All ages	255,668 (153,970)

ED: emergency department, EDDC: Emergency Department Data Collection, HIS: hospital information system, NHAMCS: National Hospital Ambulatory Medical Care Survey, eHINTS: Electronic Health Intelligence System, SGH: Singapore General Hospital, MIMIC-IV-ED: Medical Information Mart for Intensive Care - Emergency Department.

Table 3. Summary of the included studies

Study, year	Outcome (rate)	Candidate predictors	Important variables	Missing data	Algorithms	Validation
Cameron et al. [23], 2015	Admission; including death	Age, sex, transport, time, referral source, triage category, NEWS, lives alone, previous admission	Age, NEWS, triage category, GP referral, arrive by ambulance, admission within 1 year	Excluded, removed, imputation	LR	Bootstrap cross validation
Cusido et al. [2], 2022	Admission (11.0%)	Accumulated visits, age, gender, CCS, CCS frequency, triage	Accumulated visits, age, gender, CCS, CCS frequency, triage	Excluded	GBM	2-fold cross-validation
Dinh et al. [3], 2016	Admission; including short stay & transfer out (40.7%)	Age, gender, referral source, mode of arrival, hospital facility, triage category, presenting problem, mode of separation	Age, arrival by ambulance, triage category, previous admission, presenting problem	Excluded	LR, LASSO regression	Bootstrap cross validation
Ebker-White et al. [6], 2018	Admission; including short stay (36.0%)	Age, ambulance arrival, triage category, previous admission, hour of presentation, presenting problem, frailty, GP referral, ED overcrowding, comorbidities	-	-	LR	External validation
Ebker-White et al. [17], 2018	Longer stay admission (18.1%)	Age, ambulance arrival, triage category, previous admission, hour of presentation, presenting problem, frailty, decreased mobility, multiple comorbidities, GP referral	Age, ambulance arrival, triage category, previous admission, hour of presentation, presenting problem, frailty, multiple comorbidities	-	LR	External validation
Graham et al. [27], 2018	Admission (24.0%)	Hospital site, date and time of attendance, age, gender, arrival model, care group, triage category, previous admission within the last week, month, or year	Age, arrival mode, triage category, care group, admission in past year	Excluded	LR, DT, GBM	-
Handly et al. [24], 2015	Admission	Age, sex, race, time, day of arrival, initial ESI level, coded chief complaint	Coded chief complaint	-	Levenberg–Marquardt neural network learning	Temporal validation

Continued on the next page.

Table 3. Continued

Study, year	Outcome (rate)	Candidate predictors	Important variables	Missing data	Algorithms	Validation
Jones et al. [5], 2019	Admission (39.6%), LOS, mortality, readmission	Age, NEWS, triage category, GP referral, arrived by ambulance, admission within 1 year	-	-		External validation
Lee et al. [28], 2021	Admission (15.8%)	Age, sex, vital signs (BT, HR, RR, SBP, DBP, MAP), medical history, chief complaints	Age, sex, heart rate, MAP, medical history, chief complaint	Excluded	NN, ML	-
Levin et al. [20], 2018	Admission (26.0%, 22.3%), critical care, emergency procedure	Age, sex, arrival mode, vital signs (BT, PR, RR, SBP, SpO_2), primary chief complaint, relevant medical history	Vital signs, chief complaint, active medical history	Excluded	RF	Bootstrap cross validation
Parker et al. [8], 2019	Admission (38.7%)	Age, gender, ethnicity, postal code, day of week, shift time, mode of arrival, time of year, triage category, fever status, number of ED visits within the previous year	Age group, race, postal code, day of week, time of day, triage category, mode of arrival, fever status	Excluded	Stepwise LR	-
Peck et al. [25], 2012	Admission	Age, primary complaint, ED provider, designation, arrival mode, ESI level	Age, primary complaint, bed type designation, arrival mode	-	Naïve Bayesian model logit-linear regression	Temporal validation & face validity
Peck et al. [18], 2013	Admission	Age, primary complaint, ED provider, designation, arrival mode, ESI level	Age, primary complaint, arrival mode, ESI level	-	LR	Temporal validation
Raita et al. [21], 2019	Admission; including transfer (16.2%), critical care	Age, sex, mode of arrival, vital signs (BT, PR, SBP, DBP, RR, SpO_2), chief complaints, comorbidities	Age, ambulance use, SBP, DBP, PR, RR, BT, CHF comorbidity	Excluded	Lasso, RF GBDT, DNN	-

Continued on the next page.

Table 3. Continued

Study, year	Outcome (rate)	Candidate predictors	Important variables	Missing data	Algorithms	Validation
Rendell et al. [26], 2019	Admission; including short stay & transfer	Age, arrival by ambulance, triage category, previous admission within the last 30 days, hour of arrival, presenting problem	Triage category, arrival by ambulance, age, previous admission	Excluded (2.95%)	BN, DT, LR, NB, MLP, NN	10-fold cross-validation
Sun et al. [9], 2011	Ward admission (30.2%)	Age, sex, ethnic, ED visit or hospital admission in the preceding 3 months, arrival mode, patient acuity category, coexisting chronic diseases (DM, HTN, dyslipidemia)	Age, patient acuity category, arrival mode	-	Stepwise LR	-
Tschoellitsch et al. [22], 2023	Ward admission ICU admission mortality	MTS triage, SBP, DBP, HR, SpO_2, BT, age, gender, serum glucose, inhaled oxygen therapy, VAS, AVPU, expedited treatment recommendation	Age, chief complaint according to the MTS, BT	Removed, imputation	LR, RF, NN, GB, DT, KNN	-
Xie et al. [19], 2022	Admission (47.3%) critical outcomes, ED revisit	Age, gender, ESI, chief complaints, V/S, pain scale, comorbidities	Age, ESI, SBP, HR, DBP, BT, pain scale, SpO_2, RR, hospitalizations in the past year	Median imputation	LR, RF, GB, AutoScore, MLP, Med2Vec, LSTM	-
Zhang et al. [4], 2017	Admission; including transfer (13.4%)	Age, sex, ethnicity, triage level, pain scale, initial vital signs, arrival mode, comorbidities, residence type, source of payment reasons for visit (NLP)	Age, arrival by ambulance, triage level, and initial vital signs (BP, RR)	Excluded, imputation	LR, MLNN	10-fold cross-validation
Zlotnik et al. [29], 2016	Admission (13.6%)	Age, sex, insurance status, visit source, visit cause, ambulance arrival, triage score, chief complaint, previous ED visits	Triage level, chief complaints, previous ED visits, visit source, ambulance arrival	Excluded	LR, ANN	-

NEWS: National Early Warning Score, GP: general practitioner, LR: logistic regression, GBM: gradient boosting machine, LASSO: least absolute shrinkage and selection operator, DT: decision tree, NN: neural network, ML: machine learning, RF: random forest, ED: emergency department, BT: body temperature, HR: heart rate, RR: respiratory rate, SBP: systolic blood pressure, DBP: diastolic blood pressure, MAP: mean arterial pressure, SpO_2: oxygen saturation, ESI: Emergency Severity Index, BN: Bayesian network, NB: naïve Bayes, MLP: multilayer perceptron, DNN: deep neural network, GBDT: gradient boosted decision trees, DM: diabetes mellitus, HTN: hypertension, VAS: visual analog scale, AVPU: alert, verbal, pain, unresponsive scale, KNN: k-nearest neighbors, LSTM: long short-term memory, MLNN: multilayer neural network, ANN: artificial neural network, NLP: natural language processing.

https://doi.org/10.4258/hir.2025.31.1.23

Table 4. Model performances of the final model

Study, year	Final model	Discrimination (AUC scores)	Calibration	Classification
Cameron et al. [23], 2015	GAPS	0.877 (95% CI, 0.875- 0.879)	HL GOF (p = 0.524)	Accuracy: 80.3%
Cusido et al. [2], 2022	GBM model	0.891 (95% CI, 0.890-0.892)	-	Accuracy: 89.8%
Dinh et al. [3], 2016	START	0.820 (95% CI, 0.810-0.820) Sensitivity: 88.0% Specificity: 67.0%	HL GOF (p < 0.001)	PPV: 86.8%, NPV: 64.3% High risk score ranges (>20)
Ebker-White et al. [6], 2018	START	0.800 (95% CI, 0.770–0.830) Sensitivity: 78.5% Specificity: 65.0%	HL GOF (p = 0.09)	Accuracy: 70.2% PPV: 56.5% NPV: 84.2%
Ebker-White et al. [17], 2018	Extended START	0.840 (95% CI, 0.810-0.880)	HL GOF (p = 0.09)	-
Graham et al. [27], 2018	GBM model	0.859	-	Accuracy: 80.3%
Handly et al. [24], 2015	Neural network-based model with CCC data	0.860 (95% CI, 0.858–0.862) IDI: 0.060 (95% CI, 0.058–0.061)	-	NRI: 0.156 (95% CI, 0.148–0.163)
Jones et al. [5], 2019	GAPS			
Lee et al. [28], 2021	NN & ML model	0.817 (95% CI 0.820–0.821) Sensitivity: 67.2% Specificity: 78.1%	Youden's index: 0.552	PPV: 36.6% NPV: 92.7%
Levin et al. [20], 2018	E-triage	0.820–0.840	-	-
Parker et al. [8], 2019	Novel prediction model	0.825 (95% CI, 0.824–0.827) Sensitivity: 77.5% Specificity: 74.8%	Calibration plot	PPV: 83.0% NPV: 67.7%
Peck et al. [25], 2012	Logit-linear regression	0.887 R^2: 0.583	-	-
Peck et al. [18], 2013	LR model	0.800–0.890	HL GOF (p > 0.01)	-
Raita et al. [21], 2019	DNN model	0.820 (95% CI, 0.820–0.830) Sensitivity: 79.0% Specificity: 71.0%	-	NRI: 0.68 (p < 0.001) PPV: 35.0% NPV: 95.0%
Rendell et al. [26], 2019	START 2	0.827 (95% CI, ±0.0006)	-	Accuracy: 75.2%
Sun et al. [9], 2011	LR model	0.849 (95% CI, 0.847–0.851) Specificity: 96.8% Sensitivity: 33.4%	HL GOF (p > 0.05)	PPV: 81.6% NPV: 71.8%
Tschoellitsch et al. [22], 2023	NN model	0.842	-	F1-score: 0.706 PPV: 64.7% NPV: 84.9%
Xie et al. [19], 2022	GB model	0.819 (95% CI, 0.817–0.822)	-	-
Zhang et al. [4], 2017	LR model 3	0.846 (95% CI, 0.839–0.853)	-	-
Zlotnik et al. [29], 2016	ANN model	0.857 (95% CI, 0.854–0.861)	HL GOF (χ^2: 17.28) Calibration plot	-

GAPS: Glasgow Admission Prediction Score, GBM: gradient boosting machine, HL GOF: Hosmer-Lemeshow goodness of fit, LR: logistic regression, PPV: positive predictive value, NPV: negative predictive value, START: Sydney Triage to Admission Risk Tool, IDI: integrated discrimination improvement, NRI: net reclassification improvement, NN: neural network, ML: machine learning, DNN: deep neural network, ANN: artificial neural network, AUC: area under the curve, CI: confidence interval.

All metrics have been standardized to 3 decimal places and presented as raw values for AUC and as percentages for sensitivity, specificity, PPV and NPV to ensure consistency.

validating the START [5,6,17] and GAPS [26] models. These efforts enhanced the clinical applicability of both models. The majority of the studies employed logistic regression, valued for its simplicity and interpretability, as the primary algorithm. However, more recent studies have shifted towards ML and deep learning techniques to improve predictive accuracy. Despite the potential advantages of deep learning models, such as their ability to capture complex patterns in data, several studies noted that deep learning approaches often required substantial computing resources and time without offering significant improvements in model performance over traditional ML methods [26]. This underscores an important consideration for practical implementation, where computational efficiency is often as crucial as predictive accuracy, especially in time-sensitive environments like the ED.

The operational definition of hospital admission, which is the primary outcome variable in this review, varied across the included studies. Some studies counted transfers to other hospitals as part of the admission outcome [3,4,21,26], while another study included patients who died in the ED [23]. Other studies categorized hospital stays into different durations based on clinical objectives [3,6,17,26]. For the purposes of this review, hospital admission is defined to include both general ward and intensive care unit admissions. This definition supports the review's objective of predicting admissions during triage, which helps reduce boarding times and optimize hospital capacity management, thereby ensuring a comprehensive evaluation of admission outcomes.

The predictors identified as important variables in predictive models were largely consistent across studies, including age, sex, vital signs, and mode of arrival, which are among the most commonly used. These variables are routinely collected in the ED. Triage-based models, which do not require additional or complex data collection, provide immediate predictions that significantly contribute to timely decision-making in an emergency setting. In contrast, models that utilize laboratory or imaging data leverage detailed clinical information to achieve higher accuracy [4,21]. However, they depend on data that are not available during the triage stage. This distinction further underscores the practical value of triage-based models in real-world clinical applications.

Despite the generally high performance of the models, as indicated by AUC values typically ranging from 0.80 to 0.89, several studies have noted challenges related to missing data and varying data quality across different hospital settings [30]. While some models excelled in discrimination, others showed stronger calibration or classification performance. The models by Cusido et al. [2] and Cameron et al. [23] were particularly notable as top performers, providing a balance of high discrimination and accurate classification, making them promising tools for predicting hospital admissions from ED triage data.

Although internal validation was performed in most studies using methods like cross-validation or bootstrapping, external validation was conducted in only a few instances. This raises concerns about the generalizability of these models across various clinical contexts.

Our findings are consistent with previous systematic reviews, such as the one conducted by Brink et al. [10], which evaluated admission prediction models but noted their limited real-world application due to challenges in clinical usability and validation. Brink's study was confined to European countries, which restricts its global applicability. In contrast, our review encompasses studies from a broader range of countries, enhancing its relevance to a wider variety of clinical settings. Sanchez-Salmeron et al. [12] proposed that ML-based models hold promise as effective tools for enhancing triage-based predictions. However, the deployment of these models in EDs faces significant hurdles, especially the substantial computational resources needed for real-time predictions.

This study makes a significant contribution to the ongoing efforts to refine hospital admission prediction models by specifically focusing on those that utilize critical triage information, which is readily available upon patient arrival. By systematically reviewing the performance and applicability of these models, we offer a comprehensive evaluation of the current state of hospital admission prediction tools. This review highlights the strengths and weaknesses of various algorithms and methodologies.

The review highlights the potential of models based on triage information to enhance ED operations, especially by facilitating earlier predictions of hospital admissions. This could lead to shorter boarding times and improved patient flow. Additionally, the inclusion of studies from various healthcare systems in the review broadens the applicability of its findings, providing insights relevant to diverse clinical settings.

Despite its contributions, this review has several limitations. First, the diversity in study designs, predictors, and outcome definitions complicates direct comparisons of model performance across studies. Additionally, although many studies addressed the issue of missing data, some did not disclose their data handling strategies, which could bias their results. Future research should enhance transparency

https://doi.org/10.4258/hir.2025.31.1.23

in data handling and consider incorporating unstructured data, such as nursing assessments, which have been shown to improve predictive power in certain studies. Furthermore, the absence of external validation in most studies raises concerns about the generalizability of these models to different clinical settings. Future studies should focus on validating models across various EDs to confirm their wider applicability.

This review underscores the potential of predictive models for hospital admissions based on triage data in EDs. Models like START and GAPS, which have been subjected to both extension and external validation, are particularly promising for clinical implementation. Given the broad availability of the identified predictors in ED settings, these models show great promise in reducing boarding times and enhancing patient flow through earlier bed assignments.

Conflict of Interest

Mona Choi is an editorial member of Healthcare Informatics Research; however, she did not involve in the peer reviewer selection, evaluation, and decision process of this article. Otherwise, no potential conflict of interest relevant to this article was reported.

ORCID

Hyun A Shin (https://orcid.org/0000-0002-8142-8281)
Hyeonji Kang (https://orcid.org/0009-0009-5896-4615)
Mona Choi (https://orcid.org/0000-0003-4694-0359)

References

1. Chen W, Linthicum B, Argon NT, Bohrmann T, Lopiano K, Mehrotra A, et al. The effects of emergency department crowding on triage and hospital admission decisions. Am J Emerg Med 2020;38(4):774-9. https://doi.org/10.1016/j.ajem.2019.06.039
2. Cusido J, Comalrena J, Alavi H, Llunas L. Predicting hospital admissions to reduce crowding in the emergency departments. Appl Sci 2022;12(21):10764. https://doi.org/10.3390/app122110764
3. Dinh MM, Russell SB, Bein KJ, Rogers K, Muscatello D, Paoloni R, et al. The Sydney Triage to Admission Risk Tool (START) to predict Emergency Department Disposition: a derivation and internal validation study using retrospective state-wide data from New South Wales, Australia. BMC Emerg Med 2016;16(1):46. https://doi.org/10.1186/s12873-016-0111-4
4. Zhang X, Kim J, Patzer RE, Pitts SR, Patzer A, Schrager JD. Prediction of emergency department hospital admission based on natural language processing and neural networks. Methods Inf Med 2017;56(5):377-89. https://doi.org/10.3414/ME17-01-0024
5. Jones D, Cameron A, Lowe DJ, Mason SM, O'Keeffe CA, Logan E. Multicentre, prospective observational study of the correlation between the Glasgow Admission Prediction Score and adverse outcomes. BMJ Open 2019;9(8):e026599. https://doi.org/10.1136/bmjopen-2018-026599
6. Ebker-White AA, Bein KJ, Dinh MM. The Sydney Triage to Admission Risk Tool (START): a prospective validation study. Emerg Med Australas 2018;30(4):511-6. https://doi.org/10.1111/1742-6723.12940
7. Araz OM, Olson D, Ramirez-Nafarrate A. Predictive analytics for hospital admissions from the emergency department using triage information. Int J Prod Econ 2019;208:199-207. https://doi.org/10.1016/j.ijpe.2018.11.024
8. Parker CA, Liu N, Wu SX, Shen Y, Lam SS, Ong ME. Predicting hospital admission at the emergency department triage: a novel prediction model. Am J Emerg Med 2019;37(8):1498-504. https://doi.org/10.1016/j.ajem.2018.10.060
9. Sun Y, Heng BH, Tay SY, Seow E. Predicting hospital admissions at emergency department triage using routine administrative data. Acad Emerg Med 2011;18(8):844-50. https://doi.org/10.1111/j.1553-2712.2011.01125.x
10. Brink A, Alsma J, van Attekum LA, Bramer WM, Zietse R, Lingsma H, et al. Predicting inhospital admission at the emergency department: a systematic review. Emerg Med J 2022;39(3):191-8. https://doi.org/10.1136/emermed-2020-210902
11. Larburu N, Azkue L, Kerexeta J. Predicting hospital ward admission from the emergency department: a systematic review. J Pers Med 2023;13(5):849. https://doi.org/10.3390/jpm13050849
12. Sanchez-Salmeron R, Gomez-Urquiza JL, Albendin-Garcia L, Correa-Rodriguez M, Martos-Cabrera MB, Velando-Soriano A, et al. Machine learning methods applied to triage in emergency services: a systematic review. Int Emerg Nurs 2022;60:101109. https://doi.org/10.1016/j.ienj.2021.101109
13. Fernandes M, Vieira SM, Leite F, Palos C, Finkelstein S, Sousa JM. Clinical decision support systems for triage in the emergency department using intelligent systems:

a review. Artif Intell Med 2020;102:101762. https://doi.org/10.1016/j.artmed.2019.101762
14. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. https://doi.org/10.1136/bmj.n71
15. Moons KG, de Groot JA, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. PLoS Med 2014;11(10):e1001744. https://doi.org/10.1371/journal.pmed.1001744
16. Wolff RF, Moons KG, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. Ann Intern Med 2019;170(1):51-8. https://doi.org/10.7326/M18-1376
17. Ebker-White A, Bein KJ, Dinh MM. Extending the Sydney Triage to Admission Risk Tool (START+) to predict discharges and short stay admissions. Emerg Med J 2018;35(8):471-6. https://doi.org/10.1136/emermed-2017-207227
18. Peck JS, Gaehde SA, Nightingale DJ, Gelman DY, Huckins DS, Lemons MF, et al. Generalizability of a simple approach for predicting hospital admission from an emergency department. Acad Emerg Med 2013;20(11):1156-63. https://doi.org/10.1111/acem.12244
19. Xie F, Zhou J, Lee JW, Tan M, Li S, Rajnthern LS, et al. Benchmarking emergency department prediction models with machine learning and public electronic health records. Sci Data 2022;9(1):658. https://doi.org/10.1038/s41597-022-01782-9
20. Levin S, Toerper M, Hamrock E, Hinson JS, Barnes S, Gardner H, et al. Machine-learning-based electronic triage more accurately differentiates patients with respect to clinical outcomes compared with the emergency severity index. Ann Emerg Med 2018;71(5):565-74. https://doi.org/10.1016/j.annemergmed.2017.08.005
21. Raita Y, Goto T, Faridi MK, Brown DF, Camargo CA Jr, Hasegawa K. Emergency department triage prediction of clinical outcomes using machine learning models. Crit Care 2019;23(1):64. https://doi.org/10.1186/s13054-019-2351-7
22. Tschoellitsch T, Seidl P, Bock C, Maletzky A, Moser P, Thumfart S, et al. Using emergency department triage for machine learning-based admission and mortality prediction. Eur J Emerg Med 2023;30(6):408-16. https://doi.org/10.1097/MEJ.0000000000001068
23. Cameron A, Rodgers K, Ireland A, Jamdar R, McKay GA. A simple tool to predict admission at the time of triage. Emerg Med J 2015;32(3):174-9. https://doi.org/10.1136/emermed-2013-203200
24. Handly N, Thompson DA, Li J, Chuirazzi DM, Venkat A. Evaluation of a hospital admission prediction model adding coded chief complaint data using neural network methodology. Eur J Emerg Med 2015;22(2):87-91. https://doi.org/10.1097/MEJ.0000000000000126
25. Peck JS, Benneyan JC, Nightingale DJ, Gaehde SA. Predicting emergency department inpatient admissions to improve same-day patient flow. Acad Emerg Med 2012;19(9):E1045-54. https://doi.org/10.1111/j.1553-2712.2012.01435.x
26. Rendell K, Koprinska I, Kyme A, Ebker-White AA, Dinh MM. The Sydney Triage to Admission Risk Tool (START2) using machine learning techniques to support disposition decision-making. Emerg Med Australas 2019;31(3):429-35. https://doi.org/10.1111/1742-6723.13199
27. Graham B, Bond R, Quinn M, Mulvenna M. Using data mining to predict hospital admissions from the emergency department. IEEE Access 2018;6:10458-69. https://doi.org/10.1109/ACCESS.2018.2808843
28. Lee JT, Hsieh CC, Lin CH, Lin YJ, Kao CY. Prediction of hospitalization using artificial intelligence for urgent patients in the emergency department. Sci Rep 2021;11(1):19472. https://doi.org/10.1038/s41598-021-98961-2
29. Zlotnik A, Alfaro MC, Perez MC, Gallardo-Antolin A, Martinez JM. Building a decision support system for inpatient admission prediction with the Manchester triage system and administrative check-in variables. Comput Inform Nurs 2016;34(5):224-30. https://doi.org/10.1097/CIN.0000000000000230
30. Kim M, Park S, Kim C, Choi M. Diagnostic accuracy of clinical outcome prediction using nursing data in intensive care patients: a systematic review. Int J Nurs Stud 2023;138:104411. https://doi.org/10.1016/j.ijnurstu.2022.104411