Tazeez Public Health Journal
tazeez تعزيز
November 2025; vol. 1; Issue 1

Systematic review article

The Use of Artificial Intelligence in Emergency Case Transport, Diagnosis, and Treatment

Khaled abdullaziz Alsunidy[1]*, Mazi Mohammed Alanazi[2], Majed Alawe Alotaibi[3], Ali Khalid Alsultan[4], Ibrahim Mukhtar Mohammed[5], Ahmad Mohammed Sahhaq[6], Rakan Mashl R Alharbi[7], Fahad Mohammed Obald Almehmad [8]

[1] Medical Director, Emergency Consultant, Administration Department, First Health Cluster, Riyadh, Saudi Arabia

[2] Saudi and Jordanian Board Emergency Medicine, Emergency Department, First Health Cluster, Riyadh, Saudi Arabia

[3] Emergency Consultant, Program Director Emergency Medicine, First Health Cluster, Riyadh, Saudi Arabia

[4] MRC-MOH, Medical Referrals Center - Ministry of Health

[5] Team leader Staff Nurse, Emergency Department, First Health Cluster, Riyadh, Saudi Arabia

[6] Staff nurse, Emergency Department, First Health Cluster, Riyadh, Saudi Arabia

[7] Staff Nurse, Emergency Department, First Health Cluster, Riyadh, Saudi Arabia

[8] Anesthesia technician, Surgical Operation Department, First Health Cluster, Riyadh, Saudi Arabia

Abstract

Artificial intelligence (AI) is rapidly entering prehospital emergency care, where time-critical triage, transport, and early treatment decisions determine outcomes. We systematically reviewed original studies evaluating AI tools used before hospital arrival, focusing on prediction/triage, diagnostic support, and transport optimization, and synthesized insights from contemporary reviews to contextualize clinical adoption. Seven original studies met inclusion for quantitative results synthesis: an ensemble waveform-based triage model predicting lifesaving interventions in trauma; an AI-enhanced regional platform guiding hospital selection and first aid; two studies on prehospital ST-elevation myocardial infarction (STEMI) detection (mini-12-lead and smartphone capture); a randomized trial of AI dispatcher alerts for out-of-hospital cardiac arrest; a gradient-boosted model for dyspnea serious adverse events; and a deep-learning severity algorithm predicting need for critical care in EMS. Across studies, AI frequently achieved AUCs around or above 0.80, improved sensitivity or operational timeliness (faster ECG interpretation/feedback), and in specific subgroups reduced adverse outcomes (lower mortality when AI guided optimal hospital transfer). However, not all trials showed clinical recognition gains despite superior model sensitivity, underscoring implementation challenges. Current reviews emphasize the promise of AI alongside the need for rigorous prospective validation, workflow integration, transparency, and equity. AI can augment prehospital decision-making, but robust clinical pathways and governance remain essential.

Keywords: prehospital emergency care; artificial intelligence; triage; transport; STEMI; dispatcher; dyspnea; critical care prediction

Published 29 November 2025

Alsunidy KA, Alanazi MM, Alotaibi MA, Alsultan AK, Mohammed IM, Sahhaq AM, et al. The Use of Artificial Intelligence in Emergency Case Transport, Diagnosis, and Treatment. Tazeez Public Health Journal. 2025;1(1). 34-41

Introduction

AI applications in prehospital care have accelerated, spanning triage/prognostication, dispatch optimization, diagnostic support (ECG), and multimodal monitoring. Recent scoping and systematic reviews identify over one hundred studies with AI often outperforming traditional tools or clinicians in predictive tasks, particularly triage/prognosis and cardiac arrest detection, while highlighting limited external/prospective validation and the need for explainability and workflow fit (Chee et al. 2023; Almulihi et al. 2024; El Arab et al. 2025). A 2025 systematic literature review similarly charts rapid growth since 2018 across dispatch, on-scene care, and transport decision-support, noting rising interest in large language models (LLMs) and multimodal data pipelines but persistent barriers in data linkage, privacy, and generalizability (Elfahim et al. 2025).

In low- and middle-income countries (LMICs), AI evaluations remain sparse, with most implementations in dispatch forecasting, classification, and disease prediction; deep learning predominates, and algorithms generally outperform conventional comparators, yet local sociotechnical adaptation and dataset completeness are crucial (Mallon et al. 2025). Horizon scanning from health-technology assessors echoes that prehospital AI is early in implementation, with promising pilots in call-taking (OHCA detection) and triage during surges, but more real-world trials are needed before broad deployment (Clark & Severn 2023).

Across reviews, common themes emerge: (1) AI can enhance prehospital triage accuracy and resource allocation; (2) ECG-based AI for STEMI and audio/NLP for dispatcher support are leading use cases; (3) external validation, calibration reporting, and transparent reporting (TRIPOD-AI/CONSORT-AI) remain inconsistent; and (4) integration into EMS workflows, training, and governance (bias, privacy, accountability) are preconditions for impact (Chee et al. 2023; Almulihi et al. 2024; El Arab et al. 2025; Elfahim et al. 2025; Clark & Severn 2023).

Against this backdrop, we synthesize seven original prehospital studies spanning trauma LSI prediction, regional AI transport orchestration, STEMI detection, dispatcher support for OHCA, dyspnea risk, and critical-care prediction. We aim to present performance, operational effects, and implementation signals, and to discuss implications using insights from contemporary reviews. Our focus is on the prehospital window where seconds matter and AI may translate most directly into lives saved through better triage, faster diagnosis, and optimized transport.

Methods

We conducted this systematic review following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. The aim was to evaluate studies that investigated the use of artificial intelligence in prehospital emergency care, including applications for transport decisions, diagnostic support, and treatment optimization.

Eligibility criteria

We included original research articles that reported on artificial intelligence or machine learning models applied in prehospital settings, such as dispatch centers, ambulance services, or on-scene emergency care. Studies were eligible if they evaluated performance outcomes, clinical impact, or operational efficiency. I excluded papers that focused solely on in-hospital applications, editorials, protocols, or articles without measurable outcomes.

Information sources and search strategy

We systematically gathered the relevant literature from peer-reviewed journals, covering studies published in the last decade. Searches were conducted in key databases including PubMed, Embase, Web of Science, and IEEE Xplore. To ensure comprehensiveness, I also screened the reference lists of identified studies. Only full-text articles in English were considered.

Study selection

Titles and abstracts were screened for relevance, followed by full-text review. Articles that clearly met the inclusion criteria were retained, while duplicates and unrelated reports were removed. The selection process was performed independently to minimize bias, and disagreements were resolved by re-evaluating the full text according to the predefined criteria.

Data extraction

From each included study, I extracted details on study design, setting, population, type of artificial intelligence model, input data (such as physiological signals, electrocardiograms, or dispatch records), comparators, and reported outcomes. Key performance measures such as sensitivity, specificity, area under the curve (AUC), predictive values, timeliness, and patient-centered outcomes were noted.

Data synthesis

Because of the diversity of study designs and outcomes, we synthesized the findings narratively rather than performing a meta-analysis. Two summary tables were developed to present study characteristics and key performance results. Where appropriate, we compared the findings with recognized clinical standards, existing triage tools, or expert performance.

Reporting

The methodology was designed and reported according to PRISMA standards to ensure transparency, reproducibility, and clarity. This process allowed me to provide a structured overview of the current evidence base regarding artificial intelligence in prehospital emergency care.

Results

Study characteristics

Table 1 summarizes seven original studies spanning North America, Europe, and Asia, covering dispatch-center audio/NLP inference, physiologic waveforms, 12-lead ECGs (portable and camera-captured), and structured prehospital data. Tasks included predicting lifesaving interventions (LSI), need for critical care/ICU, severe adverse events in dyspnea, STEMI detection, and enhancing dispatcher recognition of OHCA. Comparators ranged from human experts and standard call protocols to established triage tools (RETTS-A, NEWS2, ESI, KTAS). (Weidman et al. 2025; Kim et al. 2025; Chen et al. 2022; Lee et al. 2024; Blomberg et al. 2021; Kauppi et al. 2025; Kang et al. 2020)

Using 2-minute epochs immediately preceding interventions, Weidman et al. reported AUC 0.810 (95% CI 0.782–0.842), specificity 0.960, NPV 0.953 for overall LSI prediction, with comparable or better performance for subcategories (airway, transfusion, vasopressor). Performance remained robust up to 15 minutes before intervention, indicating early decompensation signatures in waveforms (Weidman et al. 2025). This demonstrates prehospital feasibility of high-frequency physiologic AI beyond static vital signs.

Kim et al. found mixed overall effects on transport delay outliers (>75th percentile): one region improved while the other worsened; however, prespecified subgroups benefited—patients with fever/respiratory symptoms had significantly fewer delays (36.5%→30.1%, P=.01), and when “real-time acceptance” signals were used, outliers fell (27.5%→19.6%, P=.02). Importantly, among system-guided “optimal hospital” transfers, ED mortality was lower (1.54% vs 0.64%, P=.01) (Kim et al. 2025). This suggests AI-enabled bed/procedure awareness and hospital selection can be outcome-relevant in defined pathways.

Two complementary studies evaluated field ECG AI. In Taiwan, AI feedback reached EMTs in 37.2 ± 11.3 s versus 113.2 ± 369.4 s for online physicians; model metrics were excellent (accuracy 0.992; sensitivity/recall 0.941; specificity 0.994; AUC 0.997), promptly identifying ten STEMI patients who underwent PPCI with median contact-to-door time 18.5 min (IQR 16–20.8) (Chen et al. 2022). In Korea, the smartphone “qSTEMI” biomarker derived from printed ECGs achieved AUC 0.815 (0.691–0.938), sensitivity 0.750, specificity 0.862 and was non-inferior to expert consensus (AUC 0.736) (Lee et al. 2024). Collectively, these show that both sensor-

Table 1. Included studies settings, tasks, inputs, comparators, and samples.

Study (year)	Setting & design	AI task & inputs	Comparator	Sample/episodes
Weidman et al. (2025)	US critical-care air transport; retrospective cohort	Predict LSI during care using ensemble ML on continuous physiologic waveforms (ECG, PPG, $EtCO_2$, BP)	N/A (model metrics vs triage goals)	2,809 patients; 15,088 2-min epochs; 910 LSI epochs
Kim et al. (2025)	Korea; community, non-randomized 2×16-week periods in 2 regions	AI platform (CONNECT-AI): first-aid guidance, critical-illness prediction, optimal hospital recommendation; 5G/IoT data + live video	Conventional practice (control periods)	14,853 ambulance transports
Chen et al. (2022)	Taiwan; implementation study	Real-time AI STEMI detection on prehospital mini-12-lead ECG; CNN-LSTM; response time to EMTs	Remote online physicians	275 patients; 362 ECGs (AI sites) + 335 ECGs (non-AI sites)
Lee et al. (2024)	Korea; diagnostic study	Smartphone AI extracting STEMI biomarker from printed ECG images	Consensus of 5 EMS directors + 3 interventional cardiologists	53 patients (24 STEMI)

Blomberg et al. (2021)	Denmark; double-masked RCT at EMS dispatch	Real-time ML alerts for suspected OHCA during 112 calls	Standard dispatcher protocol (no alert)	169,049 calls screened; 5,242 randomized alerts; 654 confirmed OHCAs
Kauppi et al. (2025)	Sweden; retrospective	Predict serious adverse events in dyspnea using gradient boosting vs RETTS-A/NEWS2	RETTS-A, NEWS2	6,354 EMS missions (dyspnea primary symptom)
Kang et al. (2020)	Korea; dev+external validation	Deep-learning algorithm to predict need for critical care using age, sex, chief complaint, onset-to-arrival, trauma, initial vitals	ESI, KTAS, NEWS, MEWS	Dev: 8,981,181 ED visits; External EMS run-sheets: 2,604

Table 2. Key outcomes and performance metrics.

Study	Primary outcome(s)	Key results
Weidman et al. 2025	Predict LSI within 2-min epochs	AUC 0.810; spec 0.960; NPV 0.953; robust up to 15 min pre-LSI
Kim et al. 2025	Transport delay outliers; mortality	Mixed overall; fewer outliers in fever/respiratory (36.5%→30.1%, P=.01); fewer outliers with acceptance signals (27.5%→19.6%, P=.02); lower mortality with “optimal hospital” routing (1.54%→0.64%, P=.01)
Chen et al. 2022	STEMI detection; feedback time	AUC 0.997; sens 0.941; spec 0.994; EMT feedback 37.2 s vs physicians 113.2 s; 10 PPCI cases, median contact-to-door 18.5 min

Lee et al. 2024	STEMI from printed ECG images	AUC 0.815 vs experts 0.736 (non-inferior); sens 0.750; spec 0.862
Blomberg et al. 2021	Dispatcher OHCA recognition	No significant improvement with AI alert (93.1% vs 90.5%); AI sensitivity higher than dispatchers (85.0% vs 77.5%)
Kauppi et al. 2025	Dyspnea SAE prediction	AUC 0.81 vs RETTS-A 0.73, NEWS2 0.73; better calibration and sensitivity
Kang et al. 2020	Need for critical care	AUC 0.867; > ESI 0.839; KTAS 0.824; NEWS 0.741; MEWS 0.696; external EMS validation

native and image-based ECG AI can accelerate triage and meet expert-level accuracy in the field.

In a double-masked RCT, AI alerts did not significantly increase dispatcher recognition among confirmed OHCA calls (93.1% vs 90.5%, P=.15), despite the AI's higher sensitivity than dispatchers alone (85.0% vs 77.5%, P<.001) and faster early identification in prior observational work (Blomberg et al. 2021). This gap between model capability and clinical effect underscores human-factors and integration challenges at dispatch.

Gradient boosting improved discrimination for SAE (AUC 0.81, 95% CI 0.78–0.84) compared with RETTS-A (0.73, 0.70–0.76) and NEWS2 (0.73, 0.70–0.76), with better calibration and sensitivity (Kauppi et al. 2025). Given dyspnea's high 30-day mortality risk, enhanced risk stratification may better direct transport priority and pre-alert receiving teams.

Predicting need for critical care. A national-scale deep-learning model predicted critical-care needs with AUC 0.867 (0.864–0.871), outperforming ESI (0.839), KTAS (0.824), NEWS (0.741), and MEWS (0.696); external validation on EMS run-sheets confirmed strong discrimination (Kang et al. 2020). Such tools can guide bypass to higher-acuity centers and resource activation.

The seven studies show consistent model-level accuracy (AUCs around/above 0.80) and notable operational gains in specific contexts (faster AI ECG reads, subgroup mortality benefit with AI-guided routing). The RCT at dispatch illustrates that human-system interaction can limit realized impact, despite AI's superior sensitivity. Models leveraging high-resolution physiologic signals (waveforms) and tailored disease-specific features (ECG biomarkers) perform strongly, aligning with review-level observations that AI excels in prehospital prognostication and cardiac use cases (Chee et al. 2023; Elfahim et al. 2025).

Discussion

This synthesis supports three practical messages. First, AI can enhance early recognition and risk stratification in the field. Waveform-based triage predicted imminent LSIs, while dyspnea and global severity models outperformed conventional triage scores, echoing review findings that AI frequently surpasses non-AI tools for prehospital prognostication (Chee et al. 2023; Almulihi et al. 2024; El Arab et al. 2025). Second, diagnostic acceleration is feasible: prehospital ECG AI achieved expert-level STEMI performance and materially shortened interpretation/feedback time, which plausibly compresses reperfusion pathways—an archetype of AI's value where seconds matter. Third, system-level orchestration (bed/procedure awareness, hospital acceptance, routing) can translate into fewer delays and, in targeted groups, lower mortality—consistent with horizon scanning that identifies dispatch and in-ambulance decision support as early high-yield domains (Clark & Severn 2023).

Implementation determines impact. The OHCA RCT shows that adding alerts does not guarantee higher recognition, a reminder from the reviews that workflow integration, trust, alert design, dispatcher training, and organizational readiness are essential for AI to change outcomes (Chee et al. 2023; Elfahim et al. 2025). The CONNECT-AI mixed results in regions highlight context sensitivity: benefits depend on reliable hospital acceptance signals, communication infrastructure, and adherence to AI recommendations. Reviews in LMIC contexts stress that data completeness, infrastructure, and sociocultural tailoring are prerequisites; when addressed, AI typically outperforms conventional comparators but must be locally validated (Mallon et al. 2025).

Methodological considerations from the review literature apply here: external validation is uncommon, calibration is under-reported, and prospective/multi-center trials remain limited (Chee et al. 2023; El Arab et al. 2025). The included studies partially address this (an RCT at dispatch; external EMS validation for a critical-care model), but broader uptake will require TRIPOD-AI/CONSORT-AI-aligned reporting, bias audits, and health-economic evaluation.

Equity and governance are also central. As AI expands to audio (call centers), images (printed ECGs), and high-frequency signals, datasets must represent diverse accents, devices, and pathophysiology to avoid performance gaps. Explainable interfaces may support trust for paramedics and dispatchers, as recommended across reviews (Chee et al. 2023; El Arab et al. 2025; Almulihi et al. 2024). The emergence of LLMs could enhance documentation, checklists, and protocol adherence, but rigorous guardrails are needed for reliability in high-stakes settings (Elfahim et al. 2025; Clark & Severn 2023).

Services considering prehospital AI should prioritize (1) validated, high-signal tasks (STEMI, waveform-based decompensation, critical-care prediction); (2) strong socio-technical integration (training, interface design, escalation paths); (3) local pilots with outcome tracking; and (4) governance frameworks spanning bias, privacy, and accountability.

Conclusion

Across heterogeneous prehospital settings, AI tools show strong discrimination for triage/prognosis and disease-specific diagnosis, and, when well integrated, improved operational timeliness and select patient outcomes. Yet clinical impact hinges on workflow fit, reliable data flows, and rigorous validation. The path forward is purposeful deployment where time-critical decisions and high-fidelity signals meet robust integration, STEMI ECG AI, waveform-based decompensation prediction, and critical-care routing, accompanied by prospective evaluation, transparency, and governance. With these guardrails, AI can meaningfully augment prehospital transport, diagnosis, and treatment to improve patient outcomes.

References

Almulihi QA, Alquraini AA, Almulihi FAA, et al. AI/ML in Emergency Medicine Triage—Systematic Review. *Med Arch.* 2024;78(3):198-206.

Blomberg SN, Christensen HC, Lippert F, et al. Effect of ML on Dispatcher Recognition of OHCA: Randomized Clinical Trial. *JAMA Netw Open.* 2021;4(1):e2032320.

Chee ML, Chee ML, Huang H, et al. AI and ML in prehospital emergency care: Scoping Review. *iScience.* 2023;26:107407.

Chen K-W, Wang Y-C, Liu M-H, et al. AI-assisted remote detection of STEMI using a mini-12-lead ECG device in prehospital ambulance care. *Front Cardiovasc Med.* 2022;9:1001982.

Clark M, Severn M. CADTH Horizon Scan: AI in Prehospital Emergency Health Care. *Can J Health Technol.* 2023;3(8).

El Arab RA, Al Moosa OA. Role of AI in ED triage: Integrative systematic review. *Intensive Crit Care Nurs.* 2025;104058.

Elfahim O, Edjinedja KL, Cossus J, et al. Systematic Literature Review of AI in Prehospital Emergency Care. *Big Data Cogn Comput.* 2025;9:219.

Kang D-Y, Cho K-J, Kwon O, et al. AI algorithm to predict need for critical care in EMS. *Scand J Trauma Resusc Emerg Med.* 2020;28:17.

Kauppi W, Imberg H, Herlitz J, et al. ML decision support for prehospital dyspnea assessment. *BMC Emerg Med.* 2025;25:2.

Kim JH, Kim MJ, Kim HC, et al. A Novel AI–Enhanced Digital Network for Prehospital Emergency Support (CONNECT-AI): Community Intervention Study. *J Med Internet Res.* 2025;27:e58177.

Lee SH, Hong WP, Kim J, et al. Smartphone AI vs Medical Experts for Prehospital STEMI Diagnosis. *Yonsei Med J.* 2024;65(3):174-180.

Mallon O, Lippert F, Stassen W, et al. AI in prehospital emergency care systems in LMICs: Scoping Review. *Front Public Health.* 2025;13:1604231.

Weidman AC, Malakouti S, Salcido DD, et al. A Machine Learning Trauma Triage Model for Critical Care Transport. *JAMA Netw Open.* 2025;8(6):e259639.