Archives of Academic Emergency Medicine. 2023; 11(1): e38

OPEN ACCESS

REVIEW ARTICLE

The Aspects of Running Artificial Intelligence in Emergency Care; a Scoping Review

Mohsen Masoumian Hosseini[1], Seyedeh Toktam Masoumian Hosseini[2]*, Karim Qayumi[3], Soleiman Ahmady[4], Hamid Reza Koohestani[5]

1. Department of E-learning in Medical Sciences, SMART University of Medical Sciences, Tehran, Iran.
2. Department of Nursing, School of Nursing and Midwifery, Torbat Heydariyeh University of Medical Sciences, Torbat Heydariyeh, Iran.
3. Centre of Excellence for Simulation Education and Innovation, Department of Surgery, University of British Columbia, Vancouver, BC, Canada.
4. Department of Medical Education, Virtual School of Medical Education & Management, Shahid Beheshti University of Medical Sciences, Tehran, Iran.
5. Department of Nursing, Social Determinants of Health Research Center, Saveh University of Medical Sciences, Saveh, Iran.

Received: March 2023; Accepted: April 2023; Published online: 11 May 2023

Abstract: **Introduction:** Artificial Inteligence (AI) application in emergency medicine is subject to ethical and legal inconsistencies. The purposes of this study were to map the extent of AI applications in emergency medicine, to identify ethical issues related to the use of AI, and to propose an ethical framework for its use. **Methods:** A comprehensive literature collection was compiled through electronic databases/internet search engines (PubMed, Web of Science Platform, MEDLINE, Scopus, Google Scholar/Academia, and ERIC) and reference lists. We considered studies published between 1 January 2014 and 6 October 2022. Articles that did not self-classify as studies of an AI intervention, those that were not relevant to Emergency Departments (EDs), and articles that did not report outcomes or evaluations were excluded. Descriptive and thematic analyses of data extracted from the included articles were conducted. **Results:** A total of 137 out of the 2175 citations in the original database were eligible for full-text evaluation. Of these articles, 47 were included in the scoping review and considered for theme extraction. This review covers seven main areas of AI techniques in emergency medicine: Machine Learning (ML) Algorithms (10.64%), prehospital emergency management (12.76%), triage, patient acuity and disposition of patients (19.15%), disease and condition prediction (23.40%), emergency department management (17.03%), the future impact of AI on Emergency Medical Services (EMS) (8.51%), and ethical issues (8.51%). **Conclusion:** There has been a rapid increase in AI research in emergency medicine in recent years. Several studies have demonstrated the potential of AI in diverse contexts, particularly when improving patient outcomes through predictive modelling. According to the synthesis of studies in our review, AI-based decision-making lacks transparency. This feature makes AI decision-making opaque.

Keywords: Algorithms; Artificial intelligence; Emergency service, hospital; Emergency medicine; Machine learning; Neural networks, computer; Ethics

Cite this article as: Masoumian Hosseini M, Masoumian Hosseini ST, Qayumi K, AhmadyAhmady S, Koohestani HR. The Aspects of Running Artificial Intelligence in Emergency Care; a Scoping Review. Arch Acad Emerg Med. 2023; 11(1): e38. https://doi.org/10.22037/aaem.v11i1.1974.

1. Introduction

A growing need for quality care has been observed in recent years. Healthcare providers worldwide are learning that Artificial Intelligence (AI) can help them address critical healthcare challenges (1).

Emergency Departments (EDs) play an important role in admitting patients who need urgent medical attention. AI can help improve the patient experience in ED by reducing waiting times, a known factor that decreases patient satisfaction (2). Nurses and physicians in emergency departments are responsible for the initial assessment, diagnosis, rapid treatment and stabilization of patients with varying degrees of severity (3). AI can help emergency medicine departments categorize patients more quickly and accurately than cur-

*Corresponding Author: Seyedeh Toktam Masoumian Hosseini; Department of Nursing, School of Nursing and Midwifery, Torbat Heydariyeh University of Medical Sciences, Torbat Heydariyeh, Iran. TEL: 09158279655, Email: masoumiant99@gamil.com, ORCID: https://orcid.org/0000-0003-3093-8799.

rent methods, which rely heavily on subjective assessments (4). Current evidence shows that AI can significantly improve emergency medicine quality and speed. However, some argue that AI will eventually take over some of emergency personnel's work (5). The great promise of AI lies in its ability to analyse and learn from large amounts of data. It also recognizes patterns that would otherwise be impossible to discern. This ability has led to questions and concerns about liability and risk. These concerns relate in particular to the autonomy of AI applications (6). Some see AI as having a complementary role to humans, e.g. in decision support and decision augmentation. This means that humans (e.g. doctors or programmers) take over supervision and collaborate with AI (7-10).

Studies have shown that the latter approach performs better than experts working alone (11). There are also positive effects on patient outcomes, reduction of errors, optimization of the healthcare system, a reduction in costs and a higher return on investment (9). AI applications are not necessarily adopted into routine healthcare practice just because they are being developed. Research has identified several factors that influence the adoption of innovations. These factors include the context (e.g. economic and political context, laws and regulations, and sociocultural factors), the organization (e.g. organizational structure, resources and processes), the group (e.g. professional values and cultures), the individual (e.g. attitude, motivation, user satisfaction and trust), and the technology (e.g. ease of use, design, accuracy and explainability) (11, 12). With this in mind, it is necessary to learn more about how AI can be integrated into emergency medicine. This is not only an innovation, but also an aspect of its unique potential and associated concerns.

Previous reviews have only focused on some aspects of the process of implementing AI in healthcare, such as data regulation and legal concerns (13, 14), trust and ethics (15-17), clinical and patient outcomes (18-20), and economic impact (21). AI is also used in healthcare for a range of purposes, such as diagnostics, predictive medicine, and clinical decision-making (22-24). A few reviews have taken an overarching perspective and explored co-production processes (25), implementation frameworks (26), and critical implementation barriers or success factors (27) that could help in developing AI implementation strategies. In general, it is felt that the implementation of AI in emergency medicine, as a core component of the healthcare system, could significantly improve patient care overall. The purposes of this study were to map the extent of AI applications in emergency medicine, to identify ethical issues related to the use of AI, and to propose an ethical framework for its use.

2. Methods

2.1. Study design

Due to the scale and nature of the data of interest to the research question, we decided to conduct a Scoping Review. This type of literature review focuses on mapping relevant literature in the area of interest. Compared to systematic reviews and meta-analyses, this type of study addresses a wide range of topics and allows for the inclusion of different study designs. In this study, the model of Arksey and O'Malley was followed, with further clarification and recommendations from Levac et al. This scoping review was conducted according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) extension for scoping reviews (PRISMA-ScR) and established frameworks for conducting scoping reviews (28).

2.2. Stage 1: Identifying the research question

Goals and research questions

To ensure that a broad range of literature on the topic of interest is covered, we posed the following initial research questions to guide the search:

1. What areas have been addressed in the current applications of AI technology in emergency medical care?
2. Which AI algorithms are most commonly used in emergency medical care?
3. What are the current conditions and rationales for AI technologies in emergency medical care?
4. What is the future of AI development in emergency medical care?
5. What ethical issues might arise in the development of AI in emergency medical care?

2.3. Stage 2: Identifying relevant studies

In the second phase of the scoping review, we sought to identify primary studies and reviews suitable for answering our research questions. To this end, we undertook an elaborate systematic process to review a wide range of literature according to defined criteria and used a philtres that helped us identify relevant studies. In accordance with Wohlin's guidelines (29), we conducted a snowball search for further relevant studies after completing the initial literature search.

A comprehensive collection of literature was compiled via electronic databases and reference lists. Due to time and language constraints, some limitations were made. Since initial searches revealed that AI applications in healthcare are outdated, only studies published between 1 January 2014 and 30 September 2022 were considered. Since study protocols, emergency medicine and AI/machine learning have evolved, there are no application-relevant studies prior to the time period considered here. Only papers whose full text was available in English were considered. A wide range of different methods was used to develop a comprehensive set of search terms and ensure the topic was fully covered. In total, 3 methods were followed to identify the relevant data sources mentioned above. Working with a subject librarian helps optimize searches and ensure search terms are fully covered.

Electronic databases searches:

The following electronic databases/Internet search engines were used for this study: PubMed, Web of Science Platform, MEDLINE, Scopus, Google Scholar/Academia, and ERIC. A manual search for the most relevant journals was not performed. This is because we believe that electronic databases in combination with snowball searches are sufficient to obtain a comprehensive overview of the topic under study.

In this study, the keywords listed in table 1 were used as search terms. We searched the retrieved articles in the mentioned databases for titles, abstracts, and index terms. The selected articles were imported into the free online version of EndNote reference manager.

Table 1 shows an example of a search strategy developed for MEDLINE. The authors used Endnote's built-in features to eliminate duplicate articles from multiple sources. At the beginning of the scoping process and during the review process, the team met to discuss decisions, challenges, or uncertainties related to study inclusion and exclusion. The search strategy was refined as needed. It should also be noted that at the time of this search it was not possible to access foreign grey literature in English (university dissertations, government documents, organizational documents, and others), so this study does not include a grey literature review.

We applied Wohlin's snowball system to identify additional publications that could be considered as new candidates for inclusion in the database. These publications were found by searching selected services and applying eligibility criteria. A single-layer forward snowball search (citation mining) and a backward snowball search (chain search) were applied to all included articles to detect related publications that may not fall under the defined search strategy. All relevant articles identified by any of the above searches were retrieved and their reference lists were checked for other potentially eligible entries. The search was originally conducted on 10 August 2022 and updated to include articles published on 6 October 2022.

2.4. Stage 3: Study selection

Records with common publication formats were reviewed in two steps before inclusion in the study. As part of the search process, results were downloaded from the respective online databases, deduplicated and uploaded to an online review manager (EndNote) for further review. In the first step, two authors (MM and TM) independently reviewed the titles, abstracts, summaries, and synopses of the records found (each author was blind to the selection of the other author). In the second step, two independent reviewers assessed the full texts found for their suitability. One of them assessed half of the full texts and the other assessed the other half. As part of our process, we double-checked the suitability of 10% of all full texts, and in case of a poor match, (< 80%) additional double-checks were performed. The two authors of the review discussed any discrepancies identified during the screening process. If agreement could not be reached on the inclusion of studies, a third independent reviewer (SA) was consulted and, if necessary, the opinion of the whole team was sought.

Eligibility criteria

There were four criteria to determine the eligibility of a study: Population, Intervention, Comparison Group, and Outcome (30).

Non-English articles were excluded from the review because translation into English was not possible due to resource constraints.

Considering the type of publication, articles from peer-reviewed journals were included in the selection process. All conference abstracts, book reviews, commentaries and editorial articles were excluded from the review. No articles were included if the title or abstract of the study itself was not classified as a study of AI or Machine Learning (ML) interventions. In addition, studies outside the ED or hospital setting were disqualified. Studies with data sets from emergency departments that were not directly related to emergency medicine were not included. Finally, papers without outcomes or assessments were excluded from the review.

PICO

1-Population

Emergency Medical Services (EMS), Emergency Departments (EDs), urgent care centers, and trauma units worldwide providing pre-hospital or out-hospital care

2-Intervention

Any informatics intervention classified as AI by the study authors, including supervised and unsupervised ML, deep learning, and neural networks

3- Comparison group

No intervention, standard treatment, another informatics intervention, or another comparison group

4- Outcome

Any outcome described in the literature.

2.5. Stage 4: Charting the data

In this phase, the most relevant information was extracted and the data were categorized and sorted as needed. Two reviewers examined the retrieved articles and extracted the information using a predesigned form that two reviewers had jointly developed to determine the variables to be extracted. The following data were extracted: Type of study, date of publication, type of technology and application in emergency

medicine.
In accordance with the data extraction framework, the responsible team members collected data from each included study, independently. Consistency of approach to data extraction was ensured through ongoing dialogue between the two reviewers. Through content analysis, a concept and its application were identified and used to create a classification scheme. For this purpose, data items that represented discussion categories were merged or renamed. Two team members tested it on a sample of the included studies to ensure that the coding scheme was appropriate and could be applied consistently. Disagreements that arose between reviewers were resolved through discussion. A third reviewer (S.A) acted as a referee when disagreements could not be resolved. We raised missing or incomplete data with the study authors.

2.6. The following data has been extracted:

1. General information: Authors, publication year, country, clinical setting, study objective, and study design/method
2. Types and applications of AI: AI technology used, type of AI model, type of task performed by the AI, degree of autonomy of action, intended use of the AI, and intended user of the AI

2.7. Stage 5: Collating, summarizing, and reporting the results

At this stage, we summarized the results and provided an overview of the literature reviewed. The results have been summarized in tables of descriptive statistics and narrative form. The verbatim extracted data were analysed by 2 reviewers according to the principles of thematic analysis (31). The identified themes were presented in schemas. After reading and re-reading the articles, the next step was to identify the initial codes in each article. These codes were examined for similarities and differences and grouped into potential themes, which were then analysed to develop thematic maps that were then used to create clear definitions and names for each theme within each domain. Coding and data analysis were conducted in pairs. Two reviewers coded the data and met regularly to discuss the coding results. The lead authors were involved when consensus was needed.

2.8. Ethics and dissemination

Ethical approval was not required as the study includes information from previously published work. The results are summarized in a report submitted to a peer-reviewed scientific journal. After publication, the results will be shared with relevant networks and local and national organizations working in the field of digital health. The results will be disseminated in appropriate formats such as journal articles, presentations, conferences, and press releases.

2.9. Assessment of methodological quality

In accordance with the guidelines for conducting scoping reviews (32), we did not assess the methodological quality or risk of bias of the included articles.

3. Results

In the first phase, 2175 articles were found in the relevant databases. Repetitive articles were deleted and removed. At the end of this phase, 1248 articles remained. In the next phase, the articles' titles were carefully examined. Articles that were not relevant to the main question and could not answer the question were removed from the study. Finally, 454 articles were included in the study. The abstracts of the remaining articles from the last phase were carefully reviewed twice. Articles that could not answer the main question of the study were removed in this phase. At the end of this phase, 137 articles remained, whose full texts were retrieved and reviewed and finally, 47 articles that could answer the main questions of the study were selected for inclusion in the study (Figure1).
The majority of the included studies focused on ML (n = 43), and most of these papers aimed at predicting healthcare outcomes (n = 23). The remaining studies focused on AI (n = 24), neural networks (n = 11), and deep learning (n = 7). These studies examined disease management and diagnosis, as well as service cases such as triage (Figure 2). ML studies were primarily aimed at developing models that can accurately predict healthcare outcomes such as disease progression or response to treatment. The other studies focused on AI, neural networks, and deep learning and aimed to develop systems that could help manage and diagnose diseases and provide services such as triage.
This review covers seven main domains of AI techniques in emergency medicine: ML algorithms; prehospital emergency management; triage, acuity, and disposition of patients; disease and condition prediction; emergency department management; the future impact of AI on EMS; and ethical issues.

3.1. Types of ML algorithms

Deep Learning algorithms automatically learn feature hierarchies, require a large amount of data to make predictions, and take significantly more time to train than traditional ML algorithms. For an overview of artificial intelligence methods in ED, see Figure 3 and Table 2.
In supervised learning, a decision tree is a method of classification and regression structured like a flowchart. It recursively partitions the data into subsets based on the feature value that minimizes the impurity of the resulting subsets. Ensemble learning is an ML method that combines multiple algorithms to increase accuracy and decrease vari-

ance. Among the breakthroughs in deep Learning, the Convolutional Neural Network (CNN) has attracted researchers' attention due to its high performance on computer vision tasks. CNN has proven successful in overcoming the limitations of traditional ML models. Deep Learning has significantly contributed to natural language processing, including using neural network architectures to extract valuable information from unstructured text data in electronic health records. Supervised ML algorithms have replaced traditional methods in predictive studies and are expected to outperform simple scoring systems. They are also expected to reduce bias in causal analysis studies (33, 34).
In Razieh Safaripour's study, 9348 participants visited ED at least once between 2008 and 2015, with 15,627 visits. The data was stored in 16 separate datasets for the population studied. They explored five predictive modelling techniques to account for typical irregularities in health data, including logistic regression with the best subsets based on Recursive Feature Elimination (RFE) and classification algorithms. They then used logistic regression with RFE, random forest and Support Vector Machine (SVM) to identify the best-fitting regression and create a more accurate classifier. They also applied Scikit-Learn's built-in feature GridSrearchCV with 10-fold cross-validation. Each model's performance was calculated using standard classification evaluation metrics: area under the curve (AUC), sensitivity, precision, calibration error, and classification error. The values are expressed as the percentage of incorrect predictions. As a result, this study evaluated four ML algorithms for predicting frequent visits to ED in the Korean population and found that Random Forest was the best method with the highest precision and lowest classification error. Other ML algorithms also outperformed logistic regression in predicting frequent visits to ED (35).
The study by Georgios Feretzakis examined the performance of eight ML models based on data from the biochemistry and haematology departments of ED patients. The models were trained using urea, creatinine, lactate dehydrogenase, creatine kinase, C-reactive protein, and complete blood count data. All raw data were retrieved from a standard Hospital Information System (HIS) and a Laboratory Information System (LIS) and analysed using the Waikato Environment for Knowledge Analysis (WEKA). To evaluate the best model's performance, we performed 10-fold cross-validation. WEKA provides detailed results for the classifiers studied. AUC was calculated as the area under the Receiver Operating Characteristic (ROC) curve. They also kept the default settings of the original implementations of all WEKA classifier algorithms and evaluated each algorithm on two datasets. They also applied the Class-Balancer technique from WEKA to prevent overfitting. They then evaluated eight ML algorithms for predicting hospitalisation in adult patients in the ED. The models produced a yes/no result and were easy to use (36).
In another study by Georgios Feretzakis, five ML algorithms were assessed using routine ED and laboratory tests. These algorithms were evaluated to promptly predict patients' admission to and discharge from the hospital. The five algorithms were linear discriminant analysis, recursive partitioning and regression trees, support vector machines, k - nearest neighbor, and random forests. Among the different classifiers evaluated by 10-fold cross-validation, one random forest model performed better than the other models in terms of AUC-ROC (37).
According to our analysis of the themes that emerged from the studies involved in this review, AI applications in ED can be divided into five categories; we have summarized these categories in Figure 4. A summary of the articles used for AI applications in ED is presented in Table 3. This is a summary of the study's population, method and algorithm, aim, and outcomes.

3.2. Prehospital emergency management

In prehospital emergency medicine, ML and deep learning techniques detect and predict critical medical conditions before the patient arrives at the hospital (38). In the study by James Morrill, ML methods were used to develop and validate a triage method for heart failure exacerbations. The algorithm developed in this study performed exceptionally well compared to individual clinicians in assessing the likelihood of a patient experiencing an exacerbation and determining the appropriate consensus triage category. This study demonstrated that an ML approach to categorising patients with heart failure is a viable and accurate method to facilitate home triage and self-identification of exacerbations (39).
In Da-Young Kang's study, the AI algorithm was combined with conventional triage tools to develop a robust algorithm that predicted critical patients at level 3 and non-critical patients at levels 4 and 5. The AI algorithm outperformed Emergency Severity Index (ESI), Korean Triage and Acuity Scale (KTAS), National Early Warning Score (NEWS) and Modified Early Warning Score (MEWS). In addition, the AI + ESI ensemble algorithm outperformed the AI algorithm and other conventional methods. This study demonstrated that an AI algorithm could accurately predict critical care needs in the prehospital setting. The AI algorithm outperformed conventional triage tools and scoring systems and was more accurate than medical staff decision-making. Conventional triage tools have low accuracy and are in the middle range. The AI+ESI algorithm has higher accuracy than the AI algorithm alone, and the combination of expert opinion (ESI level) and the AI algorithm has more accurate performance (9). Blomberg et al. used a trained ML system to detect and identify Out-of-Hospital Cardiac Arrest (OHCA) cases using audio recordings from emergency call centres. The trained

system was significantly faster than a trained medical dispatcher. Of 918 calls with OHCA, the ML system accurately identified 84.1% of cases compared to 72.4% by a trained medical dispatcher. The median time taken by the ML system to detect an OHCA for all cases was also significantly shorter at 44 seconds compared to 54 seconds (40).

Al-Dury et al. used a Random Forest (RF) approach to assess the importance of 16 critical factors associated with 30-day survival in OHCA. According to the results, initial rhythm, age, time to cardiopulmonary resuscitation, time to ambulance arrival, and location of cardiac arrest are all significant predictors of survival (41).

An evaluation of six ML classifiers on real ambulatory and demographic datasets for predicting ambulance demand was conducted by Lin et al., including the Regional Moving Average (RMA), Linear Regression (LINR), Support Vector Regression (SVR), Multi-layer Perceptron (MLP), Radial Basis Function Network (RBF) and Light Gradient Boosting Machine (LightGBM). They found that LightGBM was the best-performing model and that the total number of demands was the critical feature for predicting next-day demand accurately (42).

3.3. Triage, acuity, and disposition of patients

In this field, ML and deep learning techniques help healthcare professionals determine the patient's triage level and care needs based on the severity of their condition. Patient disposition is sending the patient to their next destination based on their clinical status (43).

Georgios Feretzakis examined coagulation tests and biochemical markers routinely used in ED patients concerning ED outcomes. They used Multivariate Imputation by Chained Equations (MICE) to fill in missing data on age, gender, and admission in a large dataset of 13991 patients who visited the ED and were admitted or discharged. They applied and evaluated several classification models to identify the best model. A multiple decision tree algorithm showed the best performance in the pulmonology department ED with a sensitivity and specificity of 0.7168, 0.5834, 0.6778, 0.7013, 0.6969 and 0.7184, 0.7617, 0.6800, 0.7687, 0.7757, respectively (44).

A cross-sectional observational study of pediatric patients attending ED in South Korea was conducted using nationwide registry data. The KTAS was used, which showed adequate reliability and validity. The data came from National Emergency Department Information System (NEDIS), a national database that collects information from more than 400 emergency departments across South Korea. The NEDIS contains age, gender, type of insurance, mode of transport, state of consciousness on admission, and time variables. This study used RF to identify critically ill children and predict hospital stays. RF can solve the "black box" problem of ML models by calculating the importance of variables by reducing the Gini index. This study developed and compared several ML models to predict critical cases and hospitalisations in pediatric ED visitors using nationwide data. The RF model performed well in discriminating critical cases with an AUROC of 0.991 and an area under Precision-Recall curve (AUPRC) of 0.640 (45).

In a study, AI techniques were used to determine patient acuity and triage. Based on the RF model, over 14,326 ESI level 3 patients who required a higher level of care due to their condition were identified. The tool was equivalent to or better than the US ESI (46).

To predict the KTAS, Choi et al. developed three ML models. They found that the RF and XGBoost models trained on clinical data only outperformed the linear regression (LR) model trained on both clinical and free text data (47).

Using a large dataset of 11,656,559 patients from 151 emergency departments in the Korean NEDIS, Kwon et al. developed a deep learning triage and acuity score using a 5-layer multilayer perceptron (MLP)model. Unlike traditional triage tools with complex scoring methods and data requirements, this model requires only basic input parameters such as age, gender, chief complaint, time from symptom onset to ED visit, mode of arrival, trauma, initial vital signs, and mental status. A significant strength of the model is that it can be applied in different settings, e.g. prehospital emergency medicine (48). Similarly, Yu et al. have developed a first-triage system based on nurses' assessments that outperforms existing systems. The study results indicate that LR can achieve comparable results to neural networks when the dataset's dimensionality is kept to a minimum (49).

Using ML underpinnings, Farahmand et al. developed a web-based interface to correlate acute abdominal pain with the severity of the emergency. Initially, six models were individually evaluated before being ensembled; the output of one model was used as an input for another, based on the first output. The result was an improvement in AUC scores and greater accuracy (50).

The researchers used ML techniques to predict hospital admissions and admissions based on patient records and parameters collected at triage and compared nine ML and deep learning techniques. AI proved reliable in predicting hospital admissions in these studies (51, 52).

Compared to traditional triage approaches, Raita et al. and Goto et al. used ML methods (lasso regression, RF, gradient-based decision trees and deep neural networks) to predict clinical outcomes in adults and children. The results suggest that deep learning and ML models have higher discriminatory power, which may prevent over- or underestimation. Roquette et al. found that text data improved AUC by 1.9% when used to predict paediatric ED admissions via deep neural networks. This article highlights the importance of in-

cluding text data in models (53, 54).
Zhang et al. applied LR and Multi-Level Neural Networks (MLNN) to various structured and unstructured free-text data sources to predict hospital admissions or emergency room referrals. In this study, it was shown that the use of free-text data extracted using natural language processing led to better prediction results (55). Chen et al. used deep neural networks in the form of hybrid bidirectional short-term memory and CNN to predict patient intake from clinical narratives and structured data (56).
In another study, Sterling et al. developed a three-layer neural network regression model on free-text triage data as a result of natural language processing to determine patient triage based on free-text triage data (57). Fernandes et al. developed an ML and natural language processing approach to identify patients at high risk for mortality and cardiac arrest 24 hours after triage (58).

3.4. Disease and condition prediction

In developing countries with a shortage of trained cardiologists and ophthalmologists, ML and deep learning techniques can detect medical conditions and diseases (59).
In a study, Viu et al. compared ML and deep learning techniques to detect whether or not patients suffer from pain. Their study showed that AI was applicable with a macro-average F1 score of 90.96% (standard scoring metric for classification tasks) (60). In the Lopez Pineda study, eight ML models were evaluated on 31,268 influenza cases in emergency departments to determine the most accurate influenza case detection model (61). ML models were used to predict mental health problems such as post-traumatic stress disorder and death by suicide after an ED visit for parasuicide (62). Lindsey et al. have developed a model that detects and localises fractures in X-ray images (63), and Olczak et al. have shown that it can classify X-ray images of the wrist, hand and ankle into four classes (64).
Researchers have developed artificial neural networks to predict cardiac arrests, detect heart attacks, and identify five classes of heartbeats. These neural networks outperform conventional ML systems and track triggers in accurately detecting cardiac arrests in hospitals (65-68). The researchers have developed supervised artificial neural network models that predict long-term functional outcomes and the risk of poor outcomes early in patients with OHCA. These models performed better than an earlier LR-a-based study conducted with the same cohort (69). ML techniques were used to predict sepsis in patients within 24 hours, 72 hours, and 28 days (70-72).
Koh et al. developed a novel hybrid classification system to detect retinal detachments in ultrasound images and achieved a high classification accuracy of 99.13% (72). Patel et al. investigated ML methods to predict the need for hospitalisation or intensive care in asthma and chronic obstructive pulmonary disease post-triage exacerbation (73).

3.5. AI in emergency department management

Emergency department management encompasses all matters required for smooth operation, including operational and logistical matters, contingencies, etc. ML and deep learning techniques can better assess patient volumes in the emergency department and improve operational efficiency so patients receive the most appropriate care possible (74, 75).
Qing Liu's study introduced artificial intelligence to improve first aid efficiency by integrating medical information processing and care management into first aid. This study compared the average triage time for critically ill patients. The results were used to assess whether emergency care was managed effectively. In terms of first aid effectiveness and triage speed and accuracy, a clinical decision-making system improved the success rate of rescuing patients with haemorrhagic shock, coma, dyspnoea and organ injury. The results proved that the system could effectively improve first aid efficiency and accuracy (76).
Lin Wu's study used three deep-learning models to detect STEMI based on a 12-lead electrocardiogram (ECG). The models performed well, with an AUC of 0.99, and the deep-learning models could detect arrhythmias and heart attacks. Their model utilized long-term, short-term memory (LSTM)+CNN to extract ECG features from 12-lead ECG data and was verified on a real ECG dataset of heart attacks. This study used Deep Learning to detect STEMI and culprit vessels based on real ECG data (77).
The study by Yamanaka S analysed data from 13 academic and community emergency departments in Japan treating respiratory emergencies. Patients were treated at the discretion of the attending physicians, and all patients who underwent surgical intubation on the first attempt were excluded. They developed ML models for difficult airways and first-attempt intubation success; they used patient demographics, pre-intubation vital signs, modified LEMON score, and intubator's specialisation. They then measured the performance of the reference and ML models. They examined association measures: C-statistics, sensitivity, specificity, positive and negative predictive values, and positive and negative likelihood ratios. The significance of the variables for the best-performing ML model was also calculated. In this analysis of multicentre, prospective data from 10,741 ED patients, the state-of-the-art ML models demonstrated superior discriminatory performance in predicting both difficult airways and first-pass success. In addition, these ML models achieved higher specificity in predicting these two outcomes. The models performed better than conventional approaches (78).

Whitt et al. used a seasonal autoregressive integrated moving average time series model and a neural MLP network model to predict daily arrivals and hourly occupancy in real-time (79). Khaldi et al. used an MLP Feed-forward Artificial Neural Network paired with an Ensemble Empirical Mode Decomposition (EEMD) signal decomposition technique to predict weekly patient visits to ED (80). Yousefi et al. developed a prediction tool with LSTM support that outperformed other known techniques in predicting patient visits up to 7 days in advance (81).
Using a search engine and social media data to predict patient volume and air quality indices to predict asthma-related visits to the ED may offer novel, unconventional approaches to capturing and linking non-diagnostic data trends for asthma prediction (82, 83). ML algorithms with elements of systems thinking were applied to predict patient wait times in ED. The results showed that using ML reduced waiting time errors by 15 to 20% (84).
Greenbaum et al. developed a domain-specific ontology and ML-based user interface to optimise ED quality and workflow. The interface used top-5 suggestions and contextual autocompletion to help nurses document complaints (85). In the study by Tootooni et al., a natural language processing tool (CCMapper) was developed, which utilizes heuristic algorithms and neural networks to assign free-text symptoms and problems to chief complaints (86). Frost et al. used free-text mining from emergency records to predict treatment costs for patients who are assumed to visit the emergency department frequently, which will cost the healthcare system dearly (87).

3.6. Future impact of AI on EMS

AI development is currently experiencing an upsurge that will affect many aspects of emergency medicine in the near future.

3.7. Clinical image analysis

Advances have improved several areas of medical image analysis and image recognition. These include detecting pneumonia on chest X-rays, segmenting subdural haematomas from CT scans of the brain in three dimensions, and evaluating CT scans of patients with suspected acute ischemic stroke with the same accuracy as stroke specialists (88).

3.8. Clinical surveillance

Intelligent clinical monitoring can enable early detection of deteriorating patients (89). For example, a Compensatory Reserve Index (CRI) can be included in the monitoring screen to predict impending cardiovascular instability (90), and a sepsis prediction algorithm can be used hours before the clinical onset of sepsis (91, 92). Any system that attempts to raise the alarm faces the problem of false alarms (93). ML has been used to suppress false alarms in the ICU and predict blood pressure based on ECG recordings (94).

3.9. Predicting clinical outcomes

AI algorithms are getting better at predicting the future, often outperforming current clinical scoring systems. For example, ML models have been used to predict cardiac arrest within 72 hours (95), defibrillation success in out-of-hospital cardiac arrest, and 30-day mortality after a myocardial infarction with elevated ST segment (96).

3.10. Population and social media analysis

AI can be helpful in public health and disease surveillance. ML models have been developed to detect flu and suicide alerts on Twitter (97). ML is a subfield of AI that uses various methods to automatically detect patterns in data and then use these patterns to make predictions or decisions. Deep learning has recently been responsible for numerous breakthroughs in areas as diverse as image and speech recognition, speech synthesis, natural language processing, and translation (98, 99).

3.11. Home monitoring

In the future, patients may present to the ED earlier and with more information, as ML can be used to predict acute exacerbations of chronic obstructive pulmonary disease (100) and worsening asthma control (101).

3.12. Future trends

AI researchers benefit from the availability of large and open data sets. These data sets provide benchmarks for comparing different models and offer international competition. Current state-of-the-art models require fine-tuning of multiple hyperparameters and are hand designed by experts. Advances in ML automation could make it easier for non-experts to access the powerful techniques of ML (102). AI systems are more likely to be used for clinical decision support than to replace physicians. However, image analysis and warning of clinical deterioration are promising areas of application.

3.13. Ethical issues

Despite promising applications, AI is not yet widely used in hospital departments or health systems—the reasons vary. Based on the studies conducted on ethical activities in the field of AI in healthcare, the following dimensions of ethical consequences have been identified (Table 4):
• Human-AI relationships in health and care (103)
• AI's impact on health data usage, storage, and sharing (104, 105)
• Algorithmic transparency and explainability in health: eth-

ical concerns (106-108)
• AI's ability to mitigate or exacerbate existing health inequities (109)
• Trusting algorithms used in healthcare (110, 111)

3.14. Human-AI relationships in health and care

Considering Professor Margaret Boden's observation about the impact of algorithms on healthcare relationships, algorithms are often trained on "measurable" data such as images, medical records, and blood test results - there is a strong bias toward quantitative data. In reality, many healthcare interactions depend on more than just these "measurable" variables, such as nonverbal communication between patients and their social environment. Is it possible to measure data value in healthcare? (112)
As technology advances and tasks become more automatable, healthcare systems need fewer "human practitioners" (113).
Healthcare workers could be affected by this change. People could be divided into subgroups and categories based on artificial intelligence, separating them from the larger community to which they belong. Healthcare is not immune to the cultural trend toward individualism, as many in society at the large claim (114). In bioethical discussions, solidarity is invoked to balance the needs of the patient with the needs of the community. Barbara Prainsack and Alena Buyx suggest that solidarity involves a collective decision to bear the "costs" (financial, social, and emotional) of helping others (115). As AI-driven personalization becomes an essential feature of health care, how solidarity will be redefined and valued remains to be seen.

3.15. AI's impact on health data usage, storage, and sharing

Despite incredible advances in computer tools for tracking and monitoring data, human decisions still need to be made about what counts as data (in the first place) and how that data should be organized, labelled, and displayed. Because of data collection, dissemination, and linkage problems, which are different for each type of data, these decisions significantly impact reality. They can even result in highly skewed data sets that do not accurately reflect (and, in some cases, actively distort) reality (116). As a result, decisions about data governance transcend far beyond purely technical decisions in that they affect how data is used. The international community emphasizes that individuals have control over algorithms' data (117). However, does that control exist for individuals' medical data?
Medical data is not under individuals' control, as Professor Eduardo Magrani of the Department of Law, Technology, and Intellectual Property has stressed. According to Prof. Magrani, certain aspects of a patient's medical data are beyond his or her control, such as genetic sequence or medical history (118). Consent is an essential component of most healthcare practices. However, obtaining informed consent for all individual data is impossible. For retrospective, historical data, the assumption is that it can be used anonymously for research without the individual's explicit consent (119, 120), but what about data collected prospectively for AI development?
Data bias means using data sets that do not fully represent the population they are intended to model. A maxim says "garbage in, garbage out," which means algorithms will produce biased results if trained on biased data sets (121).

3.16. Algorithmic transparency and explainability in health: ethical concerns

AI ethics encompasses a set of moral principles and techniques to guide AI's development and responsible use (122). AI companies should be able to explain how their algorithms work and why they do what they do (106). For AI to be safe, it must provide a high level of traceability so that it can be traced back to the source if harm occurs (108). The definition of fairness refers to datasets that contain personal data, and there must be no racial, gender, or ethnic bias (111). Responsible use of AI is essential to ensure a positive impact on interactions between consumers and brands. The key to a successful AI model is responsible data use, i.e., respecting privacy rights. More data is often better, and better models are created when more data is available (123, 124).
However, people's rights to privacy and transparency must not be compromised by collecting more and more data. Data granularity should be as narrow as possible, and data should be collected only when needed.

3.17. AI's ability to mitigate or exacerbate existing health inequities

Health inequities are likely to become even more complex with the increasing use of AI and other technologies in healthcare (125). In a future healthcare system, the patient could be empowered by data analytics and algorithmic insights rather than just being a passive recipient of health advice (126). Wearables and apps, however, require a high level of digital literacy, and their use could prove prohibitively expensive. In low- and middle-income developing countries and emerging economies, this could limit access for poor people and entire health systems.

3.18. Trusting algorithms used in healthcare

Many discussions of medical decision-making involve situations in which an algorithm or a human makes a decision that is objectively "wrong" and results in harm (127). Algorithmic "errors" are an essential topic in AI discussions in healthcare. Obviously, we apply higher standards to algorithms than to humans, perhaps because we cannot understand and empathize with algorithms as well as humans. Dr. Debra Mathews and Dr. Travis Rieder point out that trust plays a vital role in this phenomenon - algorithmic errors can be more troubling than human errors because, in healthcare, we still have more trust in human interactions than in computer interactions (128, 129).

3.19. Legislative and regulatory obstacles

A legal framework that enables integration of AI systems into the clinical setting is crucial. This framework should start with AI systems in the narrow sense, as opposed to general AI systems, to ensure a robust end product. Given that large amounts of personal data are collected, transmitted, transferred, and stored as part of the technological infrastructure of AI systems, a discussion on data protection is essential. AI developers' motives to collect as much information as possible must be balanced with patients' privacy. Data collection in accordance with the Personal Data Protection and Electronic Documents Act and the General Data Protection Regulation is recommended but may undermine AI technologies. Paternalistic, top-down regulation will nip innovation in the bud, while lax policies that encourage breakthrough technologies will increase market penetration but jeopardise patient safety. A minimal regulatory sandbox may be optimal, but the review process must be robust enough to protect doctors and patients. As a result of the legal and ethical cases in this area, this review provides an ethical framework for developing an AI application for EMS, as shown in Figure 5.

Figure 5 shows a graphical representation of the ethical framework. This framework consists of two main elements: the designer and the user. These two elements are linked by a closed exposure/experience loop that runs through the different phases of the project. The ethical dimensions of each element need to be addressed in a specific way. From the users' perspective, we have divided ethical issues into three main areas that are inextricably linked to the overall experience through a relationship of trust that needs to be established and maintained. On the designer's side, the ethical dimensions discussed in this review have been grouped by project phase. To evaluate each ethical dimension of a decision, the Asilomar AI Principles can be used as a guide. Users' needs, preferences and experiences must be considered. Users' experience must be paramount in decision-making. All products and services must be tailored to users' needs. The proposed presentation is flexible and should be tailored to the needs of the project in accordance with the list of ethical dimensions and other aspects.

4. Discussion

We conducted this study to map the extent to which artificial intelligence has been applied to emergency medicine. We also identified ethical issues related to AI use in emergency medicine and proposed a framework for it in general. This study provided a brief overview of AI concepts, definitions, and applications in emergency medicine. It also discussed the main concerns and risks of these technologies. Then, the main ethical issues were discussed and general guidelines for ethical reflection during development were presented. These guidelines included ethical principles, qualification strategies and frameworks. This article described the themes and applications of artificial intelligence in emergency medicine, and proposed a theoretical framework for ethical considerations in implementing AI in emergency medicine as well.

Currently, there is very scarce research describing the views of patients and other stakeholders on AI use in emergency medicine.

There has also been no research into AI ethics checklist factors. There is an acute lack of scientific research addressing ethical challenges and possible solutions. AI in healthcare is leading to growing privacy concerns as confidential health data is shared with a variety of untrusted companies that have access to this data. There is no doubt that this is inevitable as ML and deep learning algorithms need large amounts of data to qualify, train, test, and validate them (20, 21, 36). To ensure that these cutting-edge technologies are integrated into medical care in a way that builds patient trust and addresses widespread ethical concerns about AI, patient involvement is essential. Since patients are the beneficiaries of all AI applications and innovations, characterising their needs, moral values, and preferences is crucial to ensure that these technologies are developed, applied, and implemented in an ethically appropriate manner and benefit patients in the long term (37, 51, 52, 54).

According to the synthesis of studies in our review, AI-based decision-making lacks transparency. As a result, AI decision-making processes are opaque due to this characteristic. The use of AI should improve patient safety and well-being (14, 20). However, AI reinforces existing biases in healthcare (22, 23). It is impossible to make accurate predictions with AI if data sets are unreliable and under-representative. This can also lead to inequity, data bias, discrimination, and misleading predictions. Health information is some of the most intimate and sensitive information. An important ethical principle is that patients' privacy must be respected, as it reflects their autonomy and is linked to their identity and well-being (22, 24).

In order to test the applicability of artificial intelligence in emergency medicine while respecting ethical principles, practical tools urgently need to be developed. Furthermore, it is crucial for future research to identify stakeholders, users, and beneficiaries of AI-related technology in medicine and to engage them in a dialogue about AI ethics and feasible solutions for the introduction of ethical AI applications. There is a need to conduct comprehensive qualitative and quantitative studies to understand the perspectives of emergency physicians and artificial intelligence users (patients). Emergency medicine specialists must have prior AI training to use AI in daily clinical practice without violating ethical principles. A thorough study of ethical and legal issues related to AI applications in different specialties and sub-specialties of healthcare is necessary.

Our study found that AI measures in ED are heterogeneous in purpose and design. For example, supervised ML was used to develop prognostic models for pediatric asthma exacerbations, predicting repeat visits, and diagnosing strokes (73, 130, 131). NLP models have been used to optimize resource allocation in resource-poor settings, classify computed tomography (CT), and predict hospital admission through Electronic Medical Records (EMR) (132).

Around 24% of interventions (23.40%) focused on prediction in ED, supporting AI's perceived superiority in this area. Several studies have shown that AI outperforms existing decision support and scoring systems based on traditional statistical models. Applications include predicting pneumonia mortality and calculating syncope risk based on clinical data (133, 134). The ability to simultaneously process multiple variables in large data sets may explain why AI is superior to humans in predictive modelling (4). A number of the studies we included showed superior performance in matching multiple data points to predict complex outcomes compared with human comparators.

Studies have shown that human decision-making can be influenced by possible biases and heuristics. The application of artificial intelligence in medicine could help mitigate spurious correlations and metacognition errors. For example, it has been hypothesized that AI's superiority in predictive modelling could be particularly useful in diagnosing sepsis. This is a syndrome that can lead to widespread organ dysfunction, high mortality, and high morbidity. As we reviewed, some studies have used ML to predict acute exacerbations of chronic obstructive pulmonary disease, worsening asthma control, sepsis, and mortality (71, 100, 101).

The results show that 12.76% of studies used artificial intelligence in the prehospital setting. In the prehospital setting, studies were conducted on demand forecasting for ambulance allocation, cardiac arrest detection by EMS, and out-of-hospital ECG classification (40, 135). Data from EMS have been used in several ML algorithms to predict prehospital cardiac arrest (136). Supervised ML was also utilized to automatically link electronic patient reports from EMS to ED datasets (137). When treating a patient in the prehospital setting, it is important to understand that only a limited number of clinical variables can be utilized to make quick decisions (e.g. whether or not to transport the patient to the hospital). Since AI predicts outcomes based on multiple data points, it is well-suited to this task. This means AI is likely to have a significant impact on prehospital care.

Although AI seems to perform better than clinicians at building predictive models, there is no evidence to support this conclusion, yet. Most of these studies did not include a human comparison group and did not include safety-related outcomes. It should be noted that most of the included studies were retrospective analyses of data sets and therefore, require further validation in controlled clinical trials. Most studies did not address technology implementation practical implications, such as training, cost, and convenience. The proprietary nature of the algorithms makes reproducibility and reliability evaluation of results difficult, because many interventions are not publicly available or transparently described.

A scoping review was the most appropriate method for searching and evaluating the available literature in AI, as it is a diverse and rapidly developing area. The flexible nature of the search allowed us to include a variety of study designs and broad intervention types. This provided a comprehensive review of the available evidence. In accordance with PRISMA ScR guidelines, we conducted a scoping review. Given the rapid growth of AI technology, this review provides a comprehensive overview that represents a timely cross-section of the available research in this area.

Currently, AI-related research is increasing in the field of emergency medicine. AI has the potential to significantly reduce the time it takes to diagnose and treat patients in emergency situations. It can use vast amounts of data to quickly identify patterns, predict outcomes, and provide more accurate diagnoses and treatments. AI algorithms are being developed to identify patterns in large amounts of medical data that may help make more accurate predictions about patient outcomes. By leveraging this data, AI-based models may be able to provide more accurate diagnoses and suggest better treatment plans than standard medical practice. However, AI models must be rigorously tested before they can be widely implemented in clinical settings, as the potential risks of using AI-based systems must be taken into account. Despite the numerous research studies and evidence-based findings on effective interventions, many practitioners lack the knowledge or resources to translate the research into practice. This creates a gap between what is known to be effective and what is actually being implemented in practice. On the other hand, AI applications in healthcare are often limited by the need to integrate the technology into existing workflows and processes, while health care systems are often slow to adapt to new technology. This creates a disconnect between research and practice, which needs to be addressed. To bridge the gap, research should emphasize the implications of AI systems with higher action autonomy, enabling practitioners to leverage the knowledge gained from existing implementations and apply it to new contexts. This is especially important given the wide variety of AI systems being used and the potential for unintended consequences as a result of their use. By researching the implementation of AI systems in emergency medicine, we can better understand and mitigate any potential risks. The first stream of research should focus on understanding the different types of AI systems being used and how they interact with healthcare professionals and patients. The second stream should focus on the ethical implications of AI, such as privacy and data security, and the potential to create algorithmic bias in decision-making.

5. Limitations

The present study encountered several limitations. Firstly, the scoping review method was used in this work. Therefore, the evaluation process was not as rigorous as a systematic review, as a scoping review provides a general perspective of the topic and can identify themes. This study provides an overview of general trends and serves as a basis for further studies on the topic under review. Second, the study topic was unique in that there were few grey papers. Therefore, no grey papers on artificial intelligence were considered in the ED search. In this paper, we attempted to conduct a complete search that did not miss any references; however, references may still have been missed. Third, only studies explicitly classified as AI in ED by study authors were eligible for our analysis. This means studies that did not directly identify themselves as AI in ED in their titles and abstracts may have been missed by our search strategy. Other limitations of our review were that we only examined English-language studies.

6. Conclusion

There has been a rapid increase in AI research in emergency medicine in recent years. Several studies have demonstrated the potential of AI in diverse contexts, particularly improving patient outcomes through predictive modelling. According to the synthesis of studies in our review, AI-based decision-making lacks transparency. This feature makes AI decision-making opaque.

7. Declarations

7.1. Acknowledgments

Not applicable.

7.2. Competing interests

The authors declare that they have no competing interests.

7.3. Fundings and supports

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

7.4. Authors' contribution

M.M.H conceived the research hypothesis. M.M.H, T.M.H, and S.A. designed the study. M.M.H, T.M.H, and K.Q performed the article screening. M.M.H, T.M.H, K.Q, and HR. K performed the data extraction and the quality assessment. M.M.H and T.M.H have shaped the manuscript with input from the entire team. All authors contributed to revising the work for important intellectual content, gave the final approval of the version to be published, and agreed on all aspects of the work, especially concerning its accuracy and integrity.

7.5. Ethics approval and consent to participate

Not applicable.

References

1. Wilson T. No longer science fiction, AI and robotics are transforming healthcare. PWC Accessed October. 2017;31:2021.
2. Langlotz CP. Will artificial intelligence replace radiologists? : Radiological Society of North America; 2019. p. e190058.
3. Shuaib A, Arian H, Shuaib A. The Increasing Role of Artificial Intelligence in Health Care: Will Robots Replace Doctors in the Future? Int J Gen Med. 2020; 13:891-6.
4. Chen M, Decary M. Artificial intelligence in healthcare: An essential guide for health leaders. Healthc Manage Forum. 2020;33(1):10-8.
5. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol. 2017;2(4): e000101-115.
6. Bitterman DS, Aerts HJ, Mak RH. Approaching autonomy in medical artificial intelligence. Lancet Digit Health. 2020;2(9):e447-e9.
7. Bin KJ, Melo AAR, da Rocha JGMF, de Almeida RP, Cobello Junior V, Maia FL, et al. The Impact of Artificial Intelligence on Waiting Time for Medical Care in an Urgent Care Service for COVID-19: Single-Center Prospective Study. JMIR Form Res. 2022;6(2):e29012-e.
8. Nyce A, Gandhi S, Freeze B, Bosire J, Ricca T, Kupersmith E, et al. Association of Emergency Department Waiting Times With Patient Experience in Admitted and Discharged Patients. J Patient Exp. 2021;8:23743735211011404-.

9. Kang D-Y, Cho K-J, Kwon O, Kwon J-m, Jeon K-H, Park H, et al. Artificial intelligence algorithm to predict the need for critical care in prehospital emergency medical services. Scand J Trauma Resusc Emerg Med. 2020;28(1):1-8.
10. Li X, Tian D, Li W, Hu Y, Dong B, Wang H, et al. Using artificial intelligence to reduce queuing time and improve satisfaction in pediatric outpatient service: A randomized clinical trial. Front Pediatr. 2022;10: 929834
11. Lorenzi NM, Riley RT, Blyth AJ, Southon G, Dixon BJ. Antecedents of the people and organizational aspects of medical informatics: review of the literature. J Am Med Inform Assoc. 1997;4(2):79-93.
12. Kukafka R, Johnson SB, Linfante A, Allegrante JP. Grounding a new information technology implementation framework in behavioral science: a systematic analysis of the literature on IT use. J Biomed Inform. 2003;36(3):218-27.
13. Gooding P, Kariotis T. Ethics and law in research on algorithmic and data-driven technology in mental health care: scoping review. JMIR Ment Health. 2021;8(6):e24668.
14. Čartolovni A, Tomičić A, Mosler EL. Ethical, legal, and social considerations of AI-based medical decision-support tools: A scoping review. Int J Healthc Inform Res. 2022;161:104738.
15. Murphy K, Di Ruggiero E, Upshur R, Willison DJ, Malhotra N, Cai JC, et al. Artificial intelligence for good health: a scoping review of the ethics literature. BMC Med Ethics. 2021;22(1):1-17.
16. Siala H, Wang Y. SHIFTing artificial intelligence to be responsible in healthcare: A systematic review. Soc Sci Med. 2022:114782.
17. Yokoi R, Eguchi Y, Fujita T, Nakayachi K. Artificial intelligence is trusted less than a doctor in medical treatment decisions: Influence of perceived care and value similarity. Int J Hum Comput Interact. 2021;37(10):981-90.
18. Choudhury A, Asan O. Role of artificial intelligence in patient safety outcomes: systematic literature review. JMIR Med Inform. 2020;8(7):e18599.
19. Yin J, Ngiam KY, Teo HH. Role of artificial intelligence applications in real-life clinical practice: systematic review. J Med Internet Res. 2021;23(4):e25759.
20. Trocin C, Mikalef P, Papamitsiou Z, Conboy K. Responsible AI for digital health: a synthesis and a research agenda. Inf Syst Front. 2021:1-19.
21. Wolff J, Pauling J, Keck A, Baumbach J. The economic impact of artificial intelligence in health care: systematic review. J Med Internet Res. 2020;22(2):e16866.
22. Secinaro S, Calandra D, Secinaro A, Muthurangu V, Biancone P. The role of artificial intelligence in healthcare: a structured literature review. BMC Med Inform Decis Mak. 2021;21:1-23.
23. de Hond AA, Leeuwenberg AM, Hooft L, Kant IM, Nijman SW, van Os HJ, et al. Guidelines and quality criteria for artificial intelligence-based prediction models in healthcare: a scoping review. NPJ Digit Med. 2022;5(1):2.
24. Mirbabaie M, Stieglitz S, Frick NR. Artificial intelligence in disease diagnostics: A critical review and classification on the current state of research guiding future direction. Health Technol (Berl). 2021;11(4):693-731.
25. Zidaru T, Morrow EM, Stockley R. Ensuring patient and public involvement in the transition to AI-assisted mental health care: A systematic scoping review and agenda for design justice. Health Expect. 2021;24(4):1072-124.
26. Gama F, Tyskbo D, Nygren J, Barlow J, Reed J, Svedberg P. Implementation frameworks for artificial intelligence translation into health care practice: scoping review. J Med Internet Res. 2022;24(1):e32215.
27. Alhashmi SF, Alshurideh M, Al Kurdi B, Salloum SA, editors. A systematic review of the factors affecting the artificial intelligence implementation in the health care sector. Proc Int Conf Artif Intell Comput Visi (AICV2020); 2020: Springer.
28. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol. 2005;8(1):19-32.
29. Wohlin C, editor Guidelines for snowballing in systematic literature studies and a replication in software engineering. Proc 18th Int Conf Eval Assess Softw Eng; 2014.
30. Eldawlatly A, Alshehri H, Alqahtani A, Ahmad A, Al-Dammas F, Marzouk A. Appearance of Population, Intervention, Comparison, and Outcome as research question in the title of articles of three different anesthesia journals: A pilot study. Saudi J Anaesth. 2018;12(2):283.
31. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol. 2006;3(2):77-101.
32. Munn Z, Peters MD, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol. 2018;18:1-7.
33. Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, et al. A guide to deep learning in healthcare. Nat Med. 2019;25(1):24-9.
34. Mzoughi H, Njeh I, Wali A, Slima MB, BenHamida A, Mhiri C, et al. Deep multi-scale 3D convolutional neural network (CNN) for MRI gliomas brain tumor classification. J Digit Imaging. 2020;33(4):903-15.
35. Safaripour R, June Lim HJ. Comparative analysis of machine learning approaches for predicting frequent emergency department visits. J Health Inform.

2022;28(2):14604582221106396.

36. Feretzakis G, Sakagianni A, Kalles D, Loupelis E, Panteris V, Tzelves L, et al. Using Machine Learning for Predicting the Hospitalization of Emergency Department Patients. Stud Health Technol Inform. 2022;295:405-8.
37. Feretzakis G, Karlis G, Loupelis E, Kalles D, Chatzikyriakou R, Trakas N, et al. Using Machine Learning Techniques to Predict Hospital Admission at the Emergency Department. J Crit Care Med (Targu Mures). 2022;8(2):107-16.
38. Wilson MH, Habig K, Wright C, Hughes A, Davies G, Imray CH. Pre-hospital emergency medicine. Lancet. 2015;386(10012):2526-34.
39. Morrill J, Qirko K, Kelly J, Ambrosy A, Toro B, Smith T, et al. A Machine Learning Methodology for Identification and Triage of Heart Failure Exacerbations. J Cardiovasc Transl Res. 2022;15(1):103-15.
40. Blomberg SN, Folke F, Ersbøll AK, Christensen HC, Torp-Pedersen C, Sayre MR, et al. Machine learning as a supportive tool to recognize cardiac arrest in emergency calls. Resuscitation. 2019;138:322-9.
41. Al-Dury N, Ravn-Fischer A, Hollenberg J, Israelsson J, Nordberg P, Strömsöe A, et al. Identifying the relative importance of predictors of survival in out of hospital cardiac arrest: a machine learning study. Scand J Trauma Resusc Emerg Med. 2020;28(1):1-8.
42. Lin AX, Ho AFW, Cheong KH, Li Z, Cai W, Chee ML, et al. Leveraging machine learning techniques and engineering of multi-nature features for national daily regional ambulance demand prediction. Int J Environ Res Public Health. 2020;17(11):4179.
43. Fernandes M, Vieira SM, Leite F, Palos C, Finkelstein S, Sousa JM. Clinical decision support systems for triage in the emergency department using intelligent systems: a review. Artif Intell Med. 2020;102:101762.
44. Feretzakis G, Sakagianni A, Loupelis E, Kalles D, Panteris V, Tzelves L, et al. Prediction of Hospitalization Using Machine Learning for Emergency Department Patients. Stud Health Technol Inform. 2022;294:145-6.
45. Hwang S, Lee B. Machine learning-based prediction of critical illness in children visiting the emergency department. Plos one. 2022;17(2):e0264184.
46. Levin S, Toerper M, Hamrock E, Hinson JS, Barnes S, Gardner H, et al. Machine-learning-based electronic triage more accurately differentiates patients with respect to clinical outcomes compared with the emergency severity index. Ann Emerg Med. 2018;71(5):565-74. e2.
47. Choi SW, Ko T, Hong KJ, Kim KH. Machine learning-based prediction of Korean triage and acuity scale level in emergency department patients. Healthc Inform Res. 2019;25(4):305-12.
48. Kwon J-m, Lee Y, Lee Y, Lee S, Park H, Park J. Validation of deep-learning-based triage and acuity score using a large national dataset. Plos one. 2018;13(10):e0205836.
49. Yu JY, Jeong GY, Jeong OS, Chang DK, Cha WC. Machine learning and initial nursing assessment-based triage system for emergency department. Healthc Inform Res. 2020;26(1):13-9.
50. Farahmand S, Shabestari O, Pakrah M, Hossein-Nejad H, Arbab M, Bagheri-Hariri S. Artificial intelligence-based triage for patients with acute abdominal pain in emergency department; a diagnostic accuracy study. Adv J Emerg Med. 2017;1(1): e5
51. Kim SW, Li JY, Hakendorf P, Teubner DJ, Ben-Tovim DI, Thompson CH. Predicting admission of patients by their presentation to the emergency department. Emerg Med Australas. 2014;26(4):361-7.
52. Hong WS, Haimovich AD, Taylor RA. Predicting hospital admission at emergency department triage using machine learning. Plos one. 2018;13(7):e0201016.
53. Raita Y, Goto T, Faridi MK, Brown DF, Camargo CA, Hasegawa K. Emergency department triage prediction of clinical outcomes using machine learning models. Crit Care. 2019;23(1):1-13.
54. Goto T, Camargo CA, Faridi MK, Freishtat RJ, Hasegawa K. Machine learning–based prediction of clinical outcomes for children during emergency department triage. JAMA Netw Open. 2019;2(1):e186937-e.
55. Zhang X, Kim J, Patzer RE, Pitts SR, Patzer A, Schrager JD. Prediction of emergency department hospital admission based on natural language processing and neural networks. Methods Inf Med. 2017;56(05):377-89.
56. Chen C-H, Hsieh J-G, Cheng S-L, Lin Y-L, Lin P-H, Jeng J-H. Emergency department disposition prediction using a deep neural network with integrated clinical narratives and structured data. Int J Med Inform. 2020;139:104146.
57. Sterling NW, Patzer RE, Di M, Schrager JD. Prediction of emergency department patient disposition based on natural language processing of triage notes. Int J Med Inform. 2019;129:184-8.
58. Fernandes M, Mendes R, Vieira SM, Leite F, Palos C, Johnson A, et al. Predicting Intensive Care Unit admission among patients presenting to the emergency department using machine learning and natural language processing. Plos one. 2020;15(3):e0229331.
59. Resnikoff S, Lansingh VC, Washburn L, Felch W, Gauthier T-M, Taylor HR, et al. Estimated number of ophthalmologists worldwide (International Council of Ophthalmology update): will we meet the needs? Br J Ophthalmol. 2020;104(4):588-92.
60. Vu T, Nguyen A, Brown N, Hughes J, editors. Identifying patients with pain in emergency departments using

conventional machine learning and deep learning. Proc 17th Annu Australas Lang Technol Assoc; 2019: Australasian Language Technology Association (ALTA).
61. Pineda AL, Ye Y, Visweswaran S, Cooper GF, Wagner MM, Tsui FR. Comparison of machine learning classifiers for influenza detection from emergency department free-text reports. J Biomed Inform. 2015;58:60-9.
62. Sanderson M, Bulloch AG, Wang J, Williams KG, Williamson T, Patten SB. Predicting death by suicide following an emergency department visit for parasuicide with administrative health care system data and machine learning. EClinicalMedicine. 2020;20:100281.
63. Lindsey R, Daluiski A, Chopra S, Lachapelle A, Mozer M, Sicular S, et al. Deep neural network improves fracture detection by clinicians. Proc Natl Acad Sci. 2018;115(45):11591-6.
64. Olczak J, Fahlberg N, Maki A, Razavian AS, Jilert A, Stark A, et al. Artificial intelligence for analyzing orthopedic trauma radiographs: deep learning algorithms—are they on par with humans for diagnosing fractures? Acta Orthop. 2017;88(6):581-6.
65. Jang D-H, Kim J, Jo YH, Lee JH, Hwang JE, Park SM, et al. Developing neural network models for early detection of cardiac arrest in emergency department. Am J Emerg Med. 2020;38(1):43-9.
66. Kwon JM, Kim KH, Jeon KH, Lee SY, Park J, Oh BH. Artificial intelligence algorithm for predicting cardiac arrest using electrocardiography. Scand J Trauma Resusc Emerg Med. 2020 Dec;28:1-10.
67. Acharya UR, Oh SL, Hagiwara Y, Tan JH, Adam M, Gertych A, et al. A deep convolutional neural network model to classify heartbeats. Comput Biol Med. 2017;89:389-96.
68. Kwon Jm, Lee Y, Lee Y, Lee S, Park J. An algorithm based on deep learning for predicting in-hospital cardiac arrest. J Am Heart Assoc. 2018;7(13):e008678.
69. Johnsson J, Björnsson O, Andersson P, Jakobsson A, Cronberg T, Lilja G, et al. Artificial neural networks improve early outcome prediction and risk classification in out-of-hospital cardiac arrest patients admitted to intensive care. Crit Care. 2020;24(1):1-12.
70. Kim J, Chang H, Kim D, Jang D-H, Park I, Kim K. Machine learning for prediction of septic shock at initial triage in emergency department. J Critical Care. 2020;55:163-70.
71. Perng J-W, Kao I-H, Kung C-T, Hung S-C, Lai Y-H, Su C-M. Mortality prediction of septic patients in the emergency department based on machine learning. J Clinical Med. 2019;8(11):1906.
72. Koh JEW, Raghavendra U, Gudigar A, Ping OC, Molinari F, Mishra S, et al. A novel hybrid approach for automated detection of retinal detachment using ultrasound images. Comput Biol Med. 2020;120:103704.
73. Patel SJ, Chamberlain DB, Chamberlain JM. A machine learning approach to predicting need for hospitalization for pediatric asthma exacerbation at the time of emergency department triage. Acad Emerg Med. 2018;25(12):1463-70.
74. Ho AFW, Zheng H, Cheong KH, En WL, Pek PP, Zhao X, et al. The relationship between air pollution and all-cause mortality in Singapore. Atmosphere (Basel). 2019;11(1):9.
75. Cheong KH, Ngiam NJ, Morgan GG, Pek PP, Tan BY-Q, Lai JW, et al. Acute health impacts of the Southeast Asian transboundary haze problem—A review. Int J Environ Res Public Health. 2019;16(18):3286.
76. Liu Q, Yang L, Peng Q. Artificial Intelligence Technology-Based Medical Information Processing and Emergency First Aid Nursing Management. Comput Math Methods Med. 2022;2022:8677118.
77. Wu L, Huang G, Yu X, Ye M, Liu L, Ling Y, et al. Deep Learning Networks Accurately Detect ST-Segment Elevation Myocardial Infarction and Culprit Vessel. Front Cardiovasc Med. 2022;9:797207.
78. Yamanaka S, Goto T, Morikawa K, Watase H, Okamoto H, Hagiwara Y, et al. Machine Learning Approaches for Predicting Difficult Airway and First-Pass Success in the Emergency Department: Multicenter Prospective Observational Study. Interact J Med Res. 2022;11(1):e28366.
79. Whitt W, Zhang X. Forecasting arrivals and occupancy levels in an emergency department. Oper Res Health Care. 2019;21:1-18.
80. Khaldi R, El Afia A, Chiheb R. Forecasting of weekly patient visits to emergency department: real case study. Procedia Comput Sci. 2019;148:532-41.
81. Yousefi M, Yousefi M, Fathi M, Fogliatto FS. Patient visit forecasting in an emergency department using a deep neural network approach. Kybernetes. 2019; 49(9): 2335-48
82. Ho AFW, To BZYS, Koh JM, Cheong KH. Forecasting hospital emergency department patient volume using internet search data. IEEE Access. 2019;7:93387-95.
83. Ram S, Zhang W, Williams M, Pengetnze Y. Predicting asthma-related emergency department visits using big data. IEEE J Biomed Health Inform. 2015;19(4):1216-23.
84. Kuo Y-H, Chan NB, Leung JM, Meng H, So AM-C, Tsoi KK, et al. An integrated approach of machine learning and systems thinking for waiting time prediction in an emergency department. Int J Med Inform. 2020;139:104143.
85. Greenbaum NR, Jernite Y, Halpern Y, Calder S, Nathanson LA, Sontag DA, et al. Improving documentation of presenting problems in the emergency department using a domain-specific ontology and

machine learning-driven user interfaces. Int J Med Inform. 2019;132:103981.

86. Tootooni MS, Pasupathy KS, Heaton HA, Clements CM, Sir MY. CCMapper: An adaptive NLP-based free-text chief complaint mapping algorithm. Comput Biol Med. 2019;113:103398.
87. Frost DW, Vembu S, Wang J, Tu K, Morris Q, Abrams HB. Using the electronic medical record to identify patients at high risk for frequent emergency department visits and high system costs. Am J Med. 2017;130(5):601. e17-. e22.
88. Yu H, Yang LT, Zhang Q, Armstrong D, Deen MJ. Convolutional neural networks for medical image analysis: state-of-the-art, comparisons, improvement and perspectives. Neurocomputing. 2021 Jul 15;444:92-110.
89. Convertino VA, Grudic G, Mulligan J, Moulton S. Estimation of individual-specific progression to impending cardiovascular instability using arterial waveforms. J Appl Physiol. 2013;115(8):1196-202.
90. Guillame-Bert M, Dubrawski A, Wang D, Hravnak M, Clermont G, Pinsky MR. Learning temporal rules to forecast instability in continuously monitored patients. J Am Med Inform Assoc. 2017;24(1):47-53.
91. Mao Q, Jay M, Hoffman JL, Calvert J, Barton C, Shimabukuro D, et al. Multicentre validation of a sepsis prediction algorithm using only vital sign data in the emergency department, general ward and ICU. BMJ open. 2018;8(1):e017833.
92. Kijpaisalratana N, Sanglertsinlapachai D, Techaratsami S, Musikatavorn K, Saoraya J. Machine learning algorithms for early sepsis detection in the emergency department: A retrospective study. Int J Med Inform. 2022;160:104689.
93. Monkaresi H, Calvo RA, Yan H. A machine learning approach to improve contactless heart rate monitoring using a webcam. IEEE J Biomed Health Inform. 2013;18(4):1153-60.
94. Simjanoska M, Gjoreski M, Gams M, Madevska Bogdanova A. Non-invasive blood pressure estimation from ECG using machine learning techniques. Sens. 2018;18(4):1160.
95. Shandilya S, Kurz MC, Ward KR, Najarian K. Integration of Attributes from Non-Linear Characterization of Cardiovascular Time-Series for Prediction of Defibrillation Outcomes. Plos one. 2016;11(1):e0141313.
96. Shouval R, Hadanny A, Shlomo N, Iakobishvili Z, Unger R, Zahger D, et al. Machine learning for prediction of 30-day mortality after ST elevation myocardial infraction: An Acute Coronary Syndrome Israeli Survey data mining study. Int J Cardiol. 2017;246:7-13.
97. Allen C, Tsou M-H, Aslam A, Nagel A, Gawron J-M. Applying GIS and machine learning methods to Twitter data for multiscale surveillance of influenza. Plos one. 2016;11(7):e0157734.
98. Young T, Hazarika D, Poria S, Cambria E. Recent trends in deep learning based natural language processing. IEEE Comput Intell Mag. 2018 Jul 20;13(3):55-75.
99. Shen J, Pang R, Weiss RJ, Schuster M, Jaitly N, Yang Z, et al., editors. Natural tts synthesis by conditioning wavenet on mel spectrogram predictions. 2018 Proc IEEE Int Conf Acoust Speech Signal Process (ICASSP); 2018: IEEE.
100. Fernandez-Granero MA, Sanchez-Morillo D, Leon-Jimenez A. Computerised analysis of telemonitored respiratory sounds for predicting acute exacerbations of COPD. Sens. 2015;15(10):26978-96.
101. Luo G, Stone BL, Fassl B, Maloney CG, Gesteland PH, Yerram SR, et al. Predicting asthma control deterioration in children. BMC Med Inform Decis Mak. 2015;15(1):1-8.
102. Schroeder M, Lodemann S. A systematic investigation of the integration of machine learning into supply chain risk management. Logist. 2021 Sep 8;5(3):62.
103. Park SY, Kuo P-Y, Barbarin A, Kaziunas E, Chow A, Singh K, et al., editors. Identifying challenges and opportunities in human-AI collaboration in healthcare. CSCW 19 Companion (2019); 2019.
104. Yu K-H, Beam AL, Kohane IS. Artificial intelligence in healthcare. Nat Biomed Eng. 2018;2(10):719-31.
105. Panagopoulos A, Minssen T, Sideri K, Yu H, Compagnucci MC. Incentivizing the sharing of healthcare data in the AI Era. Comput Law Secur Rev. 2022;45:105670.
106. Reddy S, Allan S, Coghlan S, Cooper P. A governance model for the application of AI in health care. J Am Med Inform Assoc. 2020;27(3):491-7.
107. Kundu S. AI in medicine must be explainable. Nat Med. 2021;27(8):1328-.
108. Cutillo CM, Sharma KR, Foschini L, Kundu S, Mackintosh M, Mandl KD. Machine intelligence in healthcare—perspectives on trustworthiness, explainability, usability, and transparency. NPJ Digit Med. 2020;3(1):1-5.
109. Leslie D, Mazumder A, Peppin A, Wolters MK, Hagerty A. Does “AI” stand for augmenting inequality in the era of covid-19 healthcare? BMJ. 2021;372: n304
110. Spiegelhalter D. Should We Trust Algorithms? Harv Data Sci Rev. 2020 Jan 31;2(1). Available from: https://hdsr.mitpress.mit.edu/pub/56lnenzj
111. McCradden MD, Joshi S, Mazwi M, Anderson JA. Ethical limitations of algorithmic fairness solutions in health care machine learning. Lancet Digit Health. 2020;2(5):e221-e3.
112. Nyholm S. Humans and robots: Ethics, agency, and anthropomorphism: Rowman & Littlefield Publishers;

2020.
113. McGregor L. Accountability for governance choices in artificial intelligence: afterword to Eyal Benvenisti's foreword. Eur J Soc Secur. 2018;29(4):1079-85.
114. Santos HC, Varnum ME, Grossmann I. Global increases in individualism. Psychol Sci. 2017;28(9):1228-39.
115. Prainsack B, Buyx A. Solidarity: reflections on an emerging concept in bioethics: Nuffield Council on Bioethics London; 2011.
116. Academy B, Society tR. Data management and use: governance in the 21st century. The British Academy and the Royal Society London; 2017.
117. Kotschy W. The new General Data Protection Regulation-Is there sufficient pay-off for taking the trouble to anonymize or pseudonymize data. 2016.
118. Magrani E, Oliveira RMd. We Are Big Data: New Technologies and Personal Data Management. Revista Científica sobre Cyberlaw do Centro de Investigação Jurídica do Ciberespaço. 2018; 5: 9-33
119. Lewis KM, Hardelid P. National data opt out programme: consequences for maternity statistics in England. Int J Popul Data Sci. 2020;5(1): 1126-33
120. Vayena E, Blasimme A, Cohen IG. Machine learning in medicine: addressing ethical challenges. PLoS Med. 2018;15(11):e1002689.
121. C. Weyerer J, F. Langer P, editors. Garbage in, garbage out: The vicious cycle of ai-based discrimination in the public sector. Proc 20th Int Conf Digit Gov Res; 2019.
122. Bostrom N, Yudkowsky E. The ethics of artificial intelligence. Artificial intelligence safety and security: Chapman and Hall/CRC; 2018. p. 57-69.
123. Ntoutsi E, Fafalios P, Gadiraju U, Iosifidis V, Nejdl W, Vidal ME, et al. Bias in data-driven artificial intelligence systems—An introductory survey. Wiley Interdiscip Rev Data Min Knowl Discov. 2020;10(3):e1356.
124. Jones ML, Kaufman E, Edenberg E. AI and the Ethics of Automating Consent. IEEE Secur Priv. 2018;16(3):64-72.
125. Johnson CY. America is a world leader in health inequality. Washington Post Available online: https://www washingtonpost com/news/wonk/wp/2017/06/05/america-is-a-world-leader-in-health-inequality/(accessed on 5 June 2017). 2017.
126. Dickman SL, Himmelstein DU, Woolhandler S. Inequality and the health-care system in the USA. Lancet. 2017;389(10077):1431-41.
127. Allen M, Pierce O. Medical errors are No. 3 cause of US deaths, researchers say. SAGE PUBLICATIONS INC 2455 TELLER RD, THOUSAND OAKS, CA 91320 USA; 2016.
128. Magrabi F, Ong M-S, Runciman W, Coiera E. An analysis of computer-related patient safety incidents to inform the development of a classification. J Am Med Inform Assoc. 2010;17(6):663-70.
129. Arnold T, Scheutz M. The "big red button" is too late: an alternative model for the ethical evaluation of AI systems. Ethics Inf Technol. 2018;20:59-69.
130. Abedi V, Goyal N, Tsivgoulis G, Hosseinichimeh N, Hontecillas R, Bassaganya-Riera J, et al. Novel screening tool for stroke using artificial neural network. Stroke. 2017;48(6):1678-81.
131. Hu Y-H, Tai C-T, Chen SC-C, Lee H-W, Sung S-F. Predicting return visits to the emergency department for pediatric patients: applying supervised learning techniques to the Taiwan National Health Insurance Research Database. Comput Methods Programs Biomed. 2017;144:105-12.
132. Yadav K, Sarioglu E, Choi HA, Cartwright IV WB, Hinds PS, Chamberlain JM. Automated outcome classification of computed tomography imaging reports for pediatric traumatic brain injury. Acad Emerg Med. 2016;23(2):171-8.
133. Bae Y, Moon HK, Kim SH. Predicting the mortality of pneumonia patients visiting the emergency department through machine learning. J Korean Soc Emerg Med. 2018;29(5):455-64.
134. Costantino G, Falavigna G, Solbiati M, Casagranda I, Sun BC, Grossman SA, et al. Neural networks as a tool to predict syncope risk in the Emergency Department. Ep Europace. 2017;19(11):1891-5.
135. Chen AY, Lu T-Y, Ma MH-M, Sun W-Z. Demand forecast using data analytics for the preallocation of ambulances. IEEE J Biomed Health Inform. 2015;20(4):1178-87.
136. Seki T, Tamura T, Suzuki M, Group S-KS. Outcome prediction of out-of-hospital cardiac arrest with presumed cardiac aetiology using an advanced machine learning technique. Resuscitation. 2019;141:128-35.
137. Redfield C, Tlimat A, Halpern Y, Schoenfeld DW, Ullman E, Sontag DA, et al. Derivation and validation of a machine learning record linkage algorithm between emergency medical services and the emergency department. J Am Med Inform Assoc. 2020;27(1):147-53.

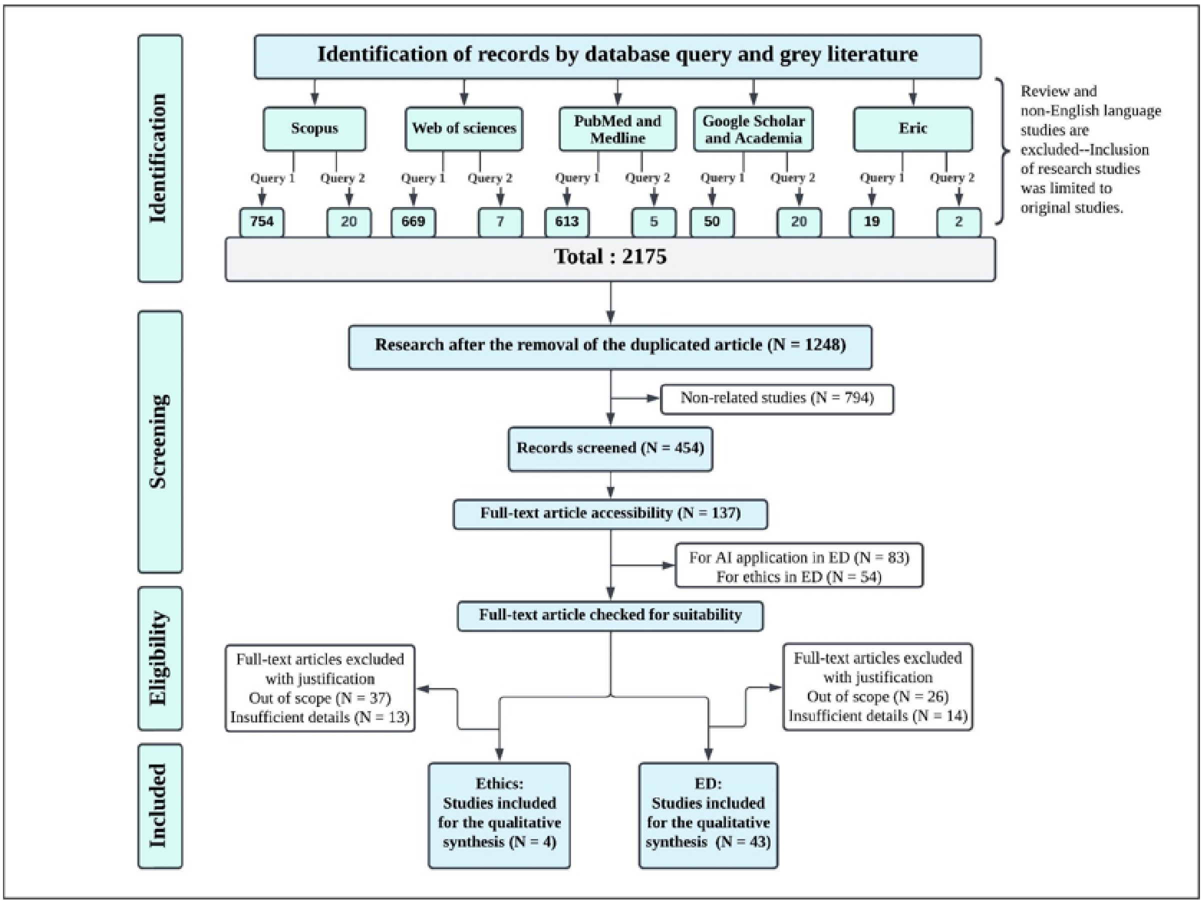

Figure 1: An overview of the manuscript selection process based on Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). ED: emergency department.

Table 1: Search strategy developed for the databases

Query (1)					
1	(((Artificial intelligen*) OR ((Machine OR Deep) adj0 learn*) OR ("Artificial neural network*"))				
2	(("Emergency Treatment") OR ("Emergency Medicine") OR ("emergency medical services") OR ("emergency service") OR (hospital) OR ("trauma centers") OR (triage) OR ("Evidence-Based Emergency Medicine") OR ("Emergency Nursing") OR (Emergencies) OR ("casualty department*") OR (ED) OR (accident) OR (ward) OR (wards) OR ("emergency unit") OR ("emergency units") OR ("emergency department*") OR (treatment*) OR (visit*) OR (critical care) OR (trauma))				
(1) AND (2)	**Scopus**	**Web of sciences**	**PubMed**	**Google Scholar and Academia**	**ERIC**
Article	754	669	613	50	19
Query (2)					
3	((ethical) OR (legal) OR (ethic*) OR ("moral code") OR (moral*) OR ("moral stand") OR ("moral principles") OR (ethos) OR (standards) OR (values) OR (beliefs) OR (manners) OR (etiquette) OR (virtuous) OR (righteous) OR (upright) OR (upstanding) OR ("high-minded") OR ("right-minded") OR (principled) OR (proper) OR (honorable) OR (honest) OR (incorruptible) OR (scrupulous) OR (respectable) OR (decent) OR (irreproachable) OR (truthful) OR (law-abiding) OR (clean-living) OR (chaste) OR (pure) OR (blameless) OR (sinless) OR ("code of ethics") OR (principles of right and) wrong OR ("principles of behavior") OR (scruples) OR (ideals)))				
((1) AND (2) AND (3))	**Scopus**	**Web of sciences**	**PubMed**	**Google Scholar and Academia**	**ERIC**
Article	20	7	5	20	2
Total	774	676	618	70	21

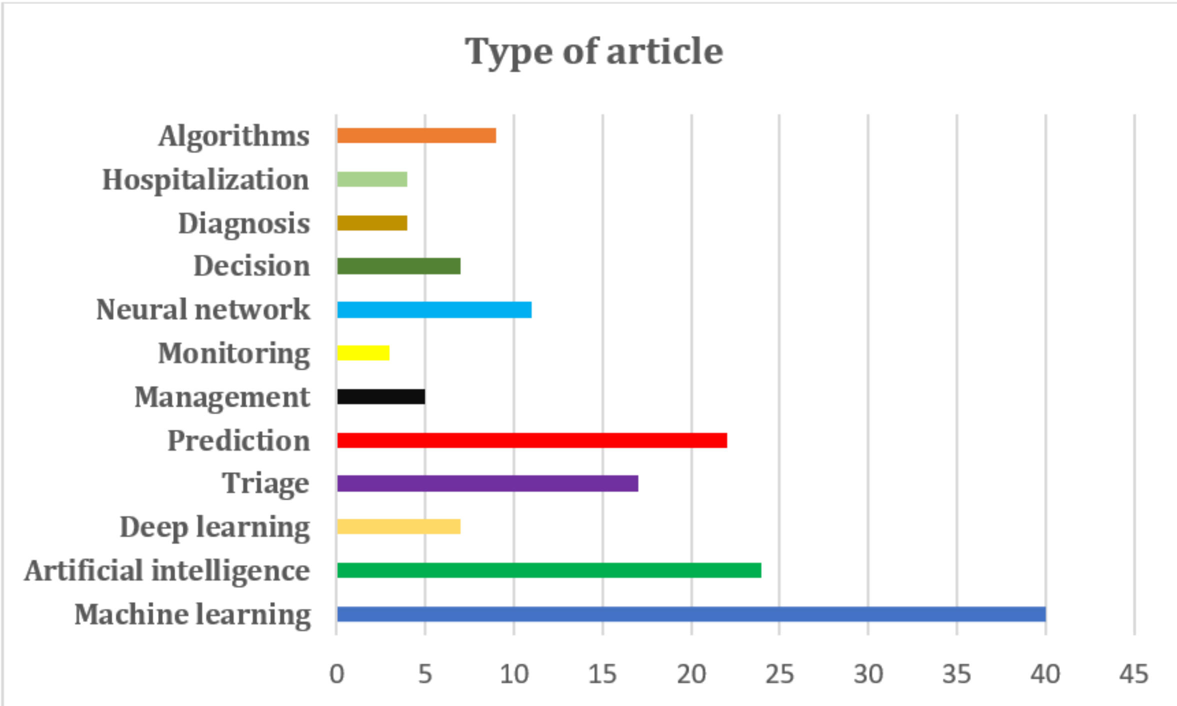

Figure 2: Study types included in this review based on the method and application of artificial intelligence in emergency department.

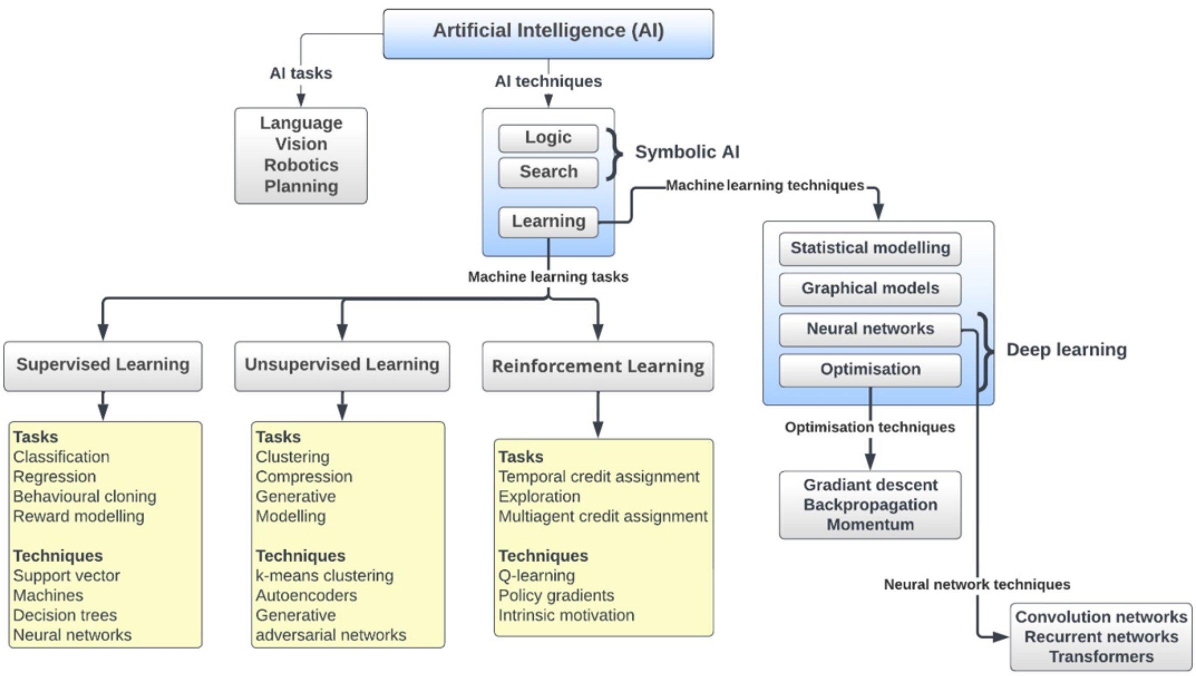

Figure 3: An overview of the AI algorithms used in the studies.

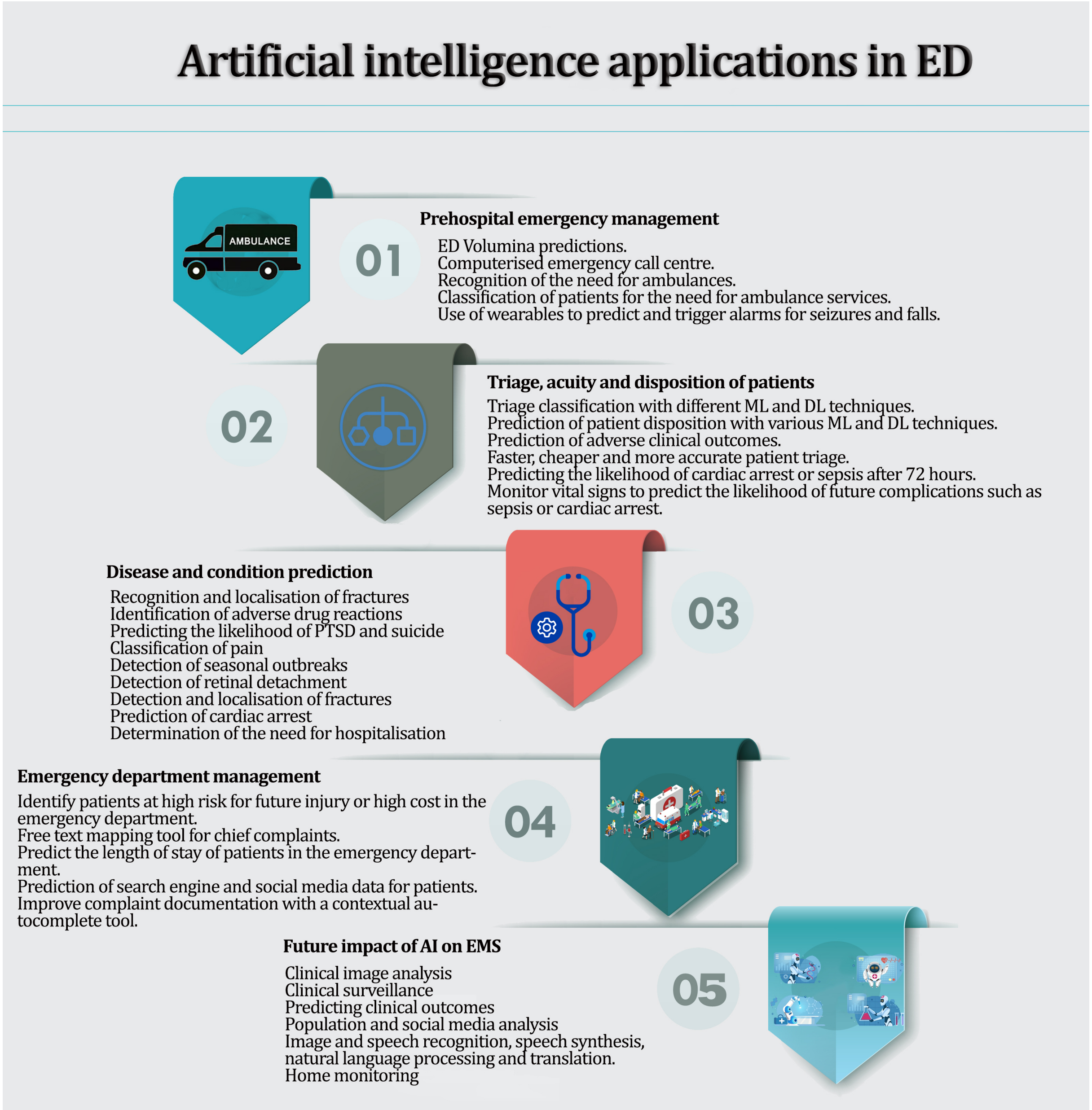

Figure 4: The applications of artificial intelligence (AI) presented in this overview in emergency department (ED). ML: machine learning; DL: deep learning; PTSD: Post-traumatic stress disorder; EMS: emergency medical services.

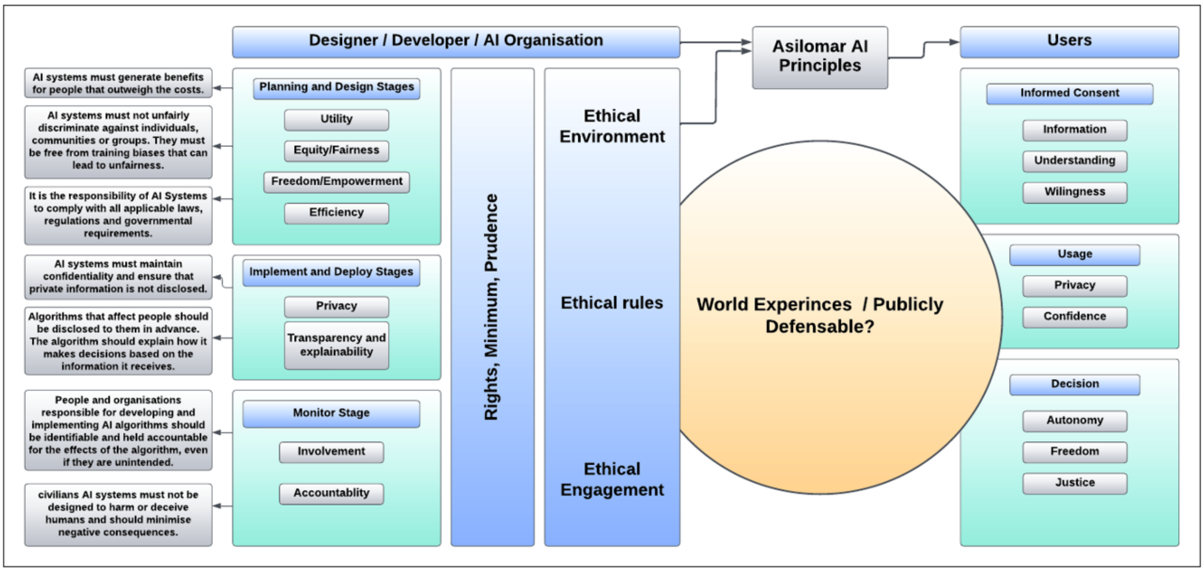

Figure 5: Proposal for an ethical framework for the use of artificial intelligence (AI) in healthcare systems.

Table 2: Summary of the report on the type of artificial intelligence (AI) technology used in emergency medicine in 2022

Population	Methods	Aims	Outcome
Razieh Safaripour (35)			
Between 2008 and 2015, a total of 9348 patients visited ED at least once, with 15,627 visits. A nationally representative population-based study was used for the study.	The study used logistic regression with RFE, Randobest-fitting SVM to identify the best fitting regression and create a more accurate classifier. It also used the Scikit-Learn integrated function GridSrearchCV with 10-fold cross-validation.	The study aimed to compare the performance of logistic regression and four other machine learning classification models in predicting frequent consumption ED.	Random Forest proved the best method with the highest precision and the lowest classification error.
Georgios Feretzakis (36)			
N= 3,204 ED visits were evaluated during the study period (14 March - 4 May 2019).	The study evaluated several classifiers, including Naive Bayes, Multinomial Logistic Regression, AdaBoost, LogitBoost, Classification via Regression, Random Forest, Bagging and Multilayer Perceptron.	The study aims to find an algorithm that uses ML techniques to support clinical decision-making in an emergency.	The study showed that logistic regression is the most frequently used technique for model design.
Georgios Feretzakis (37)			
N= 13,991 ED visits	The study used the caret package in R, a well-known framework for creating ML models. The caret package offers functions that facilitate model training for complex regression and classification tasks.	The study examined coagulation tests and biochemical markers routinely used in patients visiting ED in relation to ED outcome.	The multiple decision tree algorithm proved most suitable in pulmonology ED.

ED: emergency department; RFE: recursive feature elimination; SVM: Support Vector Machine; ML: machine learning.

Table 3: Summary of the report on artificial intelligence (AI) applications in emergency medicine (between 2014 and 2022)

Population	Methods	Amis	Outcome
Prehospital emergency management			
James Morrill (39)			
A training set of 1900-patient scenarios was developed, and a validation set of 100 cases was selected.	Machine learning	The study aimed to develop and validate a triage method for heart failure exacerbations. The methodology uses machine learning methods based on patient profile data and unique symptom responses.	The algorithm developed in this study performed exceptionally well compared to individual physicians in assessing the likelihood of exacerbation in a patient and determining the appropriate consensus triage category.
Da-Young Kang's (9)			
Multicenter retrospective cohort study; 9,304,887 adult patients visiting EDs of 151 hospitals	Deep learning	The study predicts the severity of EMS using an AI algorithm.	This study showed that an AI algorithm accurately predicted the need for critical care in a prehospital EMS. The performance of the AI algorithm exceeded that of conventional triage tools and scoring systems and was better than the accuracy of medical staff decision-making.
Stig Nikolaj Blomberg (40)			
N= 110,000	Machine learning	The aim of the study was threefold: to test whether a unique machine learning system could improve OHCA detection rates compared to trained dispatchers, to test whether the machine learning system could detect OHCA faster than the medically trained dispatchers, and to identify possible subgroups of callers or patients who were more susceptible to bias by the medical dispatchers or the machine learning system.	Machine learning overcomes hurdles in medical emergency detection to identify OHCA in raw audio files.
Al-Dury (41)			
N= 45,067 cases of OHCA recorded between 2008 and 2016.	The study used Random Forest, a machine learning algorithm.	The study examined the relative importance of 16 recognized factors in OHCA using a machine learning approach. The results suggest that such models are superior to regression models.	Using machine learning, the study has ranked the factors that influence 30-day survival after out-of-hospital cardiac arrest and found that initial rhythm is the most important predictor of survival.
Adrian Xi Lin (42)			
N= 60 ambulances	Three methods were considered in the study: Radial Basis Function Network (RBFN), Light Gradient Boosting Machine (LightGBM) and MLP with Radial Basis Function Network (RBFN).	The study proposes a novel approach to predicting ambulance demand in different regions of an urban state based on a large dataset of historical ambulance demand data.	The study used a citywide 10-year dataset of emergency ambulances to predict demand for ambulances and found that LightGBM performed best.
Triage, acuity, and disposition of patients			
Soyun Hwang (45)			
Entire dataset (n= 2621710) Under sampled dataset–Critical cases (n= 25902) Under sampled dataset–Hospitalization (n= 607616)	RF model And Random Forest Classifier; Python's Scikit-Learn library	The study developed a machine learning-based classification model to predict the clinical course of paediatric ED visitors and compared its performance with the conventional paediatric triage system of South Korea.	The model RF performed well in discriminating critical cases with an AUROC of 0.991 and an AUPRC of 0.640.

Table 3: Summary of the report on artificial intelligence (AI) applications in emergency medicine (between 2014 and 2022)

Population	Methods	Amis	Outcome
Scott Levin (46)			
retrospective cohort study N= 172,726 adult ED visits	Three distinct decision tree learning models	The study used machine learning methods to develop an electronic triage support system that predicts clinically meaningful patient outcomes.	A random forest model was developed that considered only demographics, mode of onset, vital signs, and clinical grouping of chief complaints. An updated random forest model was developed in which chief complaint was specified as the new optimal mix of clinical groupings and individual complaints recorded in the electronic health record.
Sae Won Choi (47)			
N= 142,972 patients.	The study used logistic regression, random forest, XGBoost, Ridge logistic regression, L2 penalty, NLP techniques	This study aimed to train machine learning models and compare their ability to predict KTAS scores.	The results showed that the Boost models outperformed other machine learning algorithms, with XGBoost achieving the best F1 score (=0.740).
Joon-myoung Kwon (48)			
retrospective observational cohort study; N= 11,656,559 ED visitor	The study developed a DTAS using a multilayer perceptron, a deep learning method with five hidden layers.	The study developed a Deep-learning-based TAS (DTAS)	The results showed that DTAS was more accurate than KTAS, MEWS, and RF Random Forest.
Jae Yong Yu (49)			
retrospective study; N= 145,784	The study used multivariate logistic regression analysis with the R package 'glm' and Deep Learning with the R package 'Keras'	The study aimed to evaluate an ML and INA-based ED triage system for predicting adverse clinical outcomes.	The study developed an ML and INA-based triage system for emergency departments. The novel system could predict clinical outcomes more accurately than existing triage systems KTAS and SOFA.
Shervin Farahmand (50)			
Prospective accuracy; N= 215 patients	A mixed-model approach that includes association rules (AR), clustering (CL), logistic regression (LR), decision tree (DT), Naïve Bayes (NB) and neural network (NN) algorithms.	This study was conducted to evaluate the use of AI-based tools in patients presenting with the chief complaint of acute abdominal pain to determine the ESI-4 score without estimating the resources required.	The study found that an AI-based triage model can accurately and independently identify patients with triage levels 3 and 4 without estimating resource use.
Susan W. Kim (51)			
N= 100,123	Logistic regression	The study aimed to determine the accuracy of a disposition decision made by an ED triage nurse as soon as a patient presents to the ED, and whether blood testing requested by the clinicians improved the accuracy of the model.	The model was able to predict the likelihood of hospitalisation for 72% of patients presenting to ED, which is similar to the accuracy of the nurse's triage decision. Blood test results had only a slight effect on predicting patient admission.
Woo Suk Hong (52)			
N= 560,486 patient	A set of nine binary classifiers was trained with logistic regression, gradient boosting, and deep neural networks on three types of datasets.	The study aimed to use gradient boosting and deep neural networks to model the nonlinear relationships between variables in 560,486 patient visits and present a low-dimensional model that can be used as a clinical decision support tool.	The study showed that machine learning enables robust prediction of hospital admission at emergency department (ED) triage and that adding patient history significantly improves predictive performance compared to using triage information alone, highlighting the need to include these variables in predictive models.

Table 3: Summary of the report on artificial intelligence (AI) applications in emergency medicine (between 2014 and 2022)

Population	Methods	Amis	Outcome
Yoshihiko Raita (53)			
N= 2782 ED visits	Four models were used in the study: logistic regression with lasso regularisation, random forest, gradient-boosted decision tree and deep neural network.	The study used extensive ED visit data to develop machine learning models using routinely available triage data to predict clinical outcomes after ED triage accurately. It also examined the predictive performance of these models compared to the model using the traditional five-step ESI algorithm.	Machine learning models were found to be particularly powerful in predicting outcomes of intensive care and hospitalisation.
Tadahiro Goto (54)			
N= 52,037 ED child visits	Four models were used in the study: logistic regression with lasso regularisation, random forest, gradient-boosted decision tree, and deep neural network.	The study analysed nationally representative visit data from ED to develop machine learning-based triage models that predict children's clinical course after ED triage. It compared their predictive performance with the reference model, which uses a conventional 5-level triage classification.	The machine learning models achieved higher sensitivity for predicting the outcome of intensive care and higher specificity for predicting the outcome of hospitalisation.
Xingyu Zhang (55)			
The study population of 54,320 visited patients.	Multivariable logistic regression and multi-layer neural networks.	The study described and compared logistic regression and neural network modelling strategies for predicting hospital admissions or transfers after initial presentation to ED triage.	The models performed well in predicting hospitalisation of ED paediatric patients, and accuracy improved without the influence of missing values. Logistic regression and MLNN also showed good predictive performance. MLNN performed better than logistic regression, but logistic regression requires a priori assumptions about the relationships between variables. Logistic regression and neural networks showed similar predictive power for hospital admission in ED patients.
Chien-Hua Chen (56)			
The study population was based on data from 104,803 non-traumatic adult visits ED.	The study developed four models: the REMS, the logistic regression model, the DNN model with word embedding, and the DNN model with paragraph vectors. deep learning	The study used a deep neural network model with word embedding to learn the complex clinical decision-making process to contribute to dispositional decision-making.	The study used the Deep Learning modelling strategy to develop the BiLSTM-CNN-based disposition prediction model. The proposed model better-predicted hospitalisation and discharge outcomes with a higher F1 score.
Marta Fernandes (58)			
The study population was based on two countries: the Portuguese and the US data, with a total of 599276 and 267257 ED visits in the adult population, respectively.	The study used logistic regression with L2 regularisation, RUSBoost and Random Forests Regression bootstrap aggregation of decision trees to compare a boosting classification technique - RUSBoost - and a more traditional technique - LR.	The study used machine learning to identify ED patients at high risk for ICU admission.	An isotonic calibration model for BIDMC that included clinical variables and chief complaints performed better, overall. A multi-model combining clinical variables and chief complaint for HBA achieved a good specificity.

Table 3: Summary of the report on artificial intelligence (AI) applications in emergency medicine (between 2014 and 2022)

Population	Methods	Amis	Outcome
Disease and condition prediction			
Thanh Vu (60)			
A random sampling of 2,000 ED adult patients	ML: SVM and RF DL: RNNs and CNNs	The study proposed a machine learning model to categorise patients who come to the emergency department with pain.	It has been shown that Deep Learning models can compete with or even outperform conventional models.
Robert Lindsey (63)			
Two stages: 100,855 X-rays were used in the bootstrapping phase, and 31,490 X-rays were used in the second phase	Deep learning	The study developed a deep neural network for detecting and localising fractures in X-ray images.	The study showed that a Deep Learning model can be trained to detect wrist fractures in X-ray images and improve the diagnostic accuracy of emergency physicians.
Jakub Olczak (64)			
The study evaluated 256,458 X-rays of the hand (including scaphoid projections), wrist and ankle with associated radiologists' reports.	The study selected five common, freely available deep networks: • BVLC Reference CaffeNet network (8 layers) • VGG CNN S network (8 layers) • VGG CNN (16- and 19-layers networks) • Network-in-network (14 layers)	The study aimed to determine whether standard Deep Learning networks can be trained to detect fractures in orthopaedic radiographs. It also investigated whether Deep Learning can be used to determine additional features such as body part, examination view, and laterality.	The study has shown that Deep Learning networks can be applied to X-ray images of the skeleton and that this also works for other medical fields.
Dong-Hyun Jang (65)			
The study population was 374,605 ED visits	Neural network; The three models of ANN are a multilayer perceptron model (MLP), a sequential model with long-term, short-term memory (LSTM) and a hybrid model	The study aimed to identify optimised network designs of ANNs for cardiac arrest prediction in emergency departments and to test the performance of the trained networks.	The ANN model performed significantly better in the study than other classification techniques, but the difference was relatively small. This suggests that the ANN model may be more useful in other EDs where other tools perform poorly.
U. Rajendra Acharya (67)			
N= 260 samples	Deep learning Neural network; CNN	The study proposes a deep-learning approach for automatically detecting cardiac arrhythmias, coronary artery disease, and myocardial infarction in ECG signals.	The proposed CNN model was fully automatic, insensitive to ECG signal quality, uses a tenfold cross-validation strategy, and requires long training times. Cardiologists can use this tool to diagnose ECG heartbeat signals in the clinical setting, objectively.
Jesper Johnsson (69)			
N= 939 patients participated in the study.	Artificial neural network (ANN)	The study used an artificial neural network (ANN) to predict the long-term functional outcome of OHCA survivors using background, prehospital, and admission-centred data.	The study confirmed that the machine learning model was excellent at predicting neurological recovery and survival and also provided the opportunity to determine the impact of the intervention in different subgroups.
Joonghee Kim (70)			
N= 49,299 patients	Six basic algorithms of ML were evaluated, including SVM, GBM, RF, MARS, Lasso and Ridge Regression.	The study aimed to evaluate the performance of ML-based triage tools in screening patients with septic shock in ED.	The results showed that ML classifiers have high discriminatory power, even when only baseline data are available, and that they outperform traditional scores.

Table 3: Summary of the report on artificial intelligence (AI) applications in emergency medicine (between 2014 and 2022)

Population	Methods	Amis	Outcome
Jau-Woei Perng (71)			
N= 42,220 patients	Deep learning: Convolutional Neural Networks (CNNs). PCA: traditional machine learning algorithm Machine learning methods: K nearest neighbor (KNN), Support Vector Machine (SVM), SoftMax, and Random Forest (RF).	The study used various machine learning algorithms to create a better predictive model to identify patients at increased risk of sepsis.	In the study, CNN and SoftMax showed the highest accuracy in predicting mortality in patients with suspected infection in a hospital emergency department.
Joel En Wei Koh (72)			
N= 100 patients	——-	The paper presents a system for separating PVD and RD in all situations.	The study developed a novel system automatically distinguishing PVD from RD using US images. It uses philtres, HOS cumulants and LSDA and achieves the highest classification accuracy of 99.13%.
Shilpa J Patel (73)			
N= 29,392 patient	The study trained four models with the training data set: classification trees, logistic regressions with L1 (LASSO) regularisation, random forests, and gradient boosting machines.	The study aimed to compare the performance of four commonly used machine learning approaches in predicting the need for hospitalisation for paediatric asthma at the time of triage by combining available clinical data with weather information, neighbourhood characteristics, and community viral load.	The study developed a machine learning model to predict hospital admission for paediatric asthma using only data available during triage. The model was improved by adding weight, SES, and weather data. The result showed that the best cut-off point for predicting admission would be a consensus-based decision at the institutional level. In addition, initial oxygen saturation was the most important predictor of hospital admission.
AI in emergency department management			
Qing Liu (76)			
N= 321 critically ill patients	————	The study used the internet, 4G communications, a local area network, and other information technologies to develop a system that optimises the transmission channel for medical information and improves the success rate of patient treatment.	The results prove that the system can effectively improve the efficiency and accuracy of first aid. The triage efficiency of wounded patients was improved, and the new management mode increased the survival rate.
Lin Wu (77)			
N= 793 persons with STEMI and 478 control persons without STEMI. Total: 1259	Neural network CNN, LSTM, and CNN-LSTM	The study developed a neural convolutional network, long-term memory, and CNN-LSTM models for diagnosing STEMI and diseased vessels.	The study uses Deep Learning to detect STEMI and culprit vessels based on real ECG data, allowing for more accurate remote diagnosis.
Yamanaka S (78)			
N= 10,741	Machine learning models: logistic regression model with elastic-net, random forest, gradient boosting decision tree, multilayer perceptron, k-point nearest neighbor, XGBoost, and an ensemble model.	In the study, machine learning models were developed to predict difficult airways and first-pass success with excellent accuracy, and their performance was compared with conventional techniques.	The results show that machine learning models can predict difficult airways in ED and that these models outperform conventional approaches to predicting intubation outcomes. Furthermore, the models can be used for practice and as teaching tools.

Table 3: Summary of the report on artificial intelligence (AI) applications in emergency medicine (between 2014 and 2022)

Population	Methods	Amis	Outcome
Ward Whitt (79)			
N= 443 days	Neural networks	The study aimed to develop a method for predicting future daily arrivals and hourly emergency department occupancy rates based on the recent past.	The study examined five prediction methods for daily arrival numbers and their application to predict hourly occupancy numbers. It was found that the SARIMAX time series model, which exploits both exogenous variables (temperature and holiday effects) and internal dependencies, has the best predictive power.
Rohaifa Khaldi (80)			
A time series with mean equal to1545	The model combines Artificial Neural Networks (ANNs) with a signal decomposition technique called Ensemble Empirical Mode Decomposition (EEMD).	The study's main aim was to predict patients' weekly arrival at ED.	The results show that the model used outperforms the benchmarking models in terms of approximation and generalisability.
Milad Yousefi (81)			
N= 1,095 consecutive days	Deep learning Deep neural network; ANNs	This study aimed to investigate the factors that influence daily demand in an ED and to provide a forecasting tool in a public hospital for up to seven days.	The study results show that the proposed model performs better than traditional forecasting models in terms of R2 in regression analyses and MAPE.
WAH HO (82)			
N= 3081	Multiple Regression	Predicting patient visits to the hospital using search engine data	As a result of this study, the quality of medical care can be improved, which is especially important in emergencies.
Sudha Ram (83)			
The dataset for this study contains 464 845 785 general tweets and 1 315 390 asthma-related tweets.	Decision tree and ANN techniques, Naive Bayes and SVM techniques	The study used social media, internet searches and air quality data to build a model to predict asthma-related visits to ED in a relatively discrete geographical area within a relatively short period.	The analysis of the study results showed that the Twitter/environmental data model could predict asthma ED visits with an overall fairly high accuracy.
Yong-Hong Kuo (84)			
N= 12,440 observations	The four modelling approaches used in the study were linear regression (LR), artificial neural networks (NN), support vector machines (SVM), and gradient boosting machines (GB).	The study aimed to build predictive models for estimating patient waiting time based on real-time and historical operational data using machine learning techniques and to show that data-driven approaches combined with the concept of systems thinking can contribute to better predictive performance.	The models were generally superior to the linear regression models, with a 15-20% reduction in mean squared error.
NathanielR. Greenbaum (85)			
N= 279,231 triage notes And qualitative review of 150 patients	Machine learning-driven	The aim of the study was to show that machine learning, in combination with an application-specific ontology, can help in the prospective acquisition of structured data.	A contextual autocompletion system was introduced in the study to improve structured data entry and ontology adherence and to reduce data entry time.

Table 3: Summary of the report on artificial intelligence (AI) applications in emergency medicine (between 2014 and 2022)

Population	Methods	Amis	Outcome
Mohammd Samie Tootooni (86)			
N= 147,859	An NLP-based mapping algorithm	The study aimed to show that a consensus-based validation approach can improve human judgement in practice and ensure fairness in considering misclassification and misattribution.	The study proposed a novel NLP-based heuristic algorithm, called CCMapper, to classify free-text main complaints into one or more categories within a predefined structured list.
David W. Frost (87)			
N=101,174	————	The study aimed to derive and validate a text-based predictive algorithm for frequent emergency department visits and total healthcare costs.	The study showed that machine learning techniques could analyse free text in GP electronic medical records to predict frequent emergency department visits and high total system costs.

EMS: emergency medical services; OHCA: out-of-hospital cardiac arrest; MLP: Multi-layer Perceptron; RF: random forest; AUROC: area under the receiver operating characteristic curve; AUPRC: area under Precision-Recall curve;

Table 4: An overview of the general concepts used in artificial intelligence (AI) ethics

Authors	Concept	Ref
Effy Vayena	• In order to ensure data protection and privacy, MLm data acquisition must adhere to industry standards • In order for MLm development to be successful, fairness must be upheld • Transparency standards should be followed in the deployment of MLm	(120)
Meg Leta Jones	• Digital Consent • Improving Consent • Automating Consent	(124)
Thomas Arnold	• A comparison of designing AI systems for verification with testing less accessible systems • Considering ethical principles when anticipating and planning actions • Preventing manipulation of actions and principles • Maintaining diagnostic vigilance and performing opaque self-examinations • Integrating the ethical core into the organization to ensure ethical behavior • Firewall compromise	(129)
Ryosuke Yokoi	• Trusting AI • In this study, AI was found to be less trusted than a human doctor in general, even when it learned and recommended the treatment participants desired.	(17)

MLm: Machine Learning in medicine