International Emergency Nursing 86 (2026) 101809

Contents lists available at ScienceDirect

International Emergency Nursing

journal homepage: www.elsevier.com/locate/aaen

Review

Artificial intelligence in emergency nursing: A scoping review of applications and implications

Chiyoung Cha [a], Nayeon Yi [b,c,*], Gumhee Baek [b]

[a] College of Nursing & Ewha Research Institute of Nursing Science, Ewha Womans University, Seoul, South Korea
[b] College of Nursing, Ewha Womans University, Seoul, South Korea
[c] Emergency Department, Inje University Ilsan Paik Hospital, Goyang, Gyeonggi-do, South Korea

ARTICLE INFO

Keywords:
Artificial Intelligence
Emergency Nursing
Emergency Service, Hospital
Nursing practice

ABSTRACT

Introduction: Artificial intelligence (AI) has advanced rapidly in healthcare; however, its application in emergency nursing remains underexplored. This study aimed to map and synthesise existing evidence to clarify current applications, gaps, and practical implications.
Methods: A scoping review was conducted using seven databases from inception to 22 November 2024. From 1,885 initial records, 27 studies were included in the final analysis. The review followed the PRISMA-ScR guidelines, with data extracted in standardised formats and analysed using Arksey and O'Malley's framework.
Results: The study was conducted across North America, Asia, Europe, the Middle East, and Oceania. A total of 16 studies focused on triage and decision support, while others addressed direct nursing practices or emergency department operations. Evaluation of AI performance was reported in 20 studies, revealing considerable variability across algorithms, models, and metrics, with machine learning being the predominant approach; however, operational ethical aspects were explicitly discussed in just 11 studies.
Conclusion: AI demonstrates strong potential for triage, workflow efficiency, and patient safety in emergency nursing but remains poorly integrated into clinical practice. Sustainable progress requires high-quality data, rigorous validation, auditability, and ethical safeguards. Institutional and governmental support, multidisciplinary collaboration, and nurse capacity building are critical for safe, equitable, and scalable implementation.

1. Introduction

Overcrowding resulting from increased emergency department (ED) visits [1,2] can lead to staffing shortages, limited bed availability, and prolonged length of stay (LOS), culminating in extended patient wait times and increased mortality [3]. Addressing these challenges within existing ED healthcare systems is complex, as the provision of timely and safe medical services remains paramount.

With rapid advances in data science and artificial intelligence (AI), numerous AI technologies leveraging enhanced clinical data have been integrated into healthcare systems. AI is increasingly recognised as a key driver of precision health, representing an important direction for future healthcare by enabling more personalised, predictive, and data-driven care. In emergency nursing, AI has been applied to address specific challenges and optimise care delivery. AI-driven clinical decision support systems show promise as adjunct tools for emergency nurses, facilitating rapid assessment and decision-making for diverse ED patients by enhancing identification of critical cases [4,5], optimising triage processes [6,7], and improving severity distribution and ED flow [5]. Through interaction with AI, nurses can augment and streamline professional practice to support optimal care [8]. In high-pressure acute care settings requiring clinical decision-making, such systems may support more informed decisions, simplify workflows, and enable earlier interventions.

However, AI prediction models demonstrate variable performance depending on patient characteristics, settings, and temporal factors. For instance, race has been associated with differences in predictive outcomes for specific targets [5]. In models undergoing internal validation, changes in settings have resulted in declines in system performance [9,10]. Models subjected to external validation have shown performance variation over time after development [11]. By contrast, models based on retrospective data may not fully capture the complexity of real-world ED environments [12]. Performance metrics may also not fully reflect practical utility, as evaluations are limited to specific measures

* Corresponding author at: College of Nursing, Ewha Womans University, Helen Building, 52, Ewhayeodae-gil, Seodaemun-gu, Seoul, South Korea.
E-mail address: n990011@gmail.com (N. Yi).

https://doi.org/10.1016/j.ienj.2026.101809
Received 9 October 2025; Received in revised form 23 March 2026; Accepted 26 March 2026
Available online 8 April 2026
1755-599X/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

[13–15]. These limitations underscore the need for comprehensive understanding of AI-based systems among clinical stakeholders, recognition of associated concerns, and careful approaches to implementation [9].

Accordingly, synthesising existing evidence is essential for understanding the distinctive ED context prior to clinical implementation of AI. Practical implementation also requires consideration of potential challenges, applicability, and the promotion of trust and ethical principles [16–18], grounded in patient-centred perspectives [8,17,18]. The distinctive attributes of emergency nursing, which prioritise timely human care, necessitate an understanding of AI-driven systems and their real-world applications, highlighting the need for a comprehensive literature review.

Previous reviews have focused on AI implementation in emergency medicine [19] or general nursing [20]. Emergency nursing has often been addressed within limited domains, such as triage and machine learning [6,7]. Despite its dynamic and specialised nature, few studies have comprehensively examined model evaluation and operational AI ethics for safe and practical application. This gap limits understanding of the integration and clinical utility of AI models within complex acute care workflows and their future potential.

Given the increasing significance of digital healthcare, systematic assessment of AI models based on global governance guidance and reporting standards is warranted. This review aimed to categorise and synthesise the literature on AI applications and evaluation in emergency nursing to inform safe and effective implementation. The findings are expected to provide insights for identifying research directions and improving workflows and processes that can support reliable medical models in emergency nursing.

2. Design

A scoping review was conducted using the methodological framework proposed by Arksey and O'Malley [21] to explore and synthesise literature on AI applications in emergency nursing. Results were reported in accordance with the PRISMA-ScR guidelines [22] (Table S1).

3. Methods

3.1. Research question

What is known from the existing literature regarding AI applications in emergency nursing?

3.2. Search strategy

A comprehensive literature search of seven databases was conducted between 30 October and 22 November 2024. In the initial phase, appropriate keywords were identified for each concept, and search strategies were subsequently tailored to each database. The databases included six international sources—MEDLINE, Embase, Cochrane Library, Web of Science, CINAHL, and IEEE Xplore—and the national RISS database. A manual search of references and relevant journals was also performed. The full search strategy, including database-specific keywords, is presented in Table S2.

3.3. Eligibility criteria

3.3.1. Inclusion criteria [23]

(1) Population: Patients and nurses in the ED
(2) Concept: Use of AI in emergency nursing practice
(3) Context: ED settings

3.3.2. Exclusion criteria

(1) Publications duplicating original study content, including conference abstracts and preliminary or pilot reports
(2) Non-primary research articles lacking sufficient information to identify or evaluate AI mechanisms (e.g., letters to the editor, editorial responses, and commentaries)
(3) Retracted articles
(4) Studies published in languages other than English or Korean

3.4. Study selection

The search and selection processes are presented in the PRISMA-ScR flowchart (Fig. 1). Retrieved studies were imported into EndNote 21.5 (Clarivate Analytics, Philadelphia, PA, USA) to remove duplicates and then uploaded to Covidence (Veritas Health Innovation, Melbourne, Australia) for screening. Two researchers independently screened titles and abstracts. The primary researcher subsequently reviewed full texts, and a second researcher verified accuracy and consistency. A third researcher reviewed study titles and types to ensure appropriate application of inclusion and exclusion criteria. Disagreements were resolved through discussion to achieve consensus. A manual search of references from included studies was also conducted, resulting in a final sample of 27 studies.

3.5. Data extraction

The primary researcher reviewed full texts using Covidence and extracted key data into a standardised Microsoft Excel 2019 (Microsoft Corporation, Redmond, WA, USA) template developed by the research team prior to extraction. A second researcher verified the process to minimise bias. Calibration meetings were held biweekly to address issues and refine the extraction process. The data extraction form was iteratively updated during the study. A third researcher with expertise in AI-based clinical decision-making in ED settings independently reviewed the extracted data to ensure accuracy and appropriateness. Extracted variables included AI type, data source, validation, discrimination, calibration, and operational AI ethics (Tables 1 and 2). Discrepancies were resolved through consensus. Quality appraisal was not conducted, as the aim was to map the scope of the literature.

3.6. Data summary and synthesis

The extracted data were summarised descriptively and organised into tables. A narrative synthesis was conducted to map AI application areas, evaluate model performance, and identify ethical and methodological gaps.

4. Results

4.1. Characteristics of the retrieved studies

The general characteristics of the 27 studies are presented in Table 1. Studies were conducted across multiple continents, including North America (n = 10, 37.0%) [24–33], Asia (n = 7, 25.9%) [34–40], Europe (n = 3, 11.1%) [41–43], the Middle East (n = 3, 11.1%) [44–46], and Oceania (n = 2, 7.4%) [10,47]; location was not reported in two studies [48,49]. The most common designs were quantitative, including retrospective (n = 16; 59.3%) [10,27–35,37–39,47–49], prospective (n = 6, 22.2%) [36,40–42,44,46], and comparative (n = 2, 7.4%) [43,45] approaches. The remaining three studies employed qualitative designs [24–26].

Studies were conducted in single- and multi-centre settings, involving up to 12 centres, with sample sizes ranging from 11 [24] to 2,458,857 [31]. All studies included participants: ED patients in 23 studies (85.2%) [10,25,27–40,42–44,46–49] and emergency nurses in

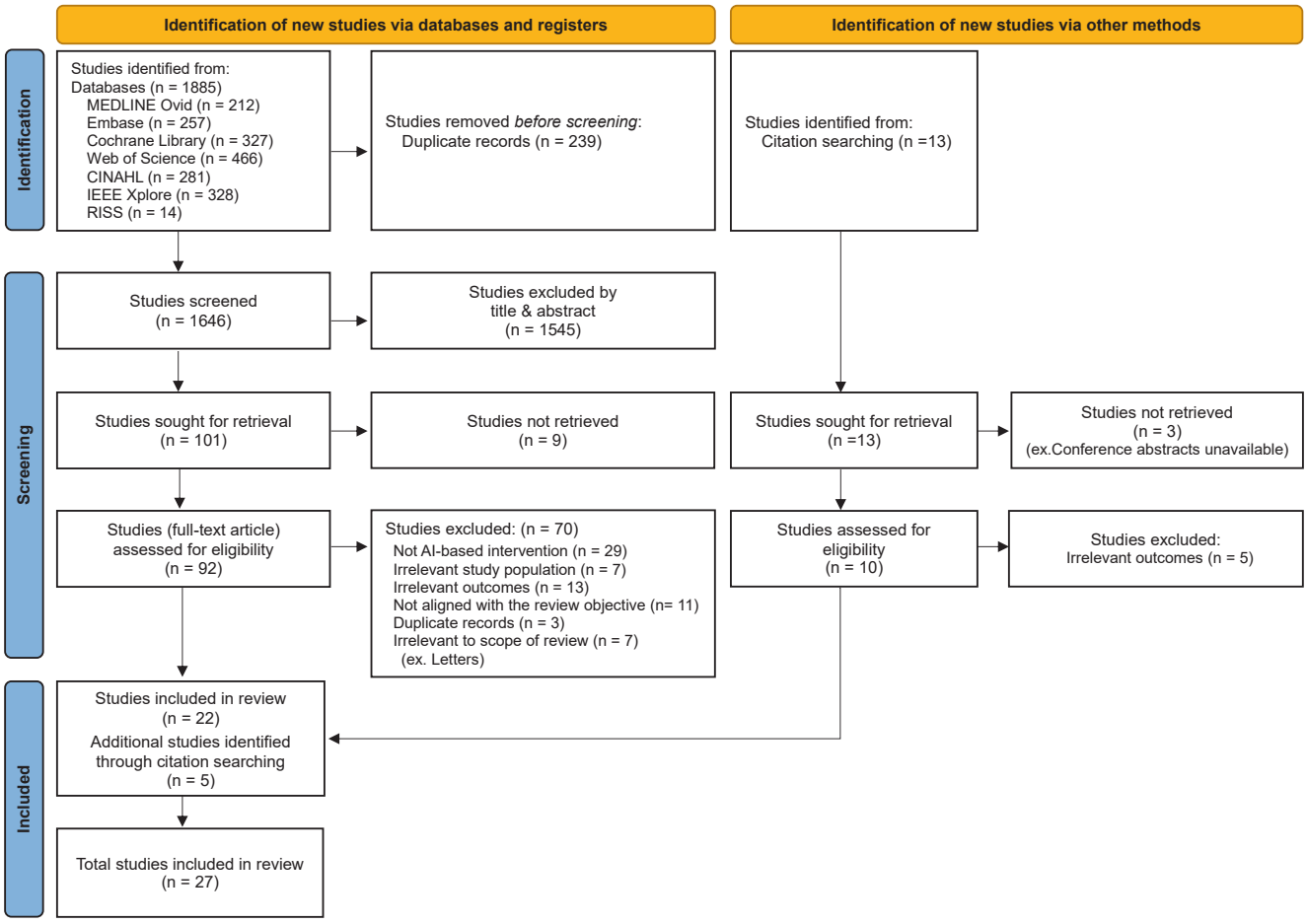


Fig. 1. Flow Diagram Corresponding to the Selection of the Studies.

four (14.8%) [24,26,41,45]. Among patient-based studies, 12 included adult patients within defined age ranges [27,28,33–35,37–39,43,44,46,47], and four focused on specific conditions, such as triage levels 1–2 [34], COVID-19 [27], or acute stroke or abdominal pain [37,44].

4.2. Scope of AI applications in emergency nursing

AI applications were categorised into three domains: triage tasks (n = 16, 59.3%) [26,28–33,37–40,42–46], nursing care (n = 6, 22.2%) [24,25,36,41,47,49], and systematic tasks (n = 5, 18.5%) [10,27,34,35,48] (Fig. 2). AI applications supporting triage decision-making included acuity level classification (n = 11) [26,31–33,39,40,42–46], chief complaint categorisation (n = 1) [30], resource prediction (n = 1) [29], electrocardiogram acquisition (n = 1) [38], sepsis identification (n = 1) [28], and acute stroke identification (n = 1) [37]. AI applications in nursing care (n = 6) addressed intravenous injections [24,25], patient-related risks [41,49], pain management [47], and medical records [36]. AI for systematic tasks (n = 5) focused on patient flow and resource allocation associated with patient volume, waiting times, and LOS (n = 4) [10,27,34,48], as well as workplace issues such as violence (n = 1) [35].

4.3. Outcomes of AI applications in emergency nursing

Outcomes associated with AI use are presented in Table 2.

4.3.1. Type of AI

Among the 27 studies, 25 (92.6%) explicitly reported the AI algorithms used; two qualitative studies did not. Machine learning was the most common approach (n = 22, 81.5%) [10,26–29,31–41,43–45,47–49], whereas expert systems, fuzzy logic, and rule-based natural language processing (NLP) were used less frequently (n = 3, 11.1%) [30,42,46]. Among machine learning studies, implementations included standalone models (n = 10) [10,27,31,33–38,41], deep learning (n = 5) [29,40,43,45,47], multimodal approaches (n = 5) [28,29,32,39,40], and one ensemble model (n = 1) [44]. Standalone models comprised decision tree (DT), support vector machine (SVM), naïve Bayes (NB), K-nearest neighbor (KNN), logistic regression (LR), and neural networks (NN), as well as extended methods such as random forest (RF), gradient boosting (GB), C4.5, classification and regression trees (CART), and elastic net (EN). Deep learning included a single-model approach (n = 1, 3.7%) [47] and multimodal approaches integrating structured and unstructured data (n = 2, 7.4%) [29,40]. Multimodal implementations combining machine learning with NLP included XGBoost with NLP (n = 2, 7.4%) [28,32] and combinations of LR, RF, and XGBoost with NLP (n = 1, 3.7%) [39]. One study used an ensemble combining DT, LR, NB, NN, association rules, and clustering algorithms [44].

4.3.2. Data source

Most studies (n = 19, 70.4%) used electronic health records (EHRs) [10,27–35,37–40,42,44,46,47,49]. Two studies (7.4%) supplemented EHRs with government datasets or electronic medication dispensing systems [10,47]. Three studies (7.4%) used observed nurse–patient interactions, real-time recordings, or historical ED data with external factors [36,41,48]. Two studies (7.4%) used case- and EHR-derived vignette ratings [43,45], and qualitative studies (n = 3, 11.1%) used

Table 1
Description of studies.

Author (year)	Country	Study design	Setting	Participant	Sample size	Application in nursing
Hudson et al. (2023)	Canada	Qualitative study	2 pediatric hospitals	ED HPs	11	Distraction support during IVI
Nishat et al. (2023)	Canada	Qualitative study	2 pediatric hospitals	ED patients (5–11 years)/ Caregivers	20	Distraction support during IVI
Jordan et al. (2023)	United States	Qualitative study	1 academic hospital	ED Nurses	13	Clinical decision support at triage
Etu et al. (2022)	United States	Retrospective study	1 academic hospital	ED Patients (adults) COVID–19 positive	3,301	Allocate resources and inform patients
Lee et al. (2023a)	South Korea	Retrospective study	1 tertiary hospital	ED Patients (≥18 years) KTAS 1–2 transferred	968	Allocate resources
Podolsky (2019)	N/R	Retrospective study	12 hospitals	ED Patients	N/R	Allocate resources
Bagnasco et al. (2015)	Italy	Prospective study	3 hospitals	ED Nurses	840	Prevent harmful communication errors
Edgcomb et al. (2023)	N/R	Retrospective study	1 academic hospital	ED Patients (10–17 years)	1,093	Prevent suicide–related emergency
Hughes et al. (2024)	Australia	Retrospective study	1 tertiary hospital	ED Patients (≥18 years)	235,789	Management of pain
Lee et al. (2023b)	South Korea	Retrospective study	1 tertiary hospital	ED Patients (adults)	7,259	Prevent workplace violence
Walker et al. (2022)	Australia	Retrospective study	12 hospitals	ED Patients	1,930,609	Inform patient flow in the post–triage queue
Cho et al. (2022)	South Korea	Prospective study	1 tertiary hospital	ED Patients	1,063	Time–critical task assistance at triage
Brann et al. (2024)	United States	Retrospective study	4 hospitals	ED Patients (≥18 years)	1,059,386	Clinical decision support at triage
Sterling et al. (2020)	United States	Retrospective study	3 hospitals	ED Patients	226,317	Clinical decision support at triage
Sung et al. (2021)	Taiwan	Retrospective study	1 academic hospital	ED Patients (adults) suspected stroke or TIA	1,361	Clinical decision support at triage
Tsai et al. (2022)	Taiwan	Retrospective study	1 tertiary hospital	ED Patients (≥20 years)	301,658	Clinical decision support at triage
Tootooni et al. (2019)	United States	Retrospective study	1 academic hospital	ED Patients	147,859	Clinical decision support at triage
Choi et al. (2019)	South Korea	Retrospective study	1 tertiary hospital	ED patients (adults)	138,022	Clinical decision support at triage
Farahmand et al. (2017)	Iran	Prospective study	1 tertiary hospital	ED patients (>18 years) acute abdominal pain	215	Clinical decision support at triage
Hall et al. (2023)	Canada	Retrospective study	3 hospitals	ED Patients	2,458,857	Clinical decision support for virtual triage
Ivanov et al. (2021)	United States	Retrospective study	2 hospitals	ED patients (≥1 year)	147,052	Clinical decision support at triage
Kipourgos et al. (2022)	Greece	Prospective study	1 tertiary hospital	ED Patients	616	Clinical decision support at triage
Lin et al. (2024)	Taiwan	Prospective study	1 tertiary hospital	ED patients	901	Clinical decision support at triage
Nasser et al. (2024)	Canada	Retrospective study	1 tertiary hospital	ED Patients (≥18 years)	21,469	Clinical decision support for virtual triage
Saban & Dubovi (2024)	Israel	Comparative vignette study	9 hospitals/ 1 Nursing school	ED Nurses/ Nursing students	30/38	Clinical decision support at triage
Soufi et al. (2018)	Iran	Prospective study	1 tertiary hospital	ED patients (>18 years)	537	Clinical decision support at triage
Zaboli et al. (2024)	Italy	Comparative vignette study	1 hospital	ED patients (≥18 years)	30	Clinical decision support at triage

Abbreviations: COVID–19, Coronavirus disease–19; ED, emergency department; HP, healthcare professional; IVI, intravenous insertion; KTAS, Korean Triage and Acuity Scale; N/R, not reported; TIA, transient ischemic attack.

interview data [24–26].

4.3.3. *Predictive performance of AI models*

Most studies (n = 20, 74.1%) reported predictive performance using metrics such as accuracy, F1 score, and area under the curve (AUC) [10,27–32,34,35,37–44,47–49]. Ten studies (37.0%) used both threshold-dependent and threshold-independent metrics, whereas nine (33.3%) primarily used threshold-dependent metrics (e.g., accuracy, F1 score). Performance varied across models, algorithms, and metrics, with over one-third focusing on predicting patient acuity [31,32,39,40,42–44]. Among single-model implementations, gradient boosting regression showed the highest F1 performance at triage level 3 but performed poorly at levels 1 and 5 (F1 = 0) [31], reflecting limitations in handling class imbalance. AUC was used to assess discrimination. One study reported that a naïve Bayes-based ensemble achieved higher AUC at triage level 4 than decision trees, logistic regression, neural networks, association rules, and clustering, as well as higher overall accuracy [44]. Among 13 studies (48.1%) using probability-based models, 10 reported discrimination via AUC [27,28,32,34,35,37–39,41,44]. However, only two studies assessed calibration, using the Hosmer-Lemeshow test in logistic regression [34,37].

Table 2
Results of the AI Application.

Author (year)	AI type	AI implementation	Data source	Validation	Main finding	Calibration	AI ethics
Hudson et al. (2023)	Unspecified	AI–enhanced Socially Assistive Robotics (SAR)	Interview data	N/A	Perspectives on AI–enhanced SAR design • Supporting autonomous adaptation • Complementing HP–led care	N/A	• Human override
Nishat et al. (2023)	Unspecified	AI–enhanced SAR	Interview data	N/A	Perspectives on AI–enhanced SAR design • Acceptance and concerns surrounding SAR	N/A	• Equity • Patient data privacy
Jordan et al. (2023)	ML	AI clinical decision support (KATE)	Interview data	N/A	Cultural impact on AI triage adoption • Decontextualization of AI • Reinforcement through AI feedback • Safety net in triage	N/A	• N/R
Etu et al. (2022)	ML LR, GB, DT, RF	*Individual models*	EHR	Internal	ED–LOS of < 4 h • Accuracy: 79 – 85% • F1 score: 0.84 – 0.88 • AUC: 0.88 – 0.93 • Performance superiority of preprocessed & balanced data	N/R	• Model explainability
Lee et al. (2023a)	ML LR, RF, DT, SVM, NB	*Individual models*	EHR	Internal (Cross–validation)	ED–LOS of ≤ 6 h • Accuracy: 75 – 77% • F1 score: 0.67 – 0.70 • AUC: 0.74 – 0.85	Limited (HL for LR only)	• N/R
Podolsky (2019)	ML (Reinforcement learning)	N/R	Historical data & external variables	N/R	Patient volume • Accuracy: 88 – 93%	N/A	• N/R
Bagnasco et al. (2015)	ML ANN–based	Multi–layered feedforward neural network	Observed nurse–patient interactions & personal data	Internal	Nurse–patient communication risk (*Guilbert's grid*) [a] • Accuracy: $>$ 80% all dimensions • AUC: 0.87 – 0.95	N/R	• N/R
Edgcomb et al. (2023)	ML RF, LASSO LR	N/R	EHR	Internal (Cross–validation)	Suicide–related emergencies (*C–CASA*) [a] • Sensitivity: 0.84 – 0.92 • Enhanced performance with additional features	N/R	• N/R
Hughes et al. (2024)	ML DL (Transformer–based)	NLP	EHR PYXIS	Internal	Time–of–triage pain (*ED nurse*) [a] • Accuracy: 93.2% • Analgesia received: <50%	N/A	• Patient data privacy
Lee et al. (2023b)	ML LR, DT, NB, RF	*Individual models*	EHR	Internal	Workplace violence • Accuracy: 72 – 86% • AUC: 0.77 – 0.90 • Performance superiority of multi–factor	N/R	• N/R
Walker et al. (2022)	ML LR, RF, EN	*Individual models*	VEMD EHR	Internal External Temporal	Triage–to–provider wait times • Accuracy: 40 – 63% • MAE: site–specific lower than global • COVID–19: minor impact	N/A	• Model transparency
Cho et al. (2022)	ML	Speech–to–Text (DNN) NLP	Real–time audio recordings	N/A	Time taken for triage task • Manual: 231 sec, model: 204 sec • Record completion: inferior to manual • Record accuracy: lower in category	N/A	• Patient data privacy
Brann et al. (2024)	ML GB	*Multi–modal* XGBoost with NLP	EHR (clinical variables + clinical text)	Internal	Time–of–triage sepsis identification (*physician*) [a] • Macro F1: 0.61→ 0.67 • AUC: 0.94→ 0.97 • Improved performance of multi–source data • Clinical relevance through early antibiotic	N/R	• Model explainability

(*continued on next page*)

Table 2 (*continued*)

Author (year)	AI type	AI implementation	Data source	Validation	Main finding	Calibration	AI ethics
Sterling et al. (2020)	ML DL (RNN–based)	*Multi–modal* LSTM with NLP	EHR (clinical variables + clinical text)	Internal	Number of resources • Accuracy: nurse 69.0%, model 65.9% • Macro F1: nurse 0.592, model 0.589	N/R	• N/R
Sung et al. (2021)	ML C4.5, CART, KNN, LR, RF, SVM	*Individual models*	EHR	Internal	Time–of–triage stroke identification (*neurologist*) [a] • F1 score: FAST 0.41–0.69, model 0.66–0.76 • AUPRC: FAST 0.56–0.74, model 0.66–0.79 • SMOTE improved performance	Limited (HL for LR only)	• N/R
Tsai et al. (2022)	ML LR, DT, RF, GB	*Individual models* XGBoost (Final)	EHR	Internal Temporal	ECG acquisition <2 h of the ED visit • F1 score: 0.757→ 0.729 • AUC: 0.891→ 0.885 • COVID–19: minor impact	N/R	• Expert autonomy
Tootooni et al. (2019)	Rule–based NLP (heuristic)	Shortcut Word & Bag–of–Words mapping	EHR	Internal	Time–of–triage CC categorization (*ED physician*) [a] • F1 score: overall 94.7%, multi–CC 82.3% • Lower performance in multi–CC cases	N/A	• Model transparency
Choi et al. (2019)	ML LR, RF, GB (XGBoost)	*Multi–modal* with NLP	EHR (clinical variables + clinical text)	N/R	KTAS level (*actual value*) [a] • F1 score: 0.623 – 0.740 • AUC: 0.842 – 0.922 • Performance superiority of multi–source data	N/R	• N/R
Farahmand et al. (2017)	ML AR, CL, DT, LR, NB, NN	*Ensemble models* Web–based interface	EHR	Internal	ESI level (*ED physician*) [a] • Overall accuracy: 71.9 – 78.1% • AUC: 0.459 – 0.839	N/R	• N/R
Hall et al. (2023)	ML DT, KNN, RF, GBR, NN	*Individual models* GBR (Final) *	EHR	Internal	CTAS level (*actual value*) [a] • Accuracy: 47.4%, under–triage 5.0% • F1 score: 0.00, 0.46, 0.74, 0.21, 0.00	N/A	• N/R
Ivanov et al. (2021)	ML GB	*Multi–modal* XGBoost with C–NLP	EHR (clinical variables + clinical text)	Internal	ESI level (*ED nurse and physician*) [a] • Accuracy: nurse 59.8%, model 75.7% • F1 score: nurse 0.43, model 0.74 • AUC: nurse 0.75, model 0.85	N/R	• N/R
Kipourgos et al. (2022)	Fuzzy Logic ML DT	Fuzzy CLIPS J–48 (WEKA)	EHR	Internal (WEKA only)	ESI level (*ED nurse*) [a] • Accuracy: Fuzzy 99%, WEKA 72 – 95%	N/A	• N/R
Lin et al. (2024)	ML DL	*Multi–modal* CNN, MLP with TabNet, NLP	EHR (clinical variables + clinical text)	Internal	TTAS level (*physician*) [a] • Accuracy: 43.9 – 63.3% • F1 score: 0.162 – 0.658 • Performance superiority of multi–source data	N/R	• Model explainability
Nasser et al. (2024)	ML GBR	VTriage (Mobile App) *	EHR	N/A	CTAS level interrater reliability (vs. ED nurse) • Unweighted kappa: 0.18 • Quadratic–weighted kappa: 0.36 • Under–triage 5.5%	N/A	• N/R
Saban & Dubovi (2024)	ML DL (LLM)	Generative AI (ChatGPT 3.5)	Literature & real case–derived vignette ratings	N/A	CTAS level agreement (vs. ED nurse, nursing student) • GPT: indecisive urgency scoring • Resource prediction at triage – GPT: tended to over–order tests	N/A	• N/R
Soufi et al. (2018)	Expert System Fuzzy Logic	*Hybrid* Rule–Based Reasoning Fuzzy Logic Classifier	EHR	N/A	ESI level agreement (vs. expert) • Accuracy: 99.44% • Record completeness improvement	N/A	• Model explainability
Zaboli et al. (2024)	ML DL (LLM)	Generative AI (ChatGPT 3.5)	EHR–derived vignette ratings	N/A	MTS level interrater reliability (vs. HP) • Unweighted kappa: 0.278 Classification performance (*ED physician*) [a] • HP outperforming GPT	N/A	• N/R

Abbreviations: AI, artificial intelligence; ANN, artificial neural network; AR, association rules; AUC, area under the receiver operating characteristic curve; AUPRC, area under the precision recall curve; CART, Classification and Regress Tree; C–CASA, Columbia Classification Algorithm for Suicide Assessment; CC, chief complaint; C–NLP, clinical natural language processing; CNN, convolutional neural network; CL, Clustering; COVID–19, Coronavirus disease–19; CTAS, Canadian Triage and Acuity Scale; DNN, deep neural network; DL, deep learning; DT, decision tree; ECG, electrocardiogram; ED, emergency department; EHR, electronic health records; EN, elastic net regression; ESI, Emergency Severity Index; FAST, face arm speech test; GB, gradient boosting; GBR, gradient boosting regressor; HL, Hosmer–Lemeshow; HP, Healthcare professional; KNN, K–Nearest Neighbor; KTAS, Korean Triage and Acuity Scale; LOS, length of stay; LLM, large language model; LR, logistic regression; LSTM, long short–term memory; MAE, median absolute error; ML, machine learning; MTS, Manchester Triage System; N/A, not applicable; NB, Naïve Bayes; NLP, natural language processing; N/R, not reported; PYXIS, electronic medication dispensing system; RF, random forest; RNN, recurrent neural networks; SMOTE, synthetic minority oversampling technique; SVM, Support Vector Machine TTAS, Taiwan Triage and Acuity Scale; VEMD, Victorian Emergency Minimum Dataset; WEKA, Waikato environment for knowledge analysis

[a]Ground truth * post–processing manual up–triaging of ≥1% CTAS level 1 complaints

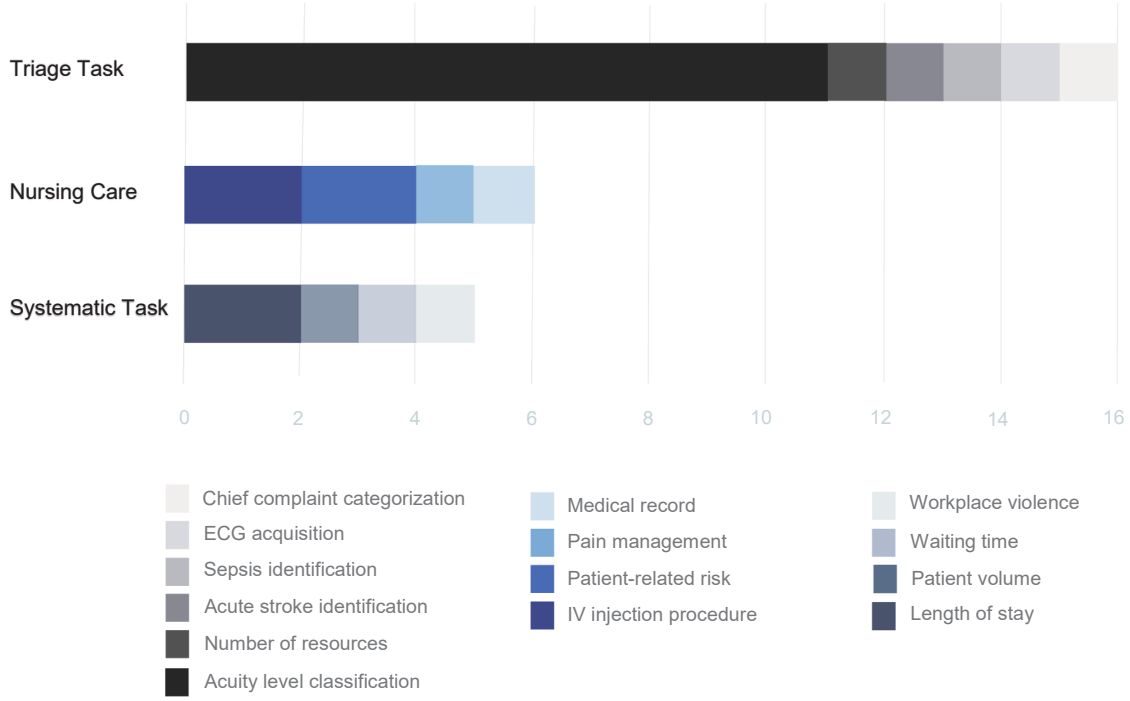


Fig. 2. Scope of AI Applications in Emergency Nursing.

4.3.4. *Evaluation of AI performance*

Ground truth was defined as expert-established standards in nine studies (33.3%), particularly for pain, sepsis, stroke, primary symptoms, and acuity classification [28,30,32,37,40,42–44,47]. Nine studies (33.3%) used observed outcomes (e.g., LOS, patient volume, workplace violence, waiting time, resource use, ECG acquisition, acuity) as ground truth [10,27,29,31,34,35,38,39,48]. Two studies (7.4%) used specific criteria for patient risk classification [41,49].

Performance was evaluated across varying data conditions and against conventional methods, indicating that data quality and context influence outcomes. Eight studies incorporated additional evaluation criteria beyond conventional accuracy measures [10,27,28,30,37–40], examining data modality (n = 3), dataset distribution (n = 2), temporal shift (n = 2), dataset complexity (n = 1), and study site (n = 1). Multimodal data outperformed single-type data [28,39,40], balanced datasets outperformed imbalanced datasets [27,37], and site-specific models outperformed global models [10]. Comparable performance was observed in some contexts, such as COVID-19 prediction [10,38], whereas weaker performance occurred in multi-symptom mapping tasks [30]. One study directly compared AI with a rule-based system, with AI showing superior performance in stroke identification [37].

Six studies (22.2%) compared AI performance with that of emergency nurses, reporting mixed results, including equivalence, superiority, and limited performance [29,32,33,36,43,45]. AI showed comparable performance in resource classification [29] and triage tasks [32]. One study reported superior AI performance in acuity prediction [32], whereas another found limited performance of a large language model in predicting acuity and resource needs [45]. Inter-rater reliability between humans and AI was low in two studies (K = 0.18–0.28), remaining below the moderate threshold [33,43].

4.3.5. *Operational AI ethics*

Ethical considerations were infrequently reported. Some studies addressed model explainability (n = 4, 14.8%), data privacy (n = 3, 11.1%), transparency (n = 2, 7.4%), population equity (n = 1, 3.7%), professional autonomy (n = 1, 3.7%), and human override (n = 1, 3.7%) [10,24,25,27,28,30,36,38,40,46,47]. Explainability was addressed through visualisation techniques such as heat maps, attention mechanisms, and model architectures [27,28,40,46]. Transparency included provision of rule-based algorithms or partial code release to support reproducibility [10,30]. Data privacy measures included on-premises deployment within local clinical environments [47] and cloud-based deployment external to hospital systems [36]. Three studies highlighted population equity and professional autonomy in practical implementation [24,25,38].

5. Discussion

This review provides a comprehensive synthesis of AI applications in emergency nursing, including triage, nursing care, and system-level tasks. The findings highlight both opportunities and limitations, offering insights to inform clinical practice and future research.

The geographic distribution of studies aligns with previous research, with most conducted in North America and Asia [6], indicating a concentration in high-income regions. Despite global momentum towards AI within the Fourth Industrial Revolution, research from South America and Africa remains limited [50]. This imbalance reflects disparities in infrastructure, access to data science resources, and medical information systems, shaped by socioeconomic constraints [51]. Such disparities may contribute to biased model development and skewed outcomes owing to unrepresentative training data [51]. Addressing these gaps requires improved access to affordable technologies and strengthened international collaboration to support equitable AI development.

AI applications in emergency nursing were most frequently focused on triage, comprising over 60% of included studies. Triage prioritises rapid identification of patients requiring urgent care [52]. This predominance may reflect both the time-sensitive nature of ED decision-making and the availability of relevant data. Early AI-supported decision-making during triage may enhance identification and management of complex or undifferentiated conditions, particularly for time-critical cases [4,5]. However, acuity classification models often demonstrate limited performance for the most severe cases (levels 1 and 2) [31]. Models predicting complementary outcomes, such as resource use, patient flow, or specific diagnoses, may help address these limitations. Furthermore, operational constraints, such as limited resources, fluctuating patient volume, and boarding, may affect both AI applicability and care quality [3]. Most studies relied on data collected during initial assessment or triage, including structured variables and clinical notes recorded before physician evaluation. These data are commonly used as model inputs and are central to ED-specific model development.

However, reliance on basic triage data may constrain predictive performance. Incorporating domain-specific clinical data, including unstructured free-text, may enhance model performance [53]. Such data combine structured elements (e.g., vital signs, pain scores) with narrative information on presenting complaints and medical history and have been effectively used in machine learning and NLP applications for risk prediction and classification tasks [4,12]. Further research is needed to optimise the integration of heterogeneous data sources.

Compared with other healthcare domains, AI applications in emergency nursing remain relatively narrow. Broader applications in healthcare include disease diagnosis, imaging, mortality prediction, ED disposition, workflow optimisation, and education [19,20]. These disparities may reflect variations in data availability, interoperability, and workforce capabilities, as well as differences in clinical environments. Effective implementation requires access to diverse data types (text, image, audio, numerical) and their integration through quantification and standardisation [54,55]. This may also enhance nurse engagement in data-driven processes. Expanding AI applications beyond current domains will require that nurses, as key stakeholders, possess appropriate competencies [9,54].

EHRs were the primary data source across studies. Although often incomplete, EHRs provide comprehensive demographic and clinical information and support large-scale analysis through structured storage [56,57]. However, reliance on single-centre data limits external validation [58], and studies used single-institution datasets, often due to retrospective designs. Variability in data structure and potential confounders across institutions further complicate generalisation. Access to sensitive medical datasets and code also remains restricted compared with other fields [58]. Registry data, when available, enable cross-site and longitudinal validation [59] but may lack detailed clinical variables aligned with specific study objectives. Understanding the strengths and limitations of each data source is essential, and careful data engineering is required to ensure reliable outcomes.

Although AI research in emergency nursing is increasing, most models remain confined to research settings with limited integration into clinical workflows. Key limitations include the predominance of retrospective designs and insufficient external validation, as highlighted in global guidelines such as WHO (World Health Organization), TRIPOD + AI (Transparent Reporting of a multivariable prediction model of Individual Prognosis Or Diagnosis), and PROBAST + AI (Prediction model Risk Of Bias Assessment Tool) [13,18,60]. Model performance often declines on independent datasets due to overfitting or domain shifts [9,58]. While external validation efforts are increasing [59], they remain limited and largely retrospective [11,59]. These constraints may hinder safe and reliable clinical implementation, underscoring the need for context-specific models and rigorous validation in future research [54].

AI models generally demonstrated favourable performance but showed weaker results in areas such as wait time prediction, resource allocation, and virtual triage. Direct comparison across studies is limited by heterogeneity in algorithms, datasets, validation approaches, and metrics. Notably, calibration was rarely assessed despite widespread reporting of discrimination metrics. The AUC, commonly used to assess discrimination, does not account for clinical consequences of false positives and false negatives [14,15]. High AUC values may therefore not translate into clinical utility. In emergency settings, calibration and clinical relevance are particularly important, especially for high-risk classifications such as under-triage at acuity levels 1 and 2 [13]. Reporting guidelines (TRIPOD + AI and PROBAST + AI) recommend comprehensive evaluation of model quality, bias, and applicability [13,60]. Decision curve analysis (DCA), which evaluates net benefits across threshold probabilities, may provide additional insight into clinical utility and support model comparison [14,15].

Ethical considerations were inconsistently addressed. Challenges include uncertainties in implementation, legal concerns, and limited integration into clinical workflows. These issues may reflect ongoing development stages and insufficient clarity in ethical frameworks [61]. AI implementation should prioritise human-centred care. Professional guidelines emphasise that AI functions as an adjunct, with nurses retaining responsibility for clinical decisions [8]. In high-pressure ED environments, unaddressed ethical risks may affect both patients and staff. Involving emergency nurses throughout AI development and implementation may help mitigate such risks [8,16,17]. Ethical principles outlined by WHO (i.e. autonomy, safety, transparency, accountability, equity, and sustainability) provide a framework for responsible use [18]. Addressing these principles may support safe AI integration while preserving professional autonomy and patient-centred care (Table S3).

6. Limitations

This study may have limitations regarding comprehensive coverage and generalisability. Although scoping reviews aim to map available evidence broadly, inclusion criteria prioritised studies with sufficient methodological detail to evaluate AI mechanisms and performance. Conference abstracts and preliminary reports were excluded only when duplicative or lacking sufficient detail, rather than based on publication type. Consequently, findings may not fully reflect real-world implementation, as grey literature and practical reports were not extensively captured. Further research is needed to address this gap. Furthermore, the scoping review design limits assessment of effectiveness, given variability in performance metrics and heterogeneity in models and algorithms across studies.

7. Implications of the findings

AI in emergency nursing has evolved from rule-based systems and conventional machine learning using structured data toward multimodal approaches that integrate heterogeneous data and deep learning to support clinical decision-making. This progression indicates a shift from retrospective prediction to real-time, context-aware support. However, many models remain opaque, highlighting the need for explainable AI to support safe use and informed decision-making by nurses [17]. The complexity of ED settings also underscores the need for

precision nursing approaches that incorporate patient-specific clinical and contextual data factors.

Effective integration of AI requires that nurses possess both domain expertise and a foundational understanding of AI, supported by targeted education and training [17]. Active involvement in system design and development is also essential. The ED environment is complex, with frequent task switching and resource constraints; AI may reduce workload and improve operational efficiency [36]. Interdisciplinary collaboration in AI implementation may further enhance care delivery [17,18]. Governmental and institutional support, along with prospective research, is necessary to validate outcomes and support implementation.

Addressing data bias through robust external validation and incorporating AI as a supportive tool within established clinical frameworks may enhance equity, safety, and applicability. Maintaining professional autonomy while integrating AI is essential. These approaches may contribute to improved ED operations and patient safety. Further investigation is needed to determine whether these effects translate into sustained nurse satisfaction and long-term clinical outcomes.

8. Conclusion

This review indicates that AI in emergency nursing, although applied across multiple areas, remains concentrated in triage. Applications in other domains are limited, and model performance varies considerably. Integration into clinical workflows remains insufficient. Advancing AI use in emergency nursing requires strengthening nurses' competencies through education and training and optimising the use of clinical data. Ensuring patient safety through ethical and responsible implementation is critical. Prospective research, rigorous validation, and transparent evaluation are essential to support evidence-based application. These efforts may expand the scope of AI in emergency nursing and facilitate its effective integration into evolving healthcare systems.

CRediT authorship contribution statement

Chiyoung Cha: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Nayeon Yi:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Gumhee Baek:** Writing – review & editing, Visualization, Validation, Supervision.

Funding

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2024–00351194).

Declaration of generative AI use

All content was developed and reviewed solely by the authors. AI tools were not used for data analysis, interpretation, or manuscript drafting.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ienj.2026.101809.

REFERENCES

[1] Centers for Disease Control and Prevention. FastStats-emergency department visits, https://www.cdc.gov/nchs/fastats/emergency-department.htm; 2025 [accessed 1 August 2025].
[2] National Emergency Medical Center. Emergency medical statistical annual report 2014–2023. https://www.e-gen.or.kr/nemc/statistics_annual_report.do; 2024 [accessed 1 August 2025].
[3] Sartini M, Carbone A, Demartini A, Giribone L, Oliva M, Spagnolo AM, et al. Overcrowding in emergency department: Causes, consequences, and solutions—A narrative review. Healthc 2022;10(9):1625. https://doi.org/10.3390/healthcare10091625.
[4] Fernandes M, Mendes R, Vieira SM, Leite F, Palos C, Johnson A, et al. Risk of mortality and cardiopulmonary arrest in critical patients presenting to the emergency department using machine learning and natural language processing. PLoS One 2020;15(4):0230876. https://doi.org/10.1371/journal.pone.0230876.
[5] Taylor RA, Chmura C, Hinson J, Steinhart B, Sangal R, Venkatesh AK, et al. Impact of artificial intelligence–based triage decision support on emergency department care. NEJM AI 2025;2(3):AIoa2400296. https://doi.org/10.1056/AIoa2400296.
[6] Sánchez-Salmerón R, Gómez-Urquiza JL, Albendín-García L, Correa-Rodríguez 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.
[7] Yi N, Baik D, Baek G. The effects of applying artificial intelligence to triage in the emergency department: a systematic review of prospective studies. J Nurs Scholarsh 2025;57(1):105–18. https://doi.org/10.1111/jnu.13024.
[8] American Nurses Association. The ethical use of artificial intelligence in nurse practice, https://www.nursingworld.org/practice-policy/nursing-excellence/official-position-statements; 2022 [accessed 15 February 2026].
[9] Finlayson SG, Subbaswamy A, Singh K, Bowers J, Kupke A, Zittrain J, et al. The clinician and dataset shift in artificial intelligence. N Engl J Med 2021;385(3):283–6. https://doi.org/10.1056/NEJMc2104626.
[10] Walker K, Jiarpakdee J, Loupis A, Tantithamthavorn C, Joe K, Ben-Meir M, et al. Emergency medicine patient wait time multivariable prediction models: a multicentre derivation and validation study. Emerg Med J 2022;39(5):386–93. https://doi.org/10.1136/emermed-2020-211000.
[11] Wong A, Otles E, Donnelly JP, Krumm A, McCullough J, DeTroyer-Cooley O, et al. External validation of a widely implemented proprietary sepsis prediction model in hospitalized patients. JAMA Intern Med 2021;181(8):1065–70. https://doi.org/10.1001/jamainternmed.2021.2626.
[12] Akhlaghi H, Freeman S, Vari C, McKenna B, Braitberg G, Karro J, et al. Machine learning in clinical practice: Evaluation of an artificial intelligence tool after implementation. Emerg Med Australas 2024;36(1):118–24. https://doi.org/10.1111/1742-6723.14325.
[13] Collins GS, Moons KG, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD + AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024;385:e078378. https://doi.org/10.1136/bmj-2023-078378.
[14] Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;26(6):565–74. https://doi.org/10.1177/0272989X06295361.
[15] Zhao L, Leng Y, Hu Y, Xiao J, Li Q, Liu C, et al. Understanding decision curve analysis in clinical prediction model research. Postgrad Med J 2024;100(1185):512–5. https://doi.org/10.1093/postmj/qgae027.
[16] Alnawafleh KA, Almagharbeh WT, Alfanash HA, Alasmari AA, Alharbi AA, Alamrani MH, et al. Exploring the ethical dimensions of AI integration in nursing practice: a systematic review. J Nurs Regul 2025;16(3):228–37. https://doi.org/10.1016/j.jnr.2025.08.001.
[17] Nasir S, Khan RA, Bai S. Ethical framework for harnessing the power of AI in healthcare and beyond. IEEE Access 2024;12:31014–35. https://doi.org/10.1109/ACCESS.2024.3369912.
[18] World Health Organization. Ethics and governance of artificial intelligence for health: Guidance on large multi-modal models, https://www.who.int/publications/i/item/9789240084759; 2024 [accessed 15 February 2026].
[19] Kirubarajan A, Taher A, Khan S, Masood S. Artificial intelligence in emergency medicine: a scoping review. J Am Coll Emerg Physicians Open 2020;1(6):1691–702. https://doi.org/10.1002/emp2.12277.
[20] Martinez-Ortigosa A, Martinez-Granados A, Gil-Hernández E, Rodriguez-Arrastia M, Ropero-Padilla C, Roman P. Applications of artificial intelligence in nursing care: a systematic review. J Nurs Manag 2023;2023(1):3219127. https://doi.org/10.1155/2023/3219127.
[21] Arksey H, O'Malley L. Scoping studies: Towards a methodological framework. Int J Soc Res Methodol 2005;8(1):19–32. https://doi.org/10.1080/1364557032000119616.
[22] Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. Ann Intern Med 2018;169(7):467–73.
[23] Peters MDJ, Godfrey C, McInerney P, Munn Z, Tricco AC, Khalil, H. Scoping Reviews. 2020. Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. JBI Manual for Evidence Synthesis. JBI, 2024. Available from: https://synthesismanual.jbi.global.
[24] Hudson S, Nishat F, Stinson J, Litwin S, Zeller F, Wiles B, et al. Perspectives of healthcare providers to inform the design of an AI-enhanced social robot in the pediatric emergency department. Child 2023;10(9):1511. https://doi.org/10.3390/children10091511.

[25] Nishat F, Hudson S, Panesar P, Ali S, Litwin S, Zeller F, et al. Exploring the needs of children and caregivers to inform design of an artificial intelligence-enhanced social robot in the pediatric emergency department. J Clin Transl Sci 2023;7(1): e191. https://doi.org/10.1017/cts.2023.608.
[26] Jordan M, Hauser J, Cota S, Li H, Wolf L. The impact of cultural embeddedness on the implementation of an artificial intelligence program at triage: a qualitative study. J Transcult Nurs 2023;34(1):32–9. https://doi.org/10.1177/10436596221129226.
[27] Etu EE, Monplaisir L, Arslanturk S, Masoud S, Aguwa C, Markevych I, et al. Prediction of length of stay in the emergency department for COVID-19 patients: a machine learning approach. IEEE Access 2022;10:42243–51. https://doi.org/10.1109/ACCESS.2022.3168045.
[28] Brann F, Sterling NW, Frisch SO, Schrager JD. Sepsis prediction at emergency department triage using natural language processing: Retrospective cohort study. JMIR AI 2024;3:e49784. https://doi.org/10.2196/49784.
[29] Sterling NW, Brann F, Patzer RE, Di M, Koebbe M, Burke M, et al. Prediction of emergency department resource requirements during triage: an application of current natural language processing techniques. J Am Coll Emerg Physicians Open 2020;1(6):1676–83. https://doi.org/10.1002/emp2.12253.
[30] Tootooni MS, Pasupathy KS, Heaton HA, Clements CM, Sir MY. CCMapper: an adaptive NLP-based free-text chief complaint mapping algorithm. Comput Biol Med 2019;113:103398. https://doi.org/10.1016/j.compbiomed.2019.103398.
[31] Hall JN, Galaev R, Gavrilov M, Mondoux S. Development of a machine learning-based acuity score prediction model for virtual care settings. BMC Med Inf Decis Making 2023;23:200. https://doi.org/10.1186/s12911-023-02307-z.
[32] Ivanov O, Wolf L, Brecher D, Lewis E, Masek K, Montgomery K, et al. Improving ED emergency severity index acuity assignment using machine learning and clinical natural language processing. J Emerg Nurs 2021;47(2):265–78. https://doi.org/10.1016/j.jen.2020.11.001.
[33] Nasser L, McLeod SL, Hall JN. Evaluating the reliability of a remote acuity prediction tool in a Canadian academic emergency department. Ann Emerg Med 2024;83(4):373–9. https://doi.org/10.1016/j.annemergmed.2023.11.018.
[34] Lee H, Lee S, Kim H. Factors affecting the length of stay in the emergency department for critically Ill patients transferred to regional emergency medical center. Nurs Open 2023;10(5):3220–31. https://doi.org/10.1002/nop2.1573.
[35] Lee H, Yun H, Choi M, Kim H. Predicting workplace violence in the emergency department based on electronic health record data. J Emerg Nurs 2023;49(3): 415–24. https://doi.org/10.1016/j.jen.2023.01.010.
[36] Cho A, Min IK, Hong S, Chung HS, Lee HS, Kim JH. Effect of applying a real-time medical record input assistance system with voice artificial intelligence on triage task performance in the emergency department: prospective interventional study. JMIR Med Inform 2022;10(8):e39892. https://doi.org/10.2196/39892.
[37] Sung SF, Hung LC, Hu YH. Developing a stroke alert trigger for clinical decision support at emergency triage using machine learning. Int J Med Inf 2021;152: 104505. https://doi.org/10.1016/j.ijmedinf.2021.104505.
[38] Tsai DJ, Tsai SH, Chiang HH, Lee CC, Chen SJ. Development and validation of an artificial intelligence electrocardiogram recommendation system in the emergency department. J Pers Med 2022;12(5):700. https://doi.org/10.3390/jpm12050700.
[39] Choi SW, Ko T, Hong KJ, Kim KH. Machine learning-based prediction of Korean Triage and Acuity Scale level in emergency department patients. Healthc Inform Res 2019;25(4):305–12. https://doi.org/10.4258/hir.2019.25.4.305.
[40] Lin YT, Deng YX, Tsai CL, Huang CH, Fu LC. Interpretable deep learning system for identifying critical patients through the prediction of triage level, hospitalization, and length of stay: prospective study. JMIR Med Inform 2024;12:e48862. https://doi.org/10.2196/48862.
[41] Bagnasco A, Siri A, Aleo G, Rocco G, Sasso L. Applying artificial neural networks to predict communication risks in the emergency department. J Adv Nurs 2015;71 (10):2293–304. https://doi.org/10.1111/jan.12691.
[42] Kipourgos G, Tzenalis A, Koutsojannis C, Hatzilygeroudis I. An artificial intelligence based application for triage nurses in emergency department, using the emergency severity index protocol. Int J Caring Sci 2022;15(3):1764–72.
[43] Zaboli A, Brigo F, Sibilio S, Mian M, Turcato G. Human intelligence versus Chat-GPT: who performs better in correctly classifying patients in triage? Am J Emerg Med 2024;79:44–7. https://doi.org/10.1016/j.ajem.2024.02.008.
[44] Farahmand S, Shabestari O, Pakrah M, Hossein-Nejad H, Arbab M, Bagheri-Hariri S. Artificial intelligence-based triage for patients with acute abdominal pain in emergency department; a diagnostic accuracy study. Adv J Emerg Med 2017;1 (1):e5.
[45] Saban M, Dubovi I. A comparative vignette study: evaluating the potential role of a generative AI model in enhancing clinical decision-making in nursing. J Adv Nurs Epub 2024 Feb 17. https://doi.org/10.1111/jan.16101.
[46] Soufi MD, Samad-Soltani T, Vahdati SS, Rezaei-Hachesu P. Decision support system for triage management: a hybrid approach using rule-based reasoning and fuzzy logic. Int J Med Inf 2018;114:35–44. https://doi.org/10.1016/j.ijmedinf.2018.03.008.
[47] Hughes JA, Wu Y, Jones L, Douglas C, Brown N, Hazelwood S, et al. Analyzing pain patterns in the emergency department: Leveraging clinical text deep learning models for real-world insights. Int J Med Inf 2024;190:105544. https://doi.org/10.1016/j.ijmedinf.2024.105544.
[48] Podolsky S. 426 Predictive analytics improves emergency department staffing and efficiency. Proceedings of the Society for Academic Emergency Medicine Annual Meeting 2019; 2019 May 14–17; Nevada, USA: 2019. p. S164. https://doi.org/10.1111/acem.13756.
[49] Edgcomb J, Choi K, Tseng CH, Klomhaus AM, Ponce C, Perez L. et al. 6.74 Computable phenotyping of children with suicide-related emergencies using nursing triage safety screening and interventions. Proceedings of the 70th Annual Meeting of the American Academy of Child and Adolescent Psychiatry; 2023 October 23–28; New York, USA: 2023. p. S309.
[50] von Gerich H, Moen H, Block LJ, Chu CH, DeForest H, Hobensack M, et al. Artificial intelligence-based technologies in nursing: a scoping literature review of the evidence. Int J Nurs Stud 2022;127:104153. https://doi.org/10.1016/j.ijnurstu.2021.104153.
[51] Wahl B, Cossy-Gantner A, Germann S, Schwalbe NR. Artificial intelligence (AI) and global health: how can AI contribute to health in resource-poor settings? BMJ Glob Health 2018;3(4):e000798. https://doi.org/10.1136/bmjgh-2018-000798.
[52] Emergency Nurses Association. Emergency Severity Index Handbook. 5th ed. https://californiaena.org/wp-content/uploads/2023/05/ESI-Handbook-5th-Edition-3-2023.pdf; 2023. [accessed 10 July 2025].
[53] Maass W, Parsons J, Purao S, Storey VC, Woo C. Data-driven meets theory-driven research in the era of big data: Opportunities and challenges for information systems research. J Assoc Inf Syst 2018;19(12):1. https://doi.org/10.17705/1jais.00526.
[54] Chew BH, Ngiam KY. Artificial intelligence tool development: what clinicians need to know? BMC Med 2025;23:244. https://doi.org/10.1186/s12916-025-04076-0.
[55] Mukhamediev RI, Popova Y, Kuchin Y, Zaitseva E, Kalimoldayev A, Symagulov A, et al. Review of artificial intelligence and machine learning technologies: Classification, restrictions, opportunities and challenges. Math 2022;10(15):2552. https://doi.org/10.3390/math10152552.
[56] Adler-Milstein J, Jha AK. HITECH Act drove large gains in hospital electronic health record adoption. Health Aff 2017;36(8):1416–22. https://doi.org/10.1377/hlthaff.2016.1651.
[57] Rose C, Chen JH. Learning from the EHR to implement AI in healthcare. NPJ Digit Med 2024;7:330. https://doi.org/10.1038/s41746-024-01340-0.
[58] McDermott MB, Wang S, Marinsek N, Ranganath R, Foschini L, Ghassemi M. Reproducibility in machine learning for health research: still a ways to go. Sci Transl Med 2021;13(586):eabb1655. https://doi.org/10.1126/scitranslmed.abb1655.
[59] Lee S, Kim DW, Oh NE, Lee H, Park S, Yon DK, et al. External validation of an artificial intelligence model using clinical variables, including ICD-10 codes, for predicting in-hospital mortality among trauma patients: a multicenter retrospective cohort study. Sci Rep 2025;15:1100. https://doi.org/10.1038/s41598-025-85420-5.
[60] Moons KG, Damen JA, Kaul T, Hooft L, Navarro CA, Dhiman P, et al. PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. BMJ 2025;388:e082505. https://doi.org/10.1136/bmj-2024-082505.
[61] Khan AA, Badshah S, Liang P, Waseem M, Khan B, Ahmad A, et al. Ethics of AI: A systematic literature review of principles and challenges. Proceedings of the 26th International Conference on Evaluation and Assessment in Software Engineering; 2022 June 13–15; Gothenburg, Sweden: 2022. p. 383–92.