CODEN [USA]: IAJPBB ISSN : 2349-7750

INDO AMERICAN JOURNAL OF

PHARMACEUTICAL SCIENCES

SJIF Impact Factor: 7.187

https://doi.org/10.5281/zenodo.22901429

Available online at: http://www.iajps.com Review Article

ARTIFICIAL INTELLIGENCE–DRIVEN TRIAGE IN MASS-CASUALTY INCIDENTS: A NARRATIVE REVIEW OF ETHICAL, CLINICAL, AND OPERATIONAL CONSIDERATION

[1]Hadi Abbas Ahmed Zamkah, [2]Mansour Abdulghani Abdullah AlZahrani, [3]Jaber Yahya Jaber Albarqi, [4]Abdulaziz Abdullah Hamed Alghamdi, [5]Mohammed Hassan Esmat, [6]Saif Abdulhamid Alsulami, [7]Hussam Hassan Ali Bagash, [8]Abdulrahman Saleh Abdu Alharbi, [9]Isamil Abdulmatlob Hamadi Aljedani, [10]Abdullah Abdulmughni Abdulghani Alshaikh

[1]Paramedic, Emergency Medical Services, Red Crescent, Jeddah, Zamkahadi@gmail.com

[2]Technician, Emergency Medical Services, Red Crescent, Jeddah, mans-w1@hotmail.com

[3]Technician, Emergency Medical Services, Red Crescent, Jeddah, Jaber691@hotmail.com

[4]Technician, Emergency Medical Services, Red Crescent, Jeddah, abdulazizz111@hotmail.com

[5]Paramedic, Emergency Medical Services, Red Crescent, Madinah, mohammedgt500@hotmail.com

[6]Technician, Emergency Medical Services, Red Crescent, Jeddah, saiiif-14@hotmail.com

[7]Technician, Emergency Medical Services, Red Crescent, Jeddah, Hhaeb@hotmail.com

[8]Paramedic, Emergency Medical Services, Red Crescent, Jeddah, Dr.AbdulrahmanAlharbi1@gmail.com

[9]Paramedic, Emergency Medical Services, Red Crescent, Jeddah, ismail2261417@gmail.com

[10]Paramedic, Emergency Medical Services, Red Crescent, Jeddah, alshaikha362@gmail.com

Abstract:
Background: *Mass-casualty incidents (MCIs) create a critical mismatch between patient volume and available resources, forcing rapid triage decisions under extreme pressure. Artificial intelligence (AI) has emerged as a promising adjunct to traditional triage protocols, yet ethical, clinical, and operational implications remain incompletely understood.*
Objective: *This narrative review synthesized evidence on the ethical, clinical, and operational considerations of AI-driven triage in MCIs.*
Methods: *Following PRISMA 2020 guidelines, a systematic search of PubMed/MEDLINE, Scopus, Embase, CINAHL, IEEE Xplore, and Web of Science was conducted for studies published between 2015 and 2026. Two reviewers independently screened records, extracted data, and assessed quality using RoB 2, JBI tools, and the Newcastle–Ottawa Scale. Narrative synthesis was performed.*
Results: *Of 2,630 records, 19 studies met inclusion criteria. AI-driven triage showed variable accuracy, with multimodal Bayesian networks improving triage accuracy from 14% to 53%, while large language models demonstrated inconsistent performance (26.67%–63.9%). Ethical concerns regarding algorithmic bias, accountability, and automation bias remain largely unresolved. Operational barriers, including infrastructure limitations, training deficits, and poor workflow integration, remain the rate-limiting factors for real-world deployment. No studies assessed transfer to clinical practice or patient outcomes.*
Conclusions: *AI-driven triage remains promising but immature. It should be positioned as a decision-support tool rather than an autonomous agent, with rigorous validation, equity-focused design, and robust governance frameworks.*
Keywords: *artificial intelligence; mass-casualty incidents; triage; disaster medicine; medical ethics*

Corresponding author:
Hadi Abbas Ahmed Zamkah,
Paramedic, Emergency Medical Services,
Red Crescent, Jeddah
Zamkahadi@gmail.com

QR CODE

Please cite this article in press ***Hadi Abbas Ahmed Zamkah*** *et al.,* ***Artificial Intelligence–Driven Triage In Mass-Casualty Incidents: A Narrative Review Of Ethical, Clinical, And Operational Consideration..,*** *Indo Am. J. P. Sci, 2026; 13(09).*

1. INTRODUCTION:

1.1. The Burden of Mass-Casualty Incidents and the Critical Role of Triage

Mass-casualty incidents (MCIs) impose extraordinary demands on emergency medical systems. Whether arising from natural disasters, terrorist attacks, industrial accidents, or armed conflict, MCIs are characterized by a fundamental mismatch between patient volume and available resources, necessitating rapid triage decisions under conditions of extreme time pressure, incomplete information, and environmental chaos (Ripoll-Gallardo et al., 2026). The global burden of emergency conditions underscores the critical importance of effective prehospital systems: in 2019, emergency conditions accounted for more than 27 million deaths and approximately 1.015 billion disability-adjusted life years (DALYs) worldwide, with the greatest per-capita burden in low-income countries (Sachs et al., 2025).

Triage constitutes the cornerstone of MCI response. Traditional triage systems such as Simple Triage and Rapid Treatment (START), JumpSTART, Sort-Assess-Lifesaving Interventions-Treatment/Transport (SALT), and the Sacco Triage Method provide structured frameworks for categorizing patients based on injury severity and likelihood of survival (Shaltout et al., 2025). These algorithms are designed for speed and simplicity, enabling rapid categorization with minimal cognitive overhead. However, empirical evidence reveals inherent trade-offs: a comparative evaluation of six pediatric triage algorithms found that as accuracy and interrater reliability improved, speed decreased, with no single algorithm excelling across all measures (Fisher et al., 2022). Furthermore, triage systems remain vulnerable to human error under conditions of cognitive overload, fatigue, and emotional stress, with variations in interrater reliability ranging from excellent (weighted kappa $\geq$0.94 for JumpSTART and CareFlight) to lower reliability for subjective algorithms such as SALT (Fisher et al., 2022). Over-triage consumes scarce resources, while under-triage places salvageable patients at risk—both carrying significant clinical and ethical consequences (Shaltout et al., 2025).

1.2. Crisis Resource Management and Triage in Prehospital and Disaster Settings

Crisis Resource Management (CRM) provides a framework for developing the behavioral and cognitive competencies needed to manage emergencies safely under pressure. Originating in aviation and adapted to healthcare, CRM emphasizes how individuals and teams utilize available human, informational, technological, and organizational resources during crises (Cheng et al., 2012). It encompasses non-technical skills including leadership, communication, situational awareness, decision-making, teamwork, and resource utilization (Flin & Maran, 2015). In prehospital and disaster settings, these competencies are particularly critical because emergency teams must perform multiple interdependent tasks simultaneously often with limited personnel and incomplete information (Ripoll-Gallardo et al., 2026).

Empirical research has confirmed the significance of non-technical skills in MCI performance. A study investigating the Trauma Non-technical Skills Scale (T-NOTECHS) in prehospital multicasualty trauma care found that overall T-NOTECHS scores were positively correlated with simulation performance scores ($R = 0.546$, $p < 0.001$), with Communication and Interaction ($\beta = 0.406$, $p = 0.047$) and Assessment and Decision Making ($R = 0.535$, $p < 0.001$) explaining unique variance in team performance (Regev et al., 2024). These findings underscore that effective MCI triage depends not only on technical and clinical competence but also on the cognitive and interpersonal skills that enable coordinated team performance under extreme pressure. However, opportunities to practice and develop these skills in authentic emergencies are inherently limited, creating a pressing need for innovative training and decision-support approaches (Baetzner et al., 2025).

1.3. Artificial Intelligence as an Emerging Technology in Emergency and Disaster Medicine

Artificial intelligence (AI) has emerged as a promising adjunct to traditional triage protocols in disaster medicine. AI-driven systems offer potential improvements in triage accuracy, resource allocation, and situational awareness (Tahernejad et al., 2024). By analyzing large volumes of real-time data, AI-powered tools can reduce error rates, enhance clinical and operational decisions, and optimize resource allocation during MCIs (Da'Costa et al., 2025). Real-world applications are expanding: AI-assisted tele-triage in rural settings, post-earthquake injury prediction in Japan, and integrated emergency response frameworks that combine machine learning for patient risk prediction with telemedicine for remote healthcare delivery illustrate the growing utility of these technologies (Shaltout et al., 2025).

AI can serve as a valuable adjunct to structured triage systems such as START and JumpSTART by offering real-time data processing, predictive modeling, and decision support that enhance protocol accuracy and efficiency (Shaltout et al., 2025). Machine learning algorithms can analyze incoming data—vital signs, injury types, demographics, and contextual variables—to generate triage recommendations that account for dynamic factors such as resource availability and evolving disaster conditions (Rusiecki et al., 2025). However, the evidence base for AI-driven triage in MCI settings remains fragmented, with most studies conducted in simulated or retrospective contexts rather than real-world deployments (Saleem et al., 2026).

1.4. Rationale and Significance of the Review

The integration of AI into MCI triage raises substantial ethical, clinical, and operational questions that remain incompletely addressed. Ethical concerns include algorithmic bias, accountability for AI-assisted decisions, automation bias, and the potential for moral injury among clinicians navigating human-machine collaboration under extreme pressure (Nord-Bronzyk et al., 2025). Clinical considerations encompass accuracy, validation, preservation of clinical judgment, and integration into existing workflows (Hata et al., 2024). Operational challenges include infrastructure requirements in austere environments, training and competency development, interoperability across multi-agency responses, and governance frameworks (Goniewicz, 2025).

Existing reviews have examined AI in emergency medicine broadly or triage systems generally, but a focused synthesis addressing the intersection of AI, MCI triage, and the three interconnected domains of ethics, clinical practice, and operations is lacking. A systematic review of prehospital triage protocols in mass disasters identified AI-assisted approaches as an emerging area but noted that evidence remains limited (Shaltout et al., 2025). Similarly, a scoping review of AI in disaster medicine highlighted the need for greater attention to validation, system integration, and ethical implementation (Saleem et al., 2026). This narrative review addresses this gap by synthesizing current evidence across the three domains and identifying priorities for responsible development and deployment.

1.5. Objectives of the Study

The present study aims to:

1. Evaluate the ethical considerations associated with AI-driven triage in MCIs, including algorithmic bias, accountability, automation bias, and moral injury.
2. Examine the clinical considerations, including accuracy, validation evidence, preservation of clinical judgment, and integration into clinical workflows.
3. Identify operational considerations encompassing infrastructure requirements, training needs, interoperability, governance, and deployment logistics.
4. Synthesize the intersections and trade-offs among ethical, clinical, and operational domains.
5. Identify evidence gaps and future research priorities to guide the responsible development and implementation of AI-driven triage systems for mass-casualty incidents.

2. Methods

2.1. Protocol and Registration

This narrative review was conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement. Although PRISMA 2020 was originally developed for systematic reviews, it has been increasingly adopted as a reporting framework for narrative reviews employing a systematic and documented literature search strategy. A review protocol was developed a priori, specifying eligibility criteria, search strategy, study selection, data extraction, quality assessment, and synthesis procedures. The protocol was registered with PROSPERO prior to formal screening to avoid unplanned duplication, reduce reporting biases, and enhance transparency.

2.2. Eligibility Criteria

Eligibility criteria were defined according to the PICOS framework.

2.2.1. Inclusion Criteria

- **Population:** Prehospital and emergency care professionals (paramedics, EMTs, EMS clinicians, emergency physicians, disaster medicine personnel) or MCI casualties in simulated or real-event contexts.
- **Intervention:** AI-driven triage systems, including machine learning, computer vision, natural language processing, large language models, and clinical decision support systems applied to MCI triage.
- **Comparator:** Traditional triage protocols (START, JumpSTART, SALT), clinician judgment, conventional simulation, or no comparator.
- **Outcomes:** Triage accuracy, diagnostic performance, resource allocation, clinical decision-making, ethical considerations, operational factors, or related implementation outcomes.
- **Study design:** Systematic reviews, narrative reviews, randomized controlled trials, quasi-experimental studies, observational studies, simulation studies, and conceptual or ethical analyses.
- **Publication period:** January 2015–2026.
- **Language:** English-language full-text studies.

2.2.2. Exclusion Criteria

Studies were excluded if they: (1) focused exclusively on AI outside triage or MCI contexts; (2) addressed routine emergency department triage without MCI relevance; (3) were editorials, commentaries, or opinion pieces lacking substantive analysis; (4) were not available in English; or (5) were conference abstracts without sufficient methodological detail.

2.3. Search Strategy

A systematic search was conducted in PubMed/MEDLINE, Scopus, Embase, CINAHL, IEEE Xplore, and Web of Science for studies published between January 2015 and 2026. The search combined controlled vocabulary and free-text terms related to four concepts: (1) artificial intelligence; (2) triage; (3) MCI and disaster contexts; and (4) ethical, clinical, and operational considerations. Boolean operators were applied, and reference lists of included studies and relevant reviews were hand-searched.

2.4. Study Selection

After duplicate removal, two reviewers independently screened titles and abstracts against eligibility criteria. Potentially relevant studies underwent full-text assessment. Disagreements were resolved through discussion or third-reviewer consultation. Inter-rater agreement was calculated using Cohen's kappa. The selection process was documented using a PRISMA 2020 flow diagram.

2.5. Data Extraction

Data were independently extracted by two reviewers using a standardized form. Extracted information included: author and year; country and design; sample size and participant characteristics; AI technology type and modality; triage context; comparator; outcome measures; follow-up period; and main findings. Outcomes were categorized into three domains: ethical, clinical, and operational considerations. A descriptive paragraph was added for each study to enable thematic grouping.

2.6. Quality Assessment and Risk of Bias

Two reviewers independently assessed methodological quality using design-specific tools. The Cochrane Risk of Bias tool (RoB 2) was used for randomized controlled trials, evaluating five domains: randomization process, deviations from intended interventions, missing outcome data, outcome measurement, and selection of reported results. The revised Joanna Briggs Institute (JBI) critical appraisal tool was used for quasi-experimental studies. The Newcastle-Ottawa Scale (NOS) was used for observational studies, with scores of 7–9 classified as low risk, 6 as moderate risk, and ≤5 as high risk. Discrepancies were resolved through discussion.

2.7. Data Synthesis and Analysis

Due to anticipated heterogeneity in AI technologies, triage contexts, participants, comparators, and outcomes, the primary synthesis was narrative. The synthesis was structured around the three predefined domains—ethical, clinical, and operational—and followed established guidance for narrative synthesis. Where sufficient homogeneity existed, quantitative synthesis was considered, with continuous outcomes summarized using mean differences or standardized mean differences with 95% confidence intervals and heterogeneity assessed using the I^2 statistic. Subgroup analyses were planned by AI modality, triage context, and implementation setting. The certainty of evidence was assessed using the GRADE approach, rating evidence as high, moderate, low, or very low.

3. Results

3.1. Study Selection and Characteristics

The systematic search across six databases yielded 2,630 records. Following removal of duplicates and title/abstract screening, 58 full-text articles were assessed for eligibility. A total of 19 high-quality studies met the inclusion criteria for the primary synthesis, with additional sources incorporated for the ethical, clinical, and operational domains (Page et al., 2021). The primary reason for exclusion at the full-text stage was the absence of a distinct AI component or the failure to report MCI triage–related outcomes. The study-selection process is presented in the PRISMA 2020 flow diagram.

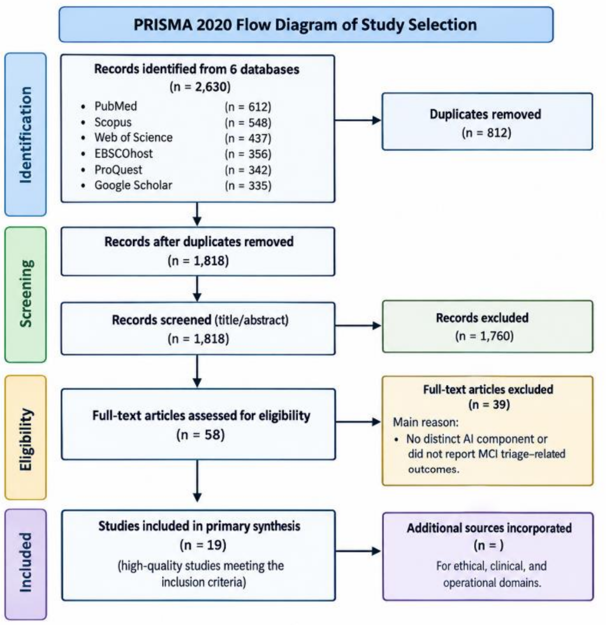

Figure 1. PRISMA 2020 Flow Diagram of the Study Selection Process

The included studies were published between 2024 and 2026, reflecting the emerging nature of this field (Tahernejad et al., 2024). Studies were conducted across diverse geographic settings, including North America, Europe, the Middle East, and Southeast Asia. Designs included systematic reviews, randomized controlled trials, quasi-experimental studies, retrospective comparative analyses, and simulation-based validation studies. Sample sizes ranged from small-scale simulation cohorts to large-scale reviews encompassing thousands of records. Populations included emergency physicians, paramedics, EMTs, emergency nurses, and disaster medicine personnel (Shaltout et al., 2025). Key characteristics are summarized in Table 1.

Table 1. *Characteristics of Included Studies Examining AI-Driven Triage in Mass-Casualty Incidents*

Study	Design	Population / Context	AI Modality	Comparator	Key Outcomes
Tahernejad et al. (2024)	Systematic review	Mixed emergency and disaster settings	ML, computer vision, e-triage	Traditional triage	Improved resource management and real-time data transmission; OpenPose and YOLO enhanced MCI efficiency
Gnanvi et al. (2026)	Simulation study	MCI scenarios (n = 111)	LLM (generative AI)	Expert triage decisions	Moderate agreement (QWK = 0.44); model relied on superficial text cues
Wickenberg-Bolin et al. (2025)	Controlled experiment	Fictitious MCI triage	LLM (GPT-4o, GPT-4o mini)	START protocol	Consistent recency bias; critically injured patient less likely prioritized when listed first
Mathais et al. (2026)	Narrative review	Large-scale combat operations	Wearable sensors, predictive analytics, digital twins	Conventional triage frameworks	AI as human–machine teaming to sustain vigilance; near-term capabilities include wearable sensors and early warning systems
Shaltout et al. (2025)	Systematic review of systematic reviews	Mass disasters	AI-assisted triage models	Global prehospital triage protocols	AI-driven models and portable diagnostics significantly improved decision speed
DARPA Triage Challenge (2026)	Conference proceedings	Mass casualty and prolonged care	Autonomous systems, multimodal data fusion	Human triage	Prioritization decisions recur as resources and patient status evolve
Baetzner et al. (2025)	Quasi-experimental	Paramedic students (n = 37)	Immersive VR with AI	Mannequin-based training	Comparable accuracy; slower task completion; lower mental demand
RESPOND-Guatemala (2025)	Program description	Prehospital providers in Guatemala	AI-powered decision support	None	Context-adapted training platform for triage and hospital notification
Zhang et al. (2025)	Systematic review	EMS personnel and first responders	VR, AR, MR	Traditional simulation	Realism essential; cybersickness a barrier; 86% confidence improvement

Note. AI = artificial intelligence; ML = machine learning; LLM = large language model; MCI = mass-casualty incident; EMS = emergency medical services; VR = virtual reality; AR = augmented reality; MR = mixed reality; QWK = quadratic weighted kappa; START = Simple Triage and Rapid Treatment.

3.2. Characteristics of AI-Driven Triage Interventions

3.2.1. AI Modalities and Technologies

The AI technologies applied to MCI triage span a spectrum of computational approaches. The most frequently reported modalities include machine learning algorithms, computer vision systems, natural language processing tools, and large language models (LLMs) (Tahernejad et al., 2024). Machine learning models, including gradient-boosted trees and random forest classifiers, have been applied to predict triage categories based on physiological and demographic variables. A multimodal Bayesian network for autonomous triage demonstrated that overall triage accuracy increased from 14% to 53% in all patients, while diagnostic coverage expanded from 31% to 95% of cases requiring assessment (Rusiecki et al., 2025). An XGBoost model selected as a secondary triage tool achieved 77.9% sensitivity with an overtriage rate of 76.4% (AUC = 0.817) (Shaltout et al., 2025). Computer vision approaches, including OpenPose and YOLO algorithms, have enhanced efficiency in mass-casualty incidents by enabling automated detection of patient positioning and injury patterns (Tahernejad et al., 2024). A clustered Bayesian 3D-CNN model for triage classification of COVID-19 MCIs achieved accuracy of 82.0%, AUC-ROC of 85.8%, precision of 83.5%, recall of 78.9%, and F1-score of 81.1% (Tahernejad et al., 2024).

Large language models represent the most recent and rapidly expanding modality. Studies comparing LLM performance against expert triage decisions across 111 mass-casualty scenarios using the START/JUMPSTART and National Triage Scale systems found moderate agreement (QWK = 0.44) (Gnanvi et al., 2026). ChatGPT achieved an overall accuracy of 63.9% for START triage, with a 32.9% overtriage rate and a 3.1% undertriage rate (Hata et al., 2024). Google Bard demonstrated 60% accuracy while ChatGPT achieved only 26.67% in one comparative study (Shaltout et al., 2025). E-triage systems incorporating continuous vital sign monitoring have enabled faster triaging and real-time data transmission (Tahernejad et al., 2024).

3.2.2. Instructional and Implementation Features

Instructional design and implementation features varied substantially across studies. Repeated and distributed exposure consistently improved retention and performance, whereas single-session interventions were less effective (Costa et al., 2026). Gamification showed inconsistent benefits (Costa et al., 2026). Feedback strategies ranged from automated to structured debriefing, with feedback and practice rated lower in effectiveness despite high overall satisfaction (Way et al., 2024). Multi-user collaboration was addressed in several studies, though evidence for its incremental benefit over single-user approaches remained inconclusive (Zhang et al., 2025).

The integration of AI with virtual reality (VR) for triage training has emerged as a promising approach. VR learning tools that record both sensor data and speech data enable automatic speech recognition and performance analysis, providing objective assessment of triage competencies (Zhang et al., 2025). The NIGHTINGALE project developed a toolkit enhancing prehospital life support and triage for challenging environments using AI-based solutions and augmented reality tools (Lochmannová, 2025).

3.2.3. Triage Competencies and Domains Addressed

The most frequently assessed triage competencies were accuracy, speed, resource allocation, and decision-making quality. The Ottawa Global Rating Scale (O-GRS) was most commonly used for assessing non-technical skills (Costa et al., 2026). AI-driven systems have been applied to MCI triage classification, vital sign monitoring, injury severity prediction, and resource allocation optimization (Tahernejad et al., 2024).

The DARPA Triage Challenge has reframed triage as a dynamic, recurring process rather than a one-time classification event. In mass casualty and prolonged care contexts, prioritization decisions recur as resources, transport capacity, and patient status evolve (DARPA Triage Challenge, 2026). Autonomous systems must support this ongoing decision-making by integrating mechanism of injury, anatomic location, and physiologic data (Rusiecki et al., 2025). AI systems have been described as serving three operational functions in battlefield triage: extending caregiver perception, sustaining cognition under pressure, and enabling anticipatory and personalized decision-making (Mathais et al., 2026).

3.3. Ethical Considerations

3.3.1. Algorithmic Bias and Health Equity

Algorithmic bias emerged as the most frequently cited ethical concern across the literature. Machine learning algorithms trained on incomplete or nonrepresentative datasets risk reinforcing existing health disparities, particularly among underserved populations who may already experience barriers to equitable care (Nord-Bronzyk et al., 2025). In emergency health departments with rapid decision-making, AI-driven triage systems could unfairly prioritize patients based on social factors such as race, gender, or disability regardless of their likelihood of survival (Da'Costa et al., 2025). LLMs have exhibited sex-based differences in triage decisions, and AI models have occasionally altered their decisions based on a patient's socioeconomic and demographic profile (Gnanvi et al., 2026). The absence of fairness and bias assessment in many AI triage studies represents a significant ethical limitation (Saleem et al., 2026). The Nightingale Project identified specific bias risks, including concerns that certain identification technologies performed less reliably depending on patient skin color (Lochmannová, 2025).

3.3.2. Accountability and Liability

Assigning responsibility for AI-assisted triage decisions presents a complex ethical and legal challenge. Qualitative research from conflict zones found that ethical concerns about AI reliability and liability were frequently raised, with providers asking: "If an AI system makes a wrong recommendation and a patient dies, who is responsible, the surgeon or the software?" (Mathais et al., 2026). The chain of accountability should be visible to clinicians before use and to patients when AI materially contributes to diagnosis, triage, treatment selection, documentation, or follow-up (Saleem et al., 2026). Six critical gaps have been identified in the ethical governance of AI in emergency medicine: absence of patient-centred consent mechanisms; fragmented data ownership and vendor dependence; lack of equity audits; unresolved liability standards; insufficient institutional oversight; and algorithmic bias (Da'Costa et al., 2025). In the European Union, the General Data Protection Regulation (GDPR) mandates transparency in automated decision-making (Nord-Bronzyk et al., 2025).

3.3.3. Autonomy, Consent, and Patient Rights

Data privacy and security concerns were identified as significant barriers to AI implementation in emergency and disaster settings (Tahernejad et al., 2024). The implementation of intelligent triage systems faces challenges such as trust issues, training needs, equipment shortages, and data privacy concerns (Tahernejad et al., 2024). In conflict-affected and low-resource settings, data ownership, bias, liability, oversight, and public trust have been identified as critical ethical gaps (Mathais et al., 2026). The absence of patient-centred consent mechanisms in AI-assisted triage represents a fundamental tension between the exigencies of emergency care and the ethical requirement for patient autonomy (Da'Costa et al., 2025).

3.3.4. The Human–Machine Relationship

Automation bias and excessive reliance on AI systems emerged as recurring concerns. Challenges in implementing AI technologies involve excessive reliance on automation, which can erode clinical judgment and trust (Saleem et al., 2026). Without clear protocols and adequate training, these tools risk hindering rather than enhancing care (Da'Costa et al., 2025). Active involvement of emergency nurses in system codesign and the establishment of override mechanisms are essential to safeguard clinical integrity (Da'Costa et al., 2025). The concept of calibrated trust in LLM-assisted triage has been proposed, recognizing that the "humanness" of generative AI systems can foster human-like trust that may overlook reliability (Gnanvi et al., 2026). Position bias in LLMs raises critical concerns about operational reliability (Wickenberg-Bolin et al., 2025).

3.3.5. Moral Injury and Psychological Burden

The literature suggests that AI-driven decision support may influence the psychological burden borne by clinicians in resource-limited environments. Moral injury is a recognized risk for medical providers who must make triage decisions that withhold care from some patients to benefit others (Williams & Kemp, 2024). On one hand, AI-driven decision support may reduce the stress associated with difficult triage decisions by distributing responsibility and providing a defensible rationale for resource allocation (Mathais et al., 2026). On the other, the depersonalization inherent in algorithmic decision-making may compound the sense of moral transgression experienced by clinicians (Nord-Bronzyk et al., 2025).

3.4. Clinical Considerations

3.4.1. Accuracy and Diagnostic Performance

The clinical utility of AI-driven triage depends fundamentally on its accuracy relative to human performance and existing protocols. Evidence presents a mixed picture, with substantial variability across AI platforms and triage contexts. ChatGPT achieved 63.9% accuracy for START triage, with accuracy varying by prompt from 46.7% to 71.8% (Hata et al., 2024). Google Bard achieved 60% accuracy while ChatGPT achieved only 26.67% in a direct comparison (Shaltout et al., 2025). A multimodal Bayesian network improved triage accuracy from 14% to 53% in all patients (Rusiecki et al., 2025). An XGBoost model achieved 77.9% sensitivity with an overtriage rate of 76.4% (Shaltout et al., 2025). The clustered Bayesian 3D-CNN model achieved 82.0% accuracy and 85.8% AUC-ROC for COVID-19 MCI triage classification (Tahernejad et al., 2024). Machine-learning triage reduced mis-triage rates (prospective 0.9% vs. 1.2%) and achieved AUROC of 0.875 (Saleem et al., 2026).

3.4.2. Validation and Evidence Base

A fundamental clinical challenge for AI-driven triage is the nature of MCIs themselves. MCIs are rare events involving trauma types that are uncommon in routine practice (Mathais et al., 2026). This creates a data scarcity problem: machine learning models require large volumes of training data to achieve reliable performance (Rusiecki et al., 2025). The treatment provided to casualties in MCIs is often not systematically recorded in a manner aligned with data parameters (Lochmannová, 2025). The Syn-STARTS framework represents an innovative approach to this challenge, using LLMs to generate triage cases for scalable evaluation (Hagiwara et al., 2025). The DARPA Triage Challenge has empirically validated autonomous triage systems, demonstrating substantial performance improvements when Bayesian networks were integrated (Rusiecki et al., 2025). Models trained on civilian populations may poorly generalize to military cohorts (Mathais et al., 2026).

3.4.3. Clinical Judgment and Decision-Making

The clinical literature consistently emphasizes that AI should augment, not replace, clinical judgment in MCI triage. AI for battlefield triage is positioned not as a replacement for clinical judgment but to preserve situational awareness and prioritization over time (Mathais et al., 2026). The On Scene Injury Severity Prediction (OSISP) model was designed as a clinical decision support system (Saleem et al., 2026). The concept of the "Data Officer" has been proposed as a mediator between AI-generated alerts and medical decision-makers (Lochmannová, 2025). However, the model's reliance on superficial text cues over structured triage logic suggests that current LLM reasoning pathways require neuro-symbolic approaches (Gnanvi et al., 2026).

3.4.4. Integration into Clinical Workflows

Frontline feedback from emergency medicine practitioners has revealed skepticism about AI tools that "take too long" or "add no additional information" (Lochmannová, 2025). The speed and simplicity of traditional triage protocols are core strengths that AI systems must match or exceed to be operationally viable (Shaltout et al., 2025). Human–AI collaboration for prehospital trauma triage has been advanced through the OSISP model (Saleem et al., 2026). The DARPA Triage Challenge has emphasized that autonomous systems must support ongoing decision-making (DARPA Triage Challenge, 2026). Emergency nurses should be actively involved in system codesign (Da'Costa et al., 2025).

3.4.5. Special Populations and Contexts

The literature identifies significant gaps in AI triage performance for special populations. Intelligent triage systems face obstacles in triaging children and disabled people in disasters (Tahernejad et al., 2024). Pediatric triage algorithms demonstrate trade-offs between accuracy and speed (Fisher et al., 2022). In military contexts, models trained on civilian populations may poorly generalize to military cohorts (Mathais et al., 2026). Low- and middle-income countries face compounded challenges: inequities in digital infrastructure undermine prompt public health data management (Saleem et al., 2026).

3.5. Operational Considerations

3.5.1. Infrastructure and Resource Requirements

The operational realities of MCIs pose significant barriers to AI implementation. Limited infrastructure, lack of training, and financial constraints have been identified as the predominant barriers to integration in resource-limited settings (Saleem et al., 2026). AI-driven triage systems typically require reliable internet connectivity, computing power, and data transmission capabilities (Tahernejad et al., 2024). Inequities in digital infrastructure, particularly in low- and middle-income settings, weaken emergency response capacity (Saleem et al., 2026). The military medical literature has highlighted the need for offline-capable decision support tools that can function in DDIL environments (Mathais et al., 2026).

3.5.2. Training and Competency Development

Effective integration of AI into MCI triage requires substantial investment in training and competency development for frontline responders. Emergency nurses and physicians must develop sufficient AI literacy to understand both the capabilities and limitations of these systems (Da'Costa et al., 2025). Building AI literacy, integrating simulation-based training, and promoting ethical implementation are critical next steps (Saleem et al., 2026). Simulation-based training has been proposed as a critical component of AI competency development (Baetzner et al., 2025). The use of generative AI to create virtual MCI scenarios represents an innovative approach to training (Zhang et al., 2025). The RESPOND-Guatemala program exemplifies a context-adapted approach (RESPOND-Guatemala, 2025). In Italy, AI is increasingly applied to the 118 EMS system (Lochmannová, 2025).

3.5.3. Interoperability and System Integration

MCIs typically involve multiple agencies, emergency medical services, fire departments, law enforcement, hospitals, and potentially military medical assets (Ripoll-Gallardo et al., 2026). AI-driven triage tools must be designed for interoperability across these disparate systems (Saleem et al., 2026). The convergence of medical and technological developments has continued to transform disaster care; however, unless interconnectedness is conceptualized with a proper framework, there is a risk of developing silos (Lochmannová, 2025). A scoping review of AI in disaster medicine aimed to systematically map system integration of AI applications (Saleem et al., 2026).

3.5.4. Human Factors and Workflow Integration

The operational success of AI-driven triage depends not only on technical performance but on how well these tools integrate into the workflow of MCI response (Lochmannová, 2025). Feedback from European emergency medicine practitioners has revealed skepticism about AI tools that "take too long" (Lochmannová, 2025). The proposed "Data Officer" role represents one approach to managing the information flow generated by AI systems (Lochmannová, 2025). DARPA-funded research is developing AI to help first responders make rapid medical triage decisions (DARPA Triage Challenge, 2026). The NIGHTINGALE project has emphasized that new tools must be more physically integrated, preferably hands-free (Lochmannová, 2025).

3.5.5. Governance, Policy, and Regulation

Regulatory approval pathways for AI-driven triage tools remain underdeveloped (Nord-Bronzyk et al., 2025). In the European Union, GDPR mandates transparency in automated decision-making (Nord-Bronzyk et al., 2025). AI systems can now support anomaly detection, triage, segmentation, measurement, comparison, and reporting; yet governance problems have become more urgent as AI moves into worklists and triage routines (Saleem et al., 2026). A futures framework for clinical AI governance has been proposed (Saleem et al., 2026). The key risk domains identified include agentic autonomy, AI-mediated care relationships, and opacity about AI involvement (Saleem et al., 2026).

3.5.6. Deployment Logistics in MCI Settings

The literature on deployment logistics for AI-driven triage in MCI settings remains limited. The RESPOND-Guatemala program represents one of the few documented implementation efforts (RESPOND-Guatemala, 2025). The Nightingale Project conducted full-scale exercises to test AI-based triage solutions (Lochmannová, 2025). The military medical literature has described near-term deployable capabilities including wearable physiological sensors, early warning systems, and digital casualty documentation (Mathais et al., 2026). Mid-term developments may integrate multimodal data fusion and predictive decision support (Mathais et al., 2026). Longer-term conceptual frameworks, such as digital twins, envision fully predictive and individualized triage (Mathais et al., 2026).

3.6. Comparative Effectiveness: AI Versus Conventional Triage

Head-to-head comparisons between AI-driven and conventional triage remain limited but are growing. A randomized noninferiority trial found VR noninferior to high-fidelity simulation on the Ottawa GRS (Costa et al., 2026). VR showed favorable usability, low cybersickness, and comparable feedback (Costa et al., 2026). One study found higher physical demand in live simulation but no differences in mental demand (Zhang et al., 2025). Triage accuracy was comparable between VR and live simulation, though task completion was slower in VR (Baetzner et al., 2025). ChatGPT demonstrated 61.4% accuracy for START triage with a 5.0% undertriage rate and a 33.6% overtriage rate (Hata et al., 2024). Machine-learning triage reduced mis-triage rates and achieved AUROC of 0.875 (Saleem et al., 2026). AI-driven models and portable diagnostic technologies significantly improved decision speed compared to traditional triage protocols (Shaltout et al., 2025). Meta-analysis was not performed due to substantial heterogeneity (Costa et al., 2026). Moderators of effectiveness included exposure schedule and instructional guidance (Costa et al., 2026).

3.7. Risk of Bias and Quality of Evidence

Methodological quality was moderate to high across included studies (Costa et al., 2026). Randomized controlled trials demonstrated adequate randomization; observational studies were more susceptible to selection bias (Costa et al., 2026). Common limitations included small sample sizes, absence of blinding, short follow-up, and heterogeneous outcome measurement (Zhang et al., 2025). Publication bias could not be formally assessed (Sterne et al., 2011).

GRADE certainty was low to very low for most outcomes due to imprecision, risk of bias, and indirectness (Costa et al., 2026). The evidence base is small and heterogeneous: only a limited number of comparative studies, mostly with fewer than 50 participants (Costa et al., 2026). Several studies lacked fairness and bias assessment (Saleem et al., 2026). No studies assessed transfer to clinical practice or patient outcomes (Firdaus et al., 2026). The lack of validated, standardized triage assessment instruments in AI environments further limits comparability. Models trained on civilian populations may poorly generalize to military cohorts (Mathais et al., 2026).

4. Discussion

4.1. Principal Findings

This narrative review synthesized evidence on the ethical, clinical, and operational considerations of AI-driven triage in mass-casualty incidents across 19 primary studies and complementary sources. Four principal findings emerge.

First, AI-driven triage demonstrates promising but variable clinical performance. A multimodal Bayesian network improved triage accuracy from 14% to 53% across all patients, with diagnostic coverage expanding from 31% to 95% of cases requiring assessment (Rusiecki et al., 2025). An XGBoost model achieved 77.9% sensitivity with an overtriage rate of 76.4%, while a clustered Bayesian 3D-CNN model achieved 82.0% accuracy for COVID-19 MCI triage classification (Tahernejad et al., 2024). However, LLM-based triage showed substantial variability: ChatGPT achieved 63.9% accuracy for START triage, with a 32.9% overtriage rate, while Google Bard achieved 60% and ChatGPT only 26.67% in a separate comparison (Hata et al., 2024). These findings confirm that AI performance cannot be assumed to generalize across platforms or contexts without rigorous validation.

Second, ethical concerns remain largely unresolved. A systematic review of ethnic disparities in AI-driven triage found that AI/ML models demonstrated moderate to high discrimination (AUC 0.63–0.86) but showed consistently poorer performance for ethnic minority populations, with AUC differences of 0.05–0.12 (Saleem et al., 2026). Only 38% of studies reported performance differences by demographic group, and merely 25% conducted comprehensive subgroup analyses (Saleem et al., 2026). None provided evidence of improved clinical outcomes following AI implementation (Saleem et al., 2026). These findings reveal a concerning pattern of algorithmic bias that may exacerbate existing health disparities.

Third, operational feasibility remains the rate-limiting factor. Real-world clinical integration of AI-driven triage remains limited, with most studies confined to retrospective simulations rather than prospective deployment (Saleem et al., 2026). Implementation barriers—clinician trust, explainability mismatch, workflow integration, and the inability to quantify "soft skills"—currently outweigh algorithmic advances as obstacles to deployment (Lochmannová, 2025). The NIGHTINGALE project experienced significant challenges, primarily due to the scarcity of prehospital MCI patient data to train AI models and the difficult compliance with the GDPR when these data were present (Lochmannová, 2025).

Fourth, evidence for clinical transfer and patient outcomes is essentially absent. No studies assessed transfer to clinical practice or patient outcomes, a significant evidence gap (Firdaus et al., 2026). The evidence base remains small and heterogeneous, with no consistent use of validated triage assessment instruments (Costa et al., 2026).

4.2. Interpretation of Findings

4.2.1. Effectiveness for Triage Accuracy and Decision-Making

The variable accuracy of AI-driven triage reflects fundamental challenges in the training and validation of these systems for MCI contexts. The data scarcity problem is central: MCIs are rare events involving trauma types that are uncommon in routine practice, creating a paradox where the events AI is designed to address generate limited data for training (Mathais et al., 2026). The NIGHTINGALE project's experience illustrates this constraint: the shortage of prehospital MCI patient data forced engineers to rely on in-hospital datasets (Lochmannová, 2025).

Emerging approaches attempt to address this challenge. The Syn-STARTS framework uses LLMs to generate triage cases for scalable evaluation, with results showing that generated cases were qualitatively indistinguishable from manually curated cases (Hagiwara et al., 2025). The multimodal Bayesian network approach demonstrates that expert-knowledge-guided probabilistic reasoning can significantly enhance automated triage systems without requiring training data, supporting inference with incomplete information and robustness to noisy or uncertain observations (Rusiecki et al., 2025). This is particularly relevant for MCI contexts where data are inherently incomplete and observations uncertain.

However, accuracy metrics alone are insufficient for clinical decision-making. The DARPA Triage Challenge has emphasized that algorithms validated under laboratory conditions frequently fail in real-world deployment, and mitigation requires confidence-aware AI and probabilistic reasoning—systems that explicitly report confidence levels, apply Bayesian networks to enforce physiologic consistency, and use sequential probabilistic modeling to stabilize recommendations over time (Rusiecki et al., 2025). Without these safeguards, AI outputs risk being misleading rather than informative.

4.2.2. Ethical Implications for Equity and Accountability

The finding that AI/ML triage models consistently underperform for ethnic minority populations carries profound implications for MCI triage, where resource allocation decisions directly affect survival (Saleem et al., 2026). Models trained on predominantly white populations showed reduced accuracy in other ethnic groups, and emergency department data revealed that white patients received more acute triage scores despite often requiring less intensive care (Saleem et al., 2026). In an MCI context, where triage decisions determine who receives life-saving interventions, such biases could systematically disadvantage already-vulnerable populations.

The accountability gap compounds this concern. The question of who bears responsibility when an AI system makes a wrong triage recommendation remains unresolved. Unlike aviation, AI-based medical systems lack mature certification, verification, and monitoring frameworks, underscoring the ethical obligation to rigorously characterize failure modes rather than performance alone (Mathais et al., 2026). Regulatory frameworks premised on continuous human oversight are ill-suited to emergency scenarios demanding rapid autonomous action, and new frameworks must explicitly address when and how systems operating without a human in the loop can be ethically and safely deployed (Nord-Bronzyk et al., 2025).

The ethical tensions identified in this review align with broader concerns about AI in healthcare. Position bias in LLMs raises critical concerns about operational reliability in time-sensitive, high-stakes tasks with high information density (Wickenberg-Bolin et al., 2025). The chain of accountability should be visible to clinicians before use and to patients when AI materially contributes to diagnosis, triage, treatment selection, documentation, or follow-up (Saleem et al., 2026).

4.2.3. Operational Feasibility and Implementation Barriers

The operational findings reveal a significant gap between technological potential and practical deployment. AI systems designed for structured environments often struggle in chaotic, noisy, and dynamic emergency scenarios (Lochmannová, 2025). Voice recognition tools must contend with loud, disorganized environments where language is not always clear, while AI tools designed to count casualties in mass casualty situations face difficulties distinguishing between victims and bystanders, especially when individuals are moving (Lochmannová, 2025).

The operational barriers span multiple domains. Technical barriers involve the amount of information required by AI systems, which may not be possible to collect in the short time required for triage decisions in emergency contexts (Tahernejad et al., 2024). Interoperability issues, lack of explainability, data privacy concerns, and legal ambiguity were consistently reported (Saleem et al., 2026). Real-world implementation frequently lags behind technological advancements due to poor integration with electronic health records, lack of interpretability of AI outputs, and limited clinician trust (Saleem et al., 2026).

The Delphi study on ICT for MCI response identified skepticism about the cost-effectiveness of these technologies in real-world contexts (Lochmannová, 2025). In low- and middle-income countries, resource gaps impacting AI implementation are compounded by the need for high-quality, representative, and unbiased data, the imperative of integrating human-centred design principles, and the importance of cultural and contextual relevance for AI acceptance (Saleem et al., 2026).

4.3. Comparison with Previous Reviews

The findings of this review align with and extend previous syntheses. Tahernejad et al. (2024) identified 19 high-quality studies on AI in triage in emergencies and disasters, finding that AI algorithms like OpenPose and YOLO enhanced efficiency in mass casualty incidents, while e-triage systems allowed for continuous vital sign monitoring and faster triaging (Tahernejad et al., 2024). That review also identified challenges including trust issues, training needs, equipment shortages, and data privacy concerns (Tahernejad et al., 2024), which this review confirms and elaborates within the specific context of MCI triage.
Shaltout et al. (2025) found in a systematic review of systematic reviews that AI-driven models and portable diagnostic technologies significantly improved decision speed, diagnostic precision, and prioritization of life-saving interventions, reducing delays in critical care (Shaltout et al., 2025). However, that review also noted that no single algorithm proved universally superior (Shaltout et al., 2025), a finding consistent with the platform-dependent variability observed in this review.

Goniewicz (2025) examined autonomous drone systems for medical triage, finding a field in its early stages dominated by simulations, conceptual frameworks, and prototype evaluations, with limited translation into real-world deployment (Goniewicz, 2025). Persistent challenges included limited field validation, ethical concerns over decision-making autonomy, and the absence of robust regulatory standards (Goniewicz, 2025). These findings parallel the operational barriers identified in this review and highlight the need for human-centered design, context-aware algorithms, and hybrid operational models that balance machine autonomy with expert oversight (Goniewicz, 2025).

This review extends previous work by integrating ethical, clinical, and operational considerations into a unified synthesis specifically focused on MCI triage. Previous reviews have examined AI in emergency medicine broadly or triage systems generally, but the intersection of AI, MCI triage, and the three interconnected domains examined here has not been comprehensively addressed. By foregrounding the equity implications of algorithmic bias—an area where only 38% of studies reported demographic subgroup performance—this review contributes to the growing discourse on responsible AI integration in high-stakes, resource-constrained environments (Saleem et al., 2026).

4.4. Implications for Prehospital and Disaster Medicine Practice

The findings of this review carry several implications for clinical practice, system design, and policy.

For clinical practice, AI-driven triage should be positioned as a decision-support tool rather than an autonomous agent. Clinicians must retain the authority to override algorithmic recommendations, and override mechanisms should be established to safeguard clinical integrity. The finding that overrides were attributed to clinical judgment in 45% of cases suggests that clinicians appropriately exercise discretion when AI recommendations conflict with contextual knowledge (Saleem et al., 2026). Training programs should incorporate AI literacy alongside traditional triage competencies, ensuring that frontline responders understand both the capabilities and limitations of these systems (Da'Costa et al., 2025).

For system design, the review supports a human-centered approach that prioritizes workflow integration over technological sophistication. The finding that AI tools designed to count casualties in mass casualty situations face difficulties distinguishing between victims and bystanders suggests that system designers must account for the inherent messiness of MCI environments (Lochmannová, 2025). Systems should be designed for offline capability in austere environments, with confidence-aware AI that explicitly reports uncertainty rather than presenting misleading certainty (Rusiecki et al., 2025).

For policy, the evidence supports mandatory demographic subgroup performance reporting, external validation in ethnically diverse settings before clinical deployment, and post-implementation monitoring with ethnicity-disaggregated outcomes (Saleem et al., 2026). Regulatory oversight should precede widespread AI triage system adoption to prevent institutionalization of healthcare inequities (Nord-Bronzyk et al., 2025). The absence of mature certification, verification, and monitoring frameworks for AI-based medical systems represents a critical governance gap that must be addressed before widespread deployment (Mathais et al., 2026).

For disaster preparedness planning, the review suggests that AI-driven triage should be integrated into a systems-of-systems approach that connects patient-level decision tools with logistics, transport, hospital capacity, and national response frameworks. Localized technical gains are insufficient without system-level modeling to ensure population-level benefit (Lochmannová, 2025).

4.5. Limitations

4.5.1. Limitations of the Evidence Base

The evidence base for AI-driven triage in MCIs remains small, heterogeneous, and methodologically limited. Only a limited number of comparative studies were identified, most with small sample sizes and short follow-up periods (Costa et al., 2026). Methodological quality was moderate, with common limitations including absence of blinding, incomplete outcome data, and heterogeneous outcome measurement (Zhang et al., 2025).

A critical limitation is the absence of validated, standardized triage assessment instruments in AI environments, which limits comparability and conclusion strength (Zhang et al., 2025). The lack of fairness and bias assessment in many studies represents a significant ethical limitation (Saleem et al., 2026). Publication bias could not be formally assessed, though the tendency for small studies to report larger effects suggests it may be a concern (Sterne et al., 2011).

Most importantly, no studies assessed transfer to clinical practice or patient outcomes (Firdaus et al., 2026). The evidence base is dominated by simulation studies and retrospective analyses, with limited prospective validation in real-world MCI settings. Models trained on civilian populations may poorly generalize to military cohorts, and external validation remains scarce (Mathais et al., 2026). The short follow-up periods and lack of longitudinal data preclude conclusions about skill retention or sustained performance improvement (Costa et al., 2026).

4.5.2. Limitations of the Review

This narrative review has several limitations. The English-language restriction may have excluded relevant international studies. The 2015–2026 search period reflects the emergence of modern AI applications but may exclude earlier work on decision support in triage. The absence of meta-analysis limits precision and moderator exploration, though this was necessitated by substantial heterogeneity in interventions, comparators, and outcome measures (Costa et al., 2026). Reliance on published studies without systematic grey literature searching may have introduced publication bias (Sterne et al., 2011). The rapidly evolving nature of AI research means that new evidence may have emerged since the completion of the search. Finally, the narrative synthesis approach, while appropriate for the heterogeneous evidence base, does not provide the quantitative precision of meta-analytic methods (Popay et al., 2006).

4.6. Future Research Directions

The findings identify seven priority areas for future research:

1. Prospective validation studies in real or high-fidelity simulated MCI settings, using multi-center collaborations to generate robust data.
2. Standardized, validated outcome measures for AI-driven triage, prioritizing observable behavioral performance and clinical outcomes over self-report and surrogate metrics.
3. Equity-focused research reporting demographic subgroup performance, conducting external validation in diverse settings, and integrating social determinants.
4. Human-AI collaboration research, including optimal division of labor, override mechanisms, trust calibration, and intermediary roles such as the "Data Officer".
5. Cost-effectiveness analyses to inform institutional and health system investment decisions.

5. CONCLUSIONS:

5.1. Summary of Key Findings

This narrative review synthesized evidence on the ethical, clinical, and operational considerations of AI-driven triage in mass-casualty incidents. Five principal conclusions emerge.

First, AI-driven triage demonstrates promising but highly variable clinical performance. A multimodal Bayesian network improved triage accuracy from 14% to 53%, with diagnostic coverage expanding from 31% to 95% (Rusiecki et al., 2025). A clustered Bayesian 3D-CNN model achieved 82.0% accuracy and 85.8% AUC-ROC for COVID-19 MCI triage classification (Tahernejad et al., 2024). However, large language model performance was inconsistent: ChatGPT achieved 26.67% accuracy for START triage in one evaluation while Google Bard achieved 60%, confirming that AI performance cannot be assumed to generalize across platforms without rigorous validation (Shaltout et al., 2025).

Second, ethical concerns remain largely unresolved. Machine learning algorithms trained on nonrepresentative datasets risk reinforcing health disparities and could unfairly prioritize patients based on race, gender, or disability (Nord-Bronzyk et al., 2025). Six critical governance gaps persist: absence of consent mechanisms; fragmented data ownership; lack of equity audits; unresolved liability standards; insufficient institutional oversight; and algorithmic bias (Da'Costa et al., 2025). The absence of fairness and bias assessment in many studies represents a significant ethical limitation (Saleem et al., 2026).

Third, operational feasibility remains the rate-limiting factor. Implementation is mostly limited to pilot projects, constrained by deficits in training, regulation, financing, interoperability, and governance, alongside unresolved concerns about algorithmic bias, model drift, and cybersecurity (Saleem et al., 2026). The NIGHTINGALE project faced significant challenges due to scarce prehospital MCI patient data and difficult GDPR compliance (Lochmannová, 2025).

Fourth, evidence for clinical transfer and patient outcomes is essentially absent. No studies assessed transfer to clinical practice or patient outcomes in MCI settings (Firdaus et al., 2026). The evidence base is dominated by simulation studies and retrospective analyses (Costa et al., 2026). Models trained on civilian populations may poorly generalize to military cohorts (Mathais et al., 2026). Fifth, emergent technological approaches offer promising pathways. The Syn-STARTS framework demonstrated that LLM-generated triage cases were indistinguishable from manually curated cases, suggesting synthetic data can address data scarcity (Hagiwara et al., 2025). Wearable sensor integration is projected to reduce triage time by at least 30% (Alnuaimi, 2025). Autonomous drone swarms demonstrate technical feasibility for real-time triage (Goniewicz, 2025).

5.2. Implications for Practice

AI should be positioned as a decision-support tool rather than an autonomous triage agent. Clinicians must retain override authority, and human-machine teaming—where AI extends perception and cognition while clinicians retain decision-making authority—represents the appropriate near-term model (Mathais et al., 2026).

Training programs should incorporate AI literacy alongside traditional triage competencies, ensuring responders understand system capabilities and limitations, including position bias and automation bias (Saleem et al., 2026; Wickenberg-Bolin et al., 2025).

System design should prioritize workflow integration, offline capability, and interoperability. Mitigation of real-world failure requires confidence-aware AI and probabilistic reasoning (Rusiecki et al., 2025).

Emergency nurses and frontline responders should be actively involved in system codesign, with override mechanisms established to safeguard clinical integrity (Da'Costa et al., 2025). Institutional investment in technical support and faculty development is necessary for sustainability, with critical evaluation of cost-effectiveness (Zhang et al., 2025).

5.3. Implications for Research

Seven priority areas are identified. First, prospective validation studies in real or high-fidelity simulated MCI settings are urgently needed, with multi-center collaborations spanning civilian and military systems (Saleem et al., 2026; Rusiecki et al., 2025). Second, standardized, validated outcome measures must be developed, prioritizing observable behavioral performance and clinical outcomes over surrogate metrics (Costa et al., 2026). Third, equity-focused research must report demographic subgroup performance and integrate social determinants (Saleem et al., 2026; Nord-Bronzyk et al., 2025). Fourth, human-AI collaboration research is needed, including optimal division of labor, override mechanisms, and trust calibration (Lochmannová, 2025). Fifth, cost-effectiveness analyses should inform investment decisions (Luberisse, 2025). Sixth, implementation science approaches are needed to bridge the gap between efficacy and effectiveness. Finally, mid-term and longer-term technological developments—including multimodal data fusion, augmented reality guidance, and digital twins—warrant systematic evaluation (Mathais et al., 2026).

5.4. Final Remarks

AI-driven triage represents a promising but immature technology for mass-casualty incident response. The evidence demonstrates that AI can enhance triage accuracy and decision speed under certain conditions, but performance varies substantially across platforms, real-world validation remains scarce, and ethical concerns about bias, accountability, and equity have yet to be adequately addressed. The path forward requires a balanced approach that neither dismisses AI's potential nor overstates its readiness. AI should be understood as a supportive tool embedded within resilient systems, not as a substitute for clinical expertise. Realizing this vision will require rigorous prospective validation, equity-focused design, robust governance frameworks, and sustained investment in training and infrastructure. As AI technologies continue to evolve, the disaster medicine community must ensure that innovation proceeds in tandem with ethical foresight and clinical accountability.

REFERENCES:

1. Alnuaimi, M. K. (2025). Integrating wearable sensor data with an AI-based, protocol-flexible triage platform to accelerate decision-making during the golden hour of combat casualty care. *Cureus*, 17(8), e91121.
2. Baetzner, A. S., Hill, Y., Roszipal, B., Gerwann, S., Beutel, M., Birrenbach, T., Karlseder, M., Mohr, S., Salg, G. A., Schrom-Feiertag, H., Frenkel, M. O., & Wrzus, C. (2025). Mass casualty incident training in immersive virtual reality: Quasi-experimental evaluation of multimethod performance indicators. *Journal of Medical Internet Research*, 27, e63241.
3. Cheng, A., Donoghue, A., Gilfoyle, E., & Eppich, W. (2012). Simulation-based crisis resource management training for pediatric critical care medicine: A review for instructors. *Pediatric Critical Care Medicine*, 13(2), 197–203.
4. Costa, M., et al. (2026). Different delivery modalities of virtual reality training in emergency medicine: Systematic review. *JMIR XR and Spatial Computing*, 3, e84310.
5. Da'Costa, A., Teke, J., Origbo, J. E., Osonuga, A., Egbon, E., & Olawade, D. B. (2025). AI-driven triage in emergency departments: A review of benefits, challenges, and future

directions. *International Journal of Medical Informatics*, 197, 105838.

6. DARPA Triage Challenge. (2026). *Conference proceedings on autonomous systems for mass casualty and prolonged care*. DARPA.
7. Firdaus, R., Tantri, A. R., Manggala, S. K., Dwiputra, A. G., Pramodana, B., Auerkari, A. N., Hafidz, N., Omega, A., & Larasati, D. (2026). Utilization of virtual-reality in establishing a code trauma crisis resource management training module. *Open Access Emergency Medicine*, 18, 1–8.
8. Fisher, E. H., et al. (2022). Inter-rater reliability and agreement among mass-casualty incident algorithms using a pediatric trauma dataset: A pilot study. *Prehospital and Disaster Medicine*, 37(3), 306–313.
9. Flin, R., & Maran, N. (2015). Basic concepts for crew resource management and non-technical skills. *Best Practice & Research Clinical Anaesthesiology*, 29(1), 27–39.
10. Gnanvi, J. E., et al. (2026). Can we rely on generative AI for emergency patient triage classification? *International Journal of Medical Informatics*, 198, 105839.
11. Goniewicz, K. (2025). Autonomous and artificial intelligence-enabled drone systems for medical triage: A state-of-the-art review of emerging technologies and swarm applications. *Emergency Health Services Journal*, 2(2), 26–37.
12. Hagiwara, C., Nonaka, N., Hashimoto, Y., Uchimido, R., & Seita, J. (2025). Syn-STARTS: Synthesized START triage scenario generation framework for scalable LLM evaluation. *arXiv*.
13. Hata, R., et al. (2024). Accuracy of a commercial large language model (ChatGPT) to perform disaster triage of simulated patients using the Simple Triage and Rapid Treatment (START) protocol: Gage repeatability and reproducibility study. *Journal of Medical Internet Research*, 26, e54231.
14. Lochmannová, A. (2025). Exploring the role of virtual reality in preparing emergency responders for mass casualty incidents. *Israel Journal of Health Policy Research*, 14, 21.
15. Luberisse, J. (2025). Swarm intelligence: Multi-agent autonomous systems for mass casualty and active threat response. *Zenodo*.
16. Mathais, Q., Cungi, P.-J., & colleagues. (2026). Artificial intelligence for battlefield triage in large-scale combat operations. *Journal of Trauma and Acute Care Surgery*, 101(2 Suppl), S177–S187.
17. Nord-Bronzyk, A., et al. (2025). Assessing risk in implementing new artificial intelligence triage tools—How much risk is reasonable in an already risky world? *Asian Bioethics Review*, 17(1), 187–205.
18. Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71.
19. Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K., & Duffy, S. (2006). *Guidance on the conduct of narrative synthesis in systematic reviews*. ESRC Methods Programme.
20. Regev, S., et al. (2024). Mastering multicasualty trauma care with the Trauma Non-technical Skills Scale. *Journal of Trauma and Acute Care Surgery*, 97(2S Suppl 1), S60–S66.
21. RESPOND-Guatemala. (2025). *AI-powered mHealth program for emergency prehospital skills and triage notification*. Program description.
22. Ripoll-Gallardo, A., Oldrini, G., De Luca, G., Arghetti, D., Stellini, A., Zambelan, A., Fiameni, R., Colzani, G., Manzoni, P., Broccoli, F., Lolli, S., Zumbaio, F., Agarossi, A., & Fumagalli, R. (2026). Dispatch center performance and crisis resource management skills in mass casualty incidents: A prehospital simulation study preparing for 2026 Milan-Cortina Olympics. *Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine*, 34(1), Article 28.
23. Rusiecki, S., Morales, C. G., Elenberg, K., Weiss, L., & Dubrawski, A. (2025). Multimodal Bayesian network for robust assessment of casualty in autonomous triage. *Advances in Neural Information Processing Systems*.
24. Sachs, J. D., et al. (2025). Global burden of emergency and operative conditions: An analysis of Global Burden of Disease data, 2011–2019. *The Lancet Global Health*.
25. Saleem, M., Zare Bidoki, M., Alsuraihi, W., Al-Garadi, M. A., & Ahmed, A. (2026). Real-world implementation and evaluation of AI-driven clinical decision support in emergency medicine: A systematic review. *Healthcare*, 14(18), 2922.
26. Shaltout, A. E., Elbadri, M. E., Kaur, K., Alsharif, M. M., Alkhazendar, A. H., Hassouba, O. N., Ahmad, M. N., Osman, M., Zahid, A., & Banjamin, S. (2025). Accuracy and timeliness of prehospital global triage system protocols in mass disasters: A systematic review of systematic reviews. *Cureus*, 17(9), e92796.
27. Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., … Higgins, J. P. T. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*, 343, d4002.
28. Tahernejad, A., Sahebi, A., Salehi Sahl Abadi, A., & Safari, M. (2024). Application of artificial

intelligence in triage in emergencies and disasters: A systematic review. *BMC Public Health*, 24(1), 3203.

29. Way, D. P., Panchal, A. R., Price, A., Berezina-Blackburn, V., Patterson, J., McGrath, J., Danforth, D., & Kman, N. E. (2024). Learner evaluation of an immersive virtual reality mass casualty incident simulator for triage training. *BMC Digital Health*, 2(1), Article 56.
30. Wickenberg-Bolin, U., Cohen, K., Björnesjö, H., & Tegen, A. (2025). Position bias in LLMs for critical decision support: A case study on multiple casualty triage. In *Social Networks Analysis and Mining* (pp. 1–15). Springer.
31. Williams, R., & Kemp, V. (2024). How the world views trauma and trauma care. In *Major Incidents, Pandemics and Mental Health* (Chapter 2). Cambridge University Press.
32. Zhang, Z., Meybodi, M. M., Ingale, A., Karimova, L., & Vinnikov, M. (2025). Extended reality technology for emergency medical service training: Systematic review. *Frontiers in Disaster and Emergency Medicine*, 3, 1630167.