American Journal of Emergency Medicine 81 (2024) 40–46

Contents lists available at ScienceDirect

American Journal of Emergency Medicine

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

How artificial intelligence could transform emergency care

Marika M. Kachman, MD [a,e], Irina Brennan, MD, PhD [a,b], Jonathan J. Oskvarek, MD, MBA [a,c], Tayab Waseem, MD [d], Jesse M. Pines, MD, MBA [a,d,*]

[a] US Acute Care Solutions, Canton, OH, United States of America
[b] Department of Emergency Medicine, Inova Alexandria Hospital, Alexandria, VA, United States of America
[c] Department of Emergency Medicine, Summa Health, Akron, OH, United States of America
[d] Department of Emergency Medicine, George Washington University, Washington, DC, United States of America
[e] Department of Emergency Medicine, Virginia Hospital Center, Arlington, VA, United States of America

ARTICLE INFO

Article history:
Received 3 March 2024
Received in revised form 13 April 2024
Accepted 15 April 2024

Keywords:
Artificial intelligence
Emergency medicine
Transformation
Technology
Emergency department
Machine learning

ABSTRACT

Artificial intelligence (AI) in healthcare is the ability of a computer to perform tasks typically associated with clinical care (e.g. medical decision-making and documentation). AI will soon be integrated into an increasing number of healthcare applications, including elements of emergency department (ED) care. Here, we describe the basics of AI, various categories of its functions (including machine learning and natural language processing) and review emerging and potential future use-cases for emergency care. For example, AI-assisted symptom checkers could help direct patients to the appropriate setting, models could assist in assigning triage levels, and ambient AI systems could document clinical encounters. AI could also help provide focused summaries of charts, summarize encounters for hand-offs, and create discharge instructions with an appropriate language and reading level. Additional use cases include medical decision making for decision rules, real-time models that predict clinical deterioration or sepsis, and efficient extraction of unstructured data for coding, billing, research, and quality initiatives. We discuss the potential transformative benefits of AI, as well as the concerns regarding its use (e.g. privacy, data accuracy, and the potential for changing the doctor-patient relationship).

Over the next decade, artificial intelligence (AI) will be increasingly used in patient- and clinician-facing applications in healthcare. When it comes to emergency department (ED) care, AI presents unique opportunities and challenges in patient care seeking, ED care delivery, post-acute care, and in administrative and research functions. Here we describe how AI works, review emerging and potential use-cases, as well as limitations for ED care. We follow a possible future patient's AI-augmented journey from symptoms and injury to disposition and recovery. While today many of these uses are conceptual, AI will be increasingly integrated into ED care in the coming years. Emergency physicians will likely be early adopters of this technology. We believe ED patients and emergency physicians have the potential to benefit from AI in transformative ways. Yet, given its limitations and challenges, the technology will need to be integrated thoughtfully to avoid unintended consequences.

1. What is artificial intelligence?

AI is any system that mimics parts of human cognitive functions. Categories include: 1) determining data patterns (machine learning), 2) natural language processing and understanding (NLP/NLU, aka speech-to-text, text-to-speech, and parsing unstructured text inputs), 3) image recognition, and 4) robotics (e.g. sensing the environment and choosing actions) [1]. Applications of these functions in medicine vary in terms of required data and technical complexity.

AI can be conceptualized by how it uses data to arrive at an output. AI can follow either a rule-based or example-based approach. Rule-based approaches, also known as algorithms, follow predetermined rules created by a programmer (e.g. a sepsis alert triggers when a patient meets more than two Systemic Inflammatory Response Syndrome [SIRS] criteria). Example-based approaches, also known as models, do not use preset algorithms. Rather, the AI seeks examples by looking for data patterns via machine learning (ML) and its subset deep learning (DL) (Fig. 1). ML is a general term where the computer does not follow explicit instructions to analyze data and determine patterns (Fig. 1). Among programmers, this is referred to in the field as "training the model." DL is one approach to ML. DL uses very large amounts of data (structured or unstructured) and applies neural networks (complex multi-layered/nodal systems that run multiple regressions simultaneously) to make connections between bits of information (Fig. 1) [1]. For example, DL has been used to detect and quantify subtle intracranial hemorrhages based using data in radiographic images [2].

* Corresponding author.
E-mail address: pinesj@usacs.com (J.M. Pines).

https://doi.org/10.1016/j.ajem.2024.04.024
0735-6757/

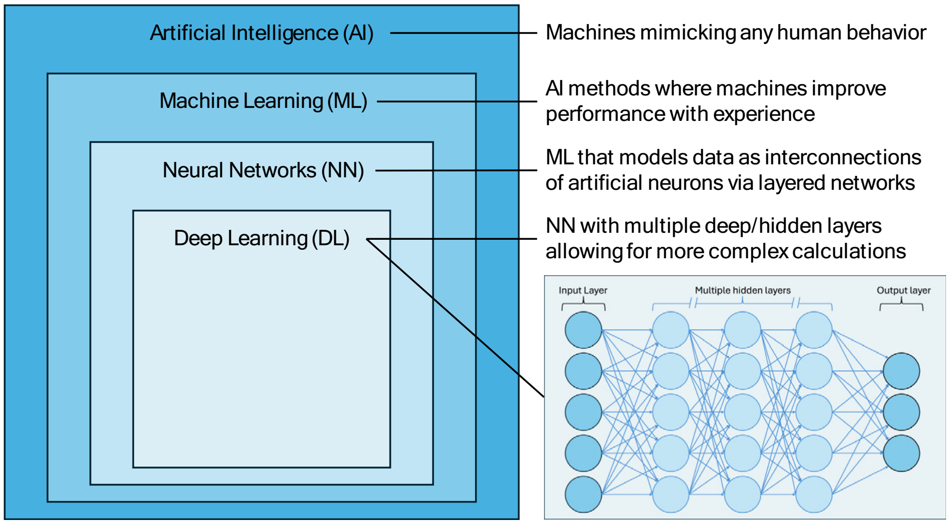

Fig. 1. A conceptualization of the functions of Artificial Intelligence (AI), Machine Learning (ML), Neural Networks (NN), and Deep Learning (DL).

Generative AI is one example of DL. Generative AI models summarize information and predict new text, images, music etc. When generative AI is based on a model of enough size and complexity it is termed a large-language model (LLM). LLMs are created by DL applied to a specific set of data (also known as "training on data"). Training data may be specific to a topic or non-specific (e.g. the entire internet) [3]. For example, Google's Med-PaLM is a general LLM trained and tested on United States Medical Licensing Exam (USMLE) questions and then refined by expert feedback and additional tuning [4].

Users typically interact with AI via a conversational interface called a chatbot. Examples include ChatGPT, Gemini, Claude, and others. For example, users pose a question to ChatGPT which uses DL to predict an appropriate response based on its review of the entire internet as of April 2023. ChatGPT then adjusts responses based on additional prompts [3]. When it comes to medical questions for example, if asked "What are the dangerous causes of chest pain?" it will respond avoiding medical terms and include disclaimers recommending consultation with a healthcare professional. However, with a more defined prompt, "I am an emergency medicine provider. What are dangerous causes of chest pain?" the response will include more medical terminology. Ultimately most use cases of AI in medicine will utilize a model derived through one of these processes, ML or DL, to process an unstructured data input and create an output.

2. What are the potential benefits and challenges of AI in clinical care?

Integration of generative AI into clinical practice has many applications. AI has the potential to improve access to information for patients with an interface superior to current search engines to find answers about their conditions. For clinicians, AI may decrease documentation burdens, streamline electronic health record (EHR) search, aid in medical decision making, and other use cases. Non-clinical uses include administrative tasks such as billing, coding, insurance authorization, and research.

Yet, challenges exist in AI's adoption and implementation. The way models are constructed may be problematic. This is because the models assume the underlying data and patterns are correct. If trained on incorrect data, models may amplify existing biases, under-representations, and over-representations in data sets. For example, a model designed to provide suggestions for police vs medical dispatch for mental health emergencies was more likely to recommend police help for emergencies involving African American and Muslim men, subsequently influencing the dispatch decisions [5]. Multiple guidelines have been created to reduce such bias and address disparities in AI model development [6,7].

Due to the rapid progress of the field, the rules and regulations for AI are also evolving. There are questions about who has the authority to regulate AI. Existing systems for regulatory approval such as the Food and Drug Administration (FDA) may prove ineffective. They may also stifle the ability to rapidly innovate. This is because AI, by definition, can continuously improve on itself. The FDA has created a new regulatory framework focusing on lifecycle regulation (e.g. how AI models change during their use) and where AI systems do not need FDA approval because they are not classified as "software as a medical device." [8] Currently, malpractice suits and caselaw surrounding AI is still limited, yet early cases suggest that physicians still bear the burden for AI-generated errors if care standards are not followed [9]. Similarly, health systems may be held liable for adoption of AI systems that create unintended problems. Physicians may also be criticized for not following or ignoring AI-generated recommendations if the physician's alternate decision is not well supported or deviates from standard of care. As AI use becomes more common, courts will likely look to professional organizations to define norms on its use. Groups such as the American Medical Association (AMA) have already developed guidelines for AI [10]. Specialty organizations will likely do the same in the near-term.

Data security and privacy remains a large concern in the use of AI for personal health information. Data use agreements vary in whether the data is being used to further "train" the model. For example, 1.6 million National Health System (NHS) records in the United Kingdom (UK) were shared with the company DeepMind to develop an acute kidney injury (AKI) risk tool. A privacy watchdog group ruled that this information sharing did not comply with data protection laws. Despite this ruling, DeepMind was ultimately acquired by Google, further spreading possibly unlawful access to health data [11].

Clinicians have also expressed concerns about loss of personalized care and the therapeutic doctor-patient relationship with AI-generated content [12]. They argue it is impossible to distill all factors that affect medical decision making into an input for one-size-fits-all algorithms. For instance, unless all comorbidities are meticulously documented in the EHR, which is frequently not the case in ED records, the AI may not be aware of certain contraindications. Thus, AI could generate discharge instructions for pancreatitis that recommend opioids which may not be appropriate for a patient with an undocumented history of opioid use disorder [13]. However, evidence also exists that AI-generated content is more empathetic than physician responses to medical questions in an online forum [14]. How consumer and physician trust and comfort with AI affects the doctor-patient relationship remains an open question.

By their nature, the connections in DL models are so numerous they cannot be explained, even by the AI. This creates a "black box" problem. It is therefore difficult, if not impossible, for humans to verify the logic behind the AI's processes. This may make it difficult for clinicians to understand and confidently act on the recommendations provided. What is more concerning is the phenomenon of "hallucinations", where an AI model generates incorrect or fictitious content and presents it as fact. For instance, an AI may report that a patient with heart failure has a particular ejection fraction when no recent echocardiogram is found in the EHR. AI chatbots have even fabricated scientific references [15]. Large efforts have been invested in eliminating hallucinations from LLMs and newer models are proving more accurate [16,17]. For example, when ChatGPT was asked 284 medical questions, GPT-4–the newer version–performed better than the older GPT-3.5. Both models improved when retested at a later date [18]. Yet, risks of hallucinations cannot be entirely eliminated. It is therefore critical that all AI-produced content be vetted by a clinician for accuracy.

2.1. What are existing and potential uses of AI in emergency care?

Here we describe uses of AI for acute, unscheduled care across the continuum of ED care from first symptoms to recovery (Fig. 2). We also describe other related uses, including quality improvement, billing, and research. We describe examples of existing and potential future modalities. Importantly, AI is rapidly evolving. Therefore, uses of AI will likely expand beyond the examples below as the technology evolves.

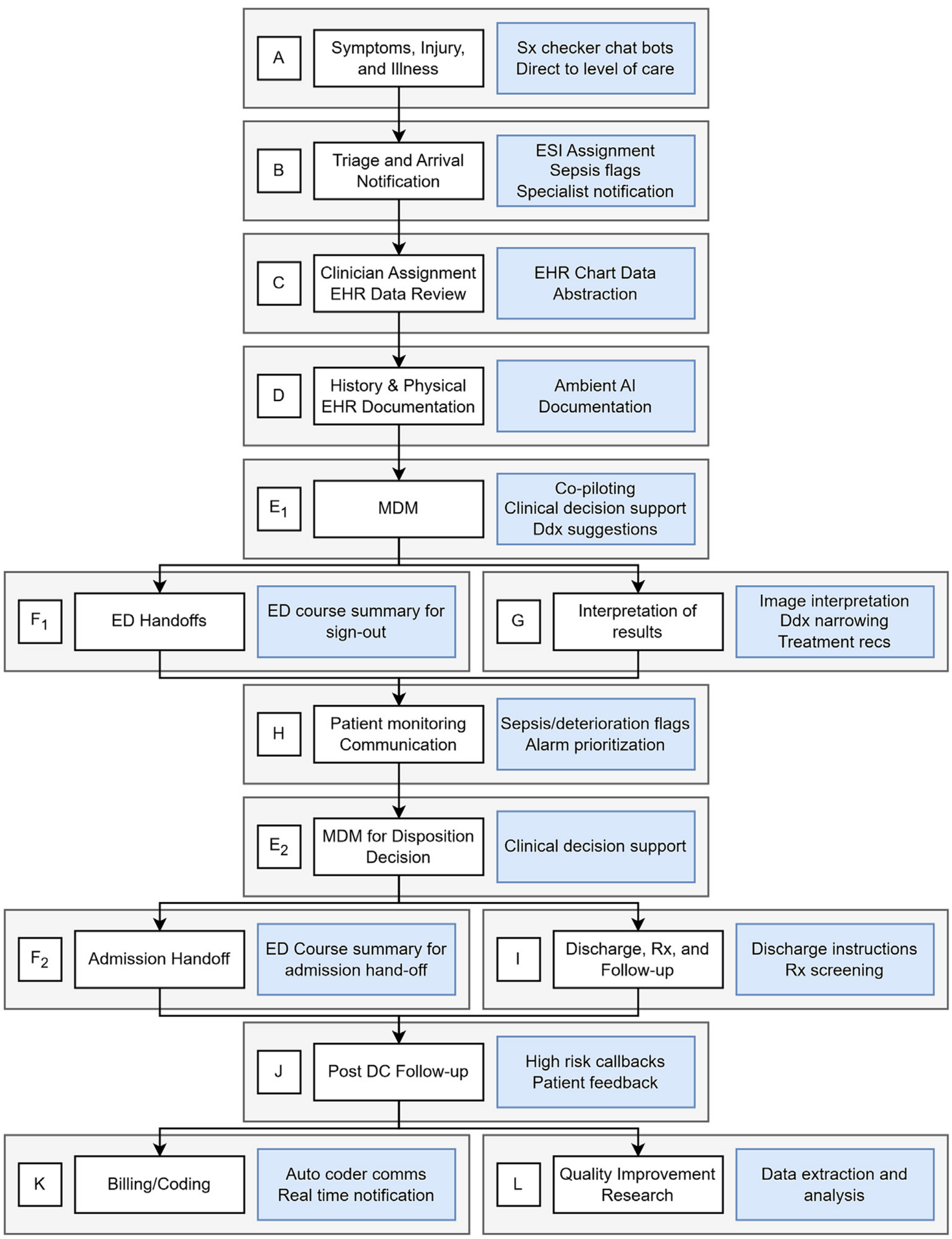

Fig. 2. Potential uses for artificial intelligence in acute and emergency care.

2.2. Symptom, injury, and illness (Fig. 2A)

When patients experience symptoms, suffer an injury, or become ill, it prompts a decision about whether, when, and where to seek care. In cases where the decision is obvious to the patient, they seek care based on the acute or perceived presence of an emergency [19]. Some may self-diagnose or self-triage with online resources, such as symptom checkers. These tools help with self-triage and generally rely on proprietary algorithmic approaches. Yet, they have been shown to have poor sensitivity and specificity [20]. The result is high rates of both over- and under-triage [21,22].

AI-powered chatbots offer advantages over symptom checkers. AI chatbots can take in more detailed, unstructured information including symptoms, physiological information, voice information, and camera data. Additionally, AI chatbots are conversational [23]. The combination of these attributes could improve and individualize recommendations for self-triage. However, the accuracy of these systems will be sensitive to the completeness and correctness of the data provided by the patient. Importantly, they may not properly assess subtle or atypical symptoms (e.g. an acute myocardial infarction presenting as shortness of breath).

Nevertheless, such systems have already been deployed and proven accurate in triage and even diagnosis compared to physicians. For example, the Babylon Triage and Diagnostic System, which uses a Bayesian network of primary care medicine, performed similarly to a group of seven physicians in evaluating clinical vignettes and providing an accurate differential diagnosis (80.0% AI sensitivity and 83.9% physician sensitivity) and safe triage recommendation (97.0% AI safety rating and 93.1% physician safety rating) [24]. More work is needed for validation and assessment of patient confidence in their recommendations. AI chatbots could potentially direct the patient to a particular setting of care (i.e. call 911, go to the ED, see a physician, or self-care, etc.). AI could also provide instructions on how to access care in a way that is most convenient and safe for the patient.

2.3. Triage, waiting for care, and arrival notification (Fig. 2B)

During the process of triage, basic demographic and visit information are collected verbally and through physiological measurement (e.g. temperature, blood pressure) by a nurse. Following data gathering, a triage level, most commonly using the emergency severity index (ESI) is assigned. Yet, studies have demonstrated that ESI assignments are frequently inaccurate and can exacerbate racial and sex based-biases [25-27]. AI-based systems could augment clinical judgment by utilizing data from the patient's triage and EHR to suggest an ESI level. For instance the KATE triage model was 26.9% more accurate than the average nurse [28] and the TriageGO model identified more than 10% of ESI 3 patients requiring up-triage who had increased risk of critical care, procedures, or hospitalization [29]. Such models may be more objective and reduce bias. Currently, algorithms that use Septic Inflammatory Response Syndrome (SIRS) criteria to flag patients at high risk of sepsis are helpful with the early detection and identification of patients for time-sensitive interventions [30]. However, their high sensitivity and low specificity raises concerns about unnecessary antibiotic and resource use [31]. AI models could utilize available unstructured data (e.g. comparison of vitals to a previous encounter or a concern about mental status raised in a nursing note) to more accurately identify patients with sepsis, decreasing false positive flag rates.

After triage, a patient may be given periodic updates on their care. Current systems exist to notify patients via text message of routine clinical actions (e.g. room assignments, order placement, discharge) [32], but this can also occur through an AI chatbot that could provide context and allow patients to communicate questions. Predictions of wait times can be provided by the AI based on previous trends, and taking into account current volumes, boarding, and staffing [33,34]. Chatbot communications may also include the status of any triage labs or the opportunity to complete registration.

Consultants or primary care physician could also be automatically notified by AI if their patient has presented to the ED with a relevant complaint. For instance, a surgeon may be notified if their patient presents after an operation, or an oncologist if their patient presents with neutropenic fever. AI models could identify such cases automatically, but with emphasis on specialist personalization, thereby decreasing the administrative burden of notifications by the ED care team while avoiding over-notification.

2.4. EHR data review (Fig. 2C)

Past information available for review in ED patients may range from none to years of visits and admissions. The inefficiency of EHR data gathering may lead physicians to proceed directly to the patient interview and rely on the patient provided information. The trouble is that patient knowledge of past medical history and medications is frequently incomplete [35,36]. This can lead to miscommunication of key information about relevant risk factors and prior care. The inefficiency of this process can lead to repetition of diagnostic tests already performed at another facility [37]. An alternative is that based on chief complaint, the AI generates a brief relevant report for the clinician. This information is pulled from the EHR and any connected health information exchanges (HIEs) (e.g. CareEverywhere) and includes relevant history, testing, and consultations. For example, for a patient with chest pain, the most recent stress test, echocardiogram, and cardiology consult may be presented, as well as prior visits for similar symptoms and the results of those evaluations. This information could be provided with links back to the original information location in the EHR to allow the clinician to verify and find further details.

2.5. Electronic Health record documentation (Fig. 2D)

Clinicians document their EHR notes in a variety of ways: during, after their encounter, sometimes after their shift, or via use of a paid human scribe. Clinicians documenting directly in the EHR during an encounter have been shown to have mixed effects on the patient-provider relationship [38]. By contrast, documenting after the shift may result in unrecalled critical information loss. An emerging AI use case that exists today is ambient AI, where the clinician-patient interaction is recorded and processed from speech-to-text. This is then converted into an appropriately formatted note which can be sent directly or copied into the EHR, for the clinician to verify, edit, and sign. Such products report increased efficiency and reduced burnout [39,40]. There are multiple pilots of such systems on-going, but thus far there is limited data on their use in ED settings.

2.6. Augmenting medical decision making (Fig. 2E)

When it comes to medical decision making (MDM), the AI may suggest tests or consultations. It could suggest up high-risk diagnoses from the pattern of data (e.g. the “tearing pain” of an aortic dissection) or suggest potential diagnoses for rarer presentations (e.g. fever and rash in a patient with travel from Africa). Suggestions would need to be designed in a non-intrusive way that integrates with workflow, thereby avoiding the disruptions and fatigue associated with traditional best practice advisories (BPAs) [41]. The AI would be able to monitor how these suggestions are dismissed or addressed by clinicians, thereby adjusting and personalizing further recommendation timing and content.

Information from the EHR and the patient interview can be utilized to assist the clinician with MDM. Validated clinical decision tools and scores (e.g. Wells for pulmonary embolism, National Emergency X-Radiography Utilization Study [NEXUS] for neck imaging) may be calculated and automatically shown to the physician. Tools could also be used at the time of disposition to determine appropriateness of outpatient vs inpatient treatment. For example, the History, Electrocardiogram (ECG), Age, Risk Factors, and Troponin (HEART) score can be used to risk stratify patients with chest pain for their likelihood of major cardiac events (MACE) post-discharge [42].

AI can already provide broad and helpful differential diagnoses even in complex cases, which is a capability that will likely improve as AIs are trained on EHR data [17]. In the future, suggestions for an expanded differential diagnosis could be provided based on available unstructured information in the note and EHR. Additionally, image recognition software can be utilized for dermatological complaints, wherein AI can assess a clinical photo to provide a differential diagnosis [43].

Decisions regarding testing and treatment can also be augmented by AI. For instance, allergy alerts and drug-drug interaction flags could be narrowed to those that are clinically appropriate/relevant (e.g. previous tolerance of cephalosporins in penicillin allergy). Or if a patient has multidrug-resistant organisms (MDROs), an alert can appear when ordering an antibiotic that patient cultures have been resistant to in the past, with links to the appropriate culture data.

2.7. Handoffs between emergency physicians (Fig. 2F)

Given the nature of shift work for both emergency physicians and nurses, handoffs frequently occur in the middle of a patient's course. Handoffs are sometimes rushed and often unstructured [44]. Unstructured handoffs have been shown to increase errors compared to a structured approach [45]. With an AI summary system, a relevant and brief sign-out could be created in a preferred format, such as the Situation, Background, Assessment, Recommendation (SBAR) or Illness severity, Patient Summary, Action list, Situation Awareness, Synthesis (IPASS) formats using the information obtained in the visit. The summary could also contain a list of pending items (e.g. imaging tests that have not been yet completed) and the current plan of care.

2.8. Interpretation of test results (Fig. 2G)

As results of tests become available, the AI can further expand or narrow the differential and assist by indicating which values are abnormal based on the patient's previous results (e.g. comparing creatinine or hemoglobin to baselines in the EHR and HIEs). AI could also be used for interpretation of radiographic images, especially in locations where a radiologist may not be immediately available. Models already assist in finding intracranial hemorrhages [2,46], pneumonias [47], pneumothoraces [48], and fractures [49]. Such models are in use now and have had recent improvements [50], but more work is needed to increase the accuracy and real-time nature of these analyses. Currently, sensitivities and specificities of such models vary, with some matching or exceeding radiologists (e.g. 91% sensitivity and specificity for fracture detection [49]), while others are less accurate (e.g. 69% sensitivity for subdural hemorrhage and 77% for acute subarachnoid hemorrhage.) [46] Additionally, further studies are needed to determine whether these technologies lead to additional confirmatory testing and increased length-of-stay. As results become available, the model could incorporate these results to narrow the differential diagnosis, calculate further risk scores and decision tools, and provide recommendations for further testing or treatment.

2.9. Patient monitoring and communication (Fig. 2H)

AI could use data such as vital signs and testing results to monitor patients in the ED for high-risk conditions (e.g. sepsis) and predict deterioration. Scoring systems, such as the electronic Cardiac Arrest Risk Triage (eCART) tool, which uses recent vital and lab trends to determine risk of inpatient deterioration requiring escalation to the intensive care unit (ICU), are already utilized [51,52]. Such tools could be applied and expanded to the ED setting. These systems could help nurses prioritize re-assessment of sicker patients. The AI model could also prioritize patients with clear dispositions to prompt the discharge or admission process.

2.10. ED discharge, prescriptions, and follow-up (Fig. 2I)

If a patient is deemed appropriate for discharge, AI can help create a personalized after visit summary or discharge instructions incorporating information from the visit, return precautions, and follow-up instructions. Despite their importance, current discharge instructions are frequently not understood [53]. AI could allow this information to be easily translated to other languages, to lower reading levels, or even converted into a short video, which may help address health and medical literacy disparities [54-56].

Prescriptions could be screened by AI for appropriateness, such as avoiding benzodiazepines in the elderly, safety of medications in pregnancy/breastfeeding, addressing previous MDROs, avoiding dangerous drug-drug interactions with patient's existing medications. They can also be screened for cost to the patient, providing alternatives if the medication would be cost-prohibitive and lead to non-adherence, a common issue with three in ten adults reporting not filling prescribed medications [57]. An appropriate automatic follow-up appointment could be scheduled or the specific specialist scheduling team could be notified to prompt an appointment to be created.

2.11. Post-discharge follow-up and recovery (Fig. 2J)

AI could automatically contact a patient after discharge to determine if they have been able to access their medications, follow-up appointments, and assess their symptoms. Unfortunately, most previous ED and inpatient readmission risk prediction models are inaccurate [58,59], which may require calling back all patients, a resource- and time-intensive task. Instead, AI models could potentially determine which patients are at high risk for being unable to access follow-up care or at high risk for return visits [59-61]. These patients can then be prioritized for check-ins by nurses or case managers, or by an AI-chatbot, which could assess the need for an immediate return visit or update the timing of longitudinal follow-up (e.g. request an earlier appointment with a specialist).

2.12. Billing and coding (Fig. 2K)

Currently, data from charts is manually reviewed by humans to determine coding. Errors and missing information are reported to the physician with requests for clarification. These requests may be delayed by weeks, at which point the physician may not recall the case. Instead, information for billing and coding could be extracted automatically and in real-time by AI. For example, as the physician signs the note, AI could inform the physician of deficiencies, such as independent interpretations of ECGs, additional diagnostic codes, or qualification for critical care time, which could be automatically estimated. Automated AI-based coding and billing could also decrease the cost of non-clinical staffing requirements for physician groups and hospitals.

2.13. Quality improvement (QI) and research (Fig. 2L)

Data on any number of research and QI questions could be pulled from the EHR from structured and unstructured sources using AI. While previously unstructured data would need to be extracted manually or using chart review, AI may be able to efficiently extract and summarize unstructured information, even in real-time [62]. For example, QI measures, such as sepsis documentation, could be rapidly assessed and reported back to clinicians efficiently to improve care. Health systems seeking to study a particular clinical issue, could use AI to query EHR data pre- and post- interventions. For example, AI could rapidly assess how often a HEART score was documented anywhere in the chart for patients with chest pain before and after an educational intervention.

3. Conclusions

In the future, AI has the potential to be integrated deeply into ED and acute unscheduled care. Importantly, despite the transformative use cases described above, we believe there will always be a role for human physicians in EDs and that AI will serve as an adjunct rather than a replacement for physician care. The industrial revolution increased specialization and decreased repetitive manual tasks. Similarly, the AI revolution will decrease rote information recall, documentation, and serve as a support for many processes as we have described. However, it will not eliminate the role of the emergency physician as the key communicator, empathizer, and ultimate decision maker in patient care. As the field evolves, physicians should strive to be informed leaders in AI development to ensure it is performed in a cautious, thoughtful, patient-centered manner. All AI-created materials must continue to be vigilantly verified by a clinician, as issues of model accuracy and bias can never be fully eliminated. Data safety, liability, and regulation will also evolve as AI adoption increases and physicians should strive to stay informed on such topics. The overall goals should focus on improving care quality and patient experience, decreasing costs, and decreasing cognitive, documentation, and emotional burdens on physicians and other ED staff.

Funding

None.

CRediT authorship contribution statement

Marika M. Kachman: Writing – original draft, Resources, Project administration, Conceptualization. **Irina Brennan:** Writing – review & editing, Writing – original draft, Conceptualization. **Jonathan J. Oskvarek:** Writing – review & editing, Writing – original draft, Conceptualization. **Tayab Waseem:** Writing – review & editing, Writing – original draft, Conceptualization. **Jesse M. Pines:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Conceptualization.

Declaration of competing interest

Dr. Pines has had funding from CSL Behring, Abbott Point-of-Care, and Astra Zeneca for unrelated work in the previous 36 months. No other authors have interests to declare. Generative AI was not used to create any part of this manuscript.

References

[1] What is Artificial Intelligence (AI)? | IBM. Accessed April 13, 2024. https://www.ibm.com/topics/artificial-intelligence; 2024.
[2] Heit JJ, Coelho H, Lima FO, et al. Automated cerebral hemorrhage detection using RAPID. Am J Neuroradiol. 2021;42(2):273–8. https://doi.org/10.3174/ajnr.A6926.
[3] What's the next word in large language models? Nat Mach Intell. 2023;5(4):331–2. https://doi.org/10.1038/s42256-023-00655-z.
[4] Singhal K, Azizi S, Tu T, et al. Large language models encode clinical knowledge. Nature. 2023;620(7972):172–80. https://doi.org/10.1038/s41586-023-06291-2.
[5] Adam H, Balagopalan A, Alsentzer E, Christia F, Ghassemi M. Mitigating the impact of biased artificial intelligence in emergency decision-making. Commun Med. 2022;2(1):149. https://doi.org/10.1038/s43856-022-00214-4.
[6] Obermeyer Z, Nissan R, Stern M, Eaneff S, Bembeneck EJ, Mullainathan S. Algorithmic bias playbook. https://www.chicagobooth.edu/-/media/project/chicago-booth/centers/caai/docs/algorithmic-bias-playbook-june-2021.pdf; 2021.
[7] Coalition for Health AI. The MITRE corporation. Blueprint for Trustworthy AI Implementation: Guidance and Assurance for Healthcare; 2022. https://www.coalitionforhealthai.org/papers/Blueprint%20for%20Trustworthy%20AI.pdf.
[8] Health C for D and R. Artificial Intelligence and Machine Learning in Software as a Medical Device. FDA; 2023. Published online October 20. Accessed January 9, 2024. https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device.
[9] Price WN, Gerke S, Cohen IG. Potential liability for physicians using artificial intelligence. JAMA. 2019;322(18):1765. https://doi.org/10.1001/jama.2019.15064.
[10] AMA Board of Directors. American Medical Association Principles for Augmented Intelligence Development, Deployment, and Use Published online; 2023.
[11] Powles J, Hodson H. Google DeepMind and healthcare in an age of algorithms. Health Technol. 2017;7(4):351–67. https://doi.org/10.1007/s12553-017-0179-1.
[12] Ray PP. A sober appraisal of artificial intelligence systems, particularly ChatGPT, in the facets of emergency medicine. Ann Emerg Med. 2023;82(6):766–7. https://doi.org/10.1016/j.annemergmed.2023.05.025.
[13] Bradshaw JC. The ChatGPT era: artificial intelligence in emergency medicine. Ann Emerg Med. 2023;81(6):764–5. https://doi.org/10.1016/j.annemergmed.2023.01.022.
[14] Ayers JW, Poliak A, Dredze M, et al. Comparing physician and artificial intelligence Chatbot responses to patient questions posted to a public social media forum. JAMA Intern Med. 2023;183(6):589. https://doi.org/10.1001/jamainternmed.2023.1838.
[15] Alkaissi H, McFarlane SI. Artificial hallucinations in ChatGPT: implications in scientific writing. Cureus. 2023. https://doi.org/10.7759/cureus.35179. Published online February 19.
[16] Ji Z, Lee N, Frieske R, et al. Survey of hallucination in natural language generation. ACM Comput Surv. 2023;55(12):1–38. https://doi.org/10.1145/3571730.
[17] Kanjee Z, Crowe B, Rodman A. Accuracy of a generative artificial intelligence model in a complex diagnostic challenge. JAMA. 2023;330(1):78. https://doi.org/10.1001/jama.2023.8288.
[18] Goodman RS, Patrinely JR, Stone CA, et al. Accuracy and reliability of Chatbot responses to physician questions. JAMA Netw Open. 2023;6(10):e2336483. https://doi.org/10.1001/jamanetworkopen.2023.36483.
[19] Ragin DF, Hwang U, Cydulka RK, et al. Reasons for using the emergency department: results of the EMPATH study. Acad Emerg Med. 2005;12(12):1158–66. https://doi.org/10.1197/j.aem.2005.06.030.
[20] Semigran HL, Linder JA, Gidengil C, Mehrotra A. Evaluation of symptom checkers for self diagnosis and triage: audit study. BMJ. 2015:h3480. https://doi.org/10.1136/bmj.h3480. Published online July 8.
[21] Hill MG, Sim M, Mills B. The quality of diagnosis and triage advice provided by free online symptom checkers and apps in Australia. Med J Aust. 2020;212(11):514–9. https://doi.org/10.5694/mja2.50600.
[22] Schmieding ML, Kopka M, Schmidt K, Schulz-Niethammer S, Balzer F, Feufel MA. Triage accuracy of symptom checker apps: 5-year follow-up evaluation. J Med Internet Res. 2022;24(5):e31810. https://doi.org/10.2196/31810.
[23] Tu T, Palepu A, Schaekermann M, et al. Towards Conversational Diagnostic AI Published online; 2024. https://doi.org/10.48550/ARXIV.2401.05654.
[24] Baker A, Perov Y, Middleton K, et al. A comparison of artificial intelligence and human doctors for the purpose of triage and diagnosis. Front Artif Intell. 2020;3:543405. https://doi.org/10.3389/frai.2020.543405.
[25] Sax DR, Warton EM, Mark DG, et al. Evaluation of the emergency severity index in US emergency departments for the rate of Mistriage. JAMA Netw Open. 2023;6(3):e233404. https://doi.org/10.1001/jamanetworkopen.2023.3404.
[26] Mistry B, Stewart De Ramirez S, Kelen G, et al. Accuracy and reliability of emergency department triage using the emergency severity index: an international multicenter assessment. Ann Emerg Med. 2018;71(5):581–587.e3. https://doi.org/10.1016/j.annemergmed.2017.09.036.
[27] Hinson JS, Martinez DA, Schmitz PSK, et al. Accuracy of emergency department triage using the emergency severity index and independent predictors of under-triage and over-triage in Brazil: a retrospective cohort analysis. Int J Emerg Med. 2018;11(1):3. https://doi.org/10.1186/s12245-017-0161-8.
[28] Ivanov O, Wolf L, Brecher D, et al. Improving ED emergency severity index acuity assignment using machine learning and clinical natural language processing. J Emerg Nurs. 2021;47(2):265–278.e7. https://doi.org/10.1016/j.jen.2020.11.001.
[29] Levin S, Toerper M, Hamrock E, 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):565–574.e2. https://doi.org/10.1016/j.annemergmed.2017.08.005.
[30] Gaieski DF, Mikkelsen ME, Band RA, et al. Impact of time to antibiotics on survival in patients with severe sepsis or septic shock in whom early goal-directed therapy was initiated in the emergency department*. Crit Care Med. 2010;38(4):1045–53. https://doi.org/10.1097/CCM.0b013e3181cc4824.
[31] Rhee C, Strich JR, Klompas M, Yealy DM, Masur H. SEP-1 has brought much needed attention to improving Sepsis care…but now is the time to improve SEP-1. Crit Care Med. 2020;48(6):779–82. https://doi.org/10.1097/CCM.0000000000004305.
[32] Brennan I, Datta M, Miner A, Dibble B. Enhancing emergency department patient experience and flow through real-time text messaging updates. NEJM Catal Innov Care Deliv. 2024.;5(4).
[33] Pak A, Gannon B, Staib A. Predicting waiting time to treatment for emergency department patients. Int J Med Inform. 2021;145:104303. https://doi.org/10.1016/j.ijmedinf.2020.104303.
[34] Trinh K, Staib A, Pak A. Forecasting emergency department waiting time using a state space representation. Stat Med. 2023;42(24):4458–83. https://doi.org/10.1002/sim.9870.
[35] Ley P. Memory for medical information. Br J Soc Clin Psychol. 1979;18(2):245–55. https://doi.org/10.1111/j.2044-8260.1979.tb00333.x.
[36] Cohen G, Java R. Memory for medical history: accuracy of recall. Appl Cogn Psychol. 1995;9(4):273–88. https://doi.org/10.1002/acp.2350090402.
[37] Stewart BA, Fernandes S, Rodriguez-Huertas E, Landzberg M. A preliminary look at duplicate testing associated with lack of electronic health record interoperability for transferred patients. J Am Med Inform Assoc. 2010;17(3):341–4. https://doi.org/10.1136/jamia.2009.001750.
[38] Alkureishi MA, Lee WW, Lyons M, et al. Impact of electronic medical record use on the patient-doctor relationship and communication: a systematic review. J Gen Intern Med. 2016;31(5):548–60. https://doi.org/10.1007/s11606-015-3582-1.

[39] Explore Nuance DAX for Clinicians | Nuance. Nuance Communications; 2024. Accessed October 29, 2023. https://www.nuance.com/healthcare/ambient-clinical-intelligence/explore-dax-for-clinicians.html.

[40] AI Medical Documentation for Physicians & Hospitals. Augmedix; 2023. Accessed October 29. https://augmedix.com/.

[41] Van Der Sijs H, Aarts J, Vulto A, Berg M. Overriding of drug safety alerts in computerized physician order entry. J Am Med Inform Assoc. 2006;13(2):138–47. https://doi.org/10.1197/jamia.M1809.

[42] Six AJ, Backus BE, Kelder JC. Chest pain in the emergency room: value of the HEART score. Neth Heart J. 2008;16(6):191–6. https://doi.org/10.1007/BF03086144.

[43] Breitbart EW, Choudhury K, Andersen AD, et al. Improved patient satisfaction and diagnostic accuracy in skin diseases with a Visual Clinical Decision Support System—A feasibility study with general practitioners. Houwink EJF, ed. PloS One. 2020; 15(7):e0235410. https://doi.org/10.1371/journal.pone.0235410.

[44] Cheung DS, Kelly JJ, Beach C, et al. Improving handoffs in the emergency department. Ann Emerg Med. 2010;55(2):171–80. https://doi.org/10.1016/j.annemergmed.2009.07.016.

[45] Starmer AJ, Spector ND, Srivastava R, et al. Changes in medical errors after implementation of a handoff program. N Engl J Med. 2014;371(19):1803–12. https://doi.org/10.1056/NEJMsa1405556.

[46] Seyam M, Weikert T, Sauter A, Brehm A, Psychogios MN, Blackham KA. Utilization of artificial intelligence-based intracranial hemorrhage detection on emergent noncontrast CT images in clinical workflow. Radiol Artif Intell. 2022;4(2):e210168. https://doi.org/10.1148/ryai.210168.

[47] Becker J, Decker JA, Römmele C, et al. Artificial intelligence-based detection of pneumonia in chest radiographs. Diagn Basel Switz. 2022;12(6):1465. https://doi.org/10.3390/diagnostics12061465.

[48] Plesner LL, Müller FC, Nybing JD, et al. Autonomous chest radiograph reporting using AI: estimation of clinical impact. Radiology. 2023;307(3):e222268. https://doi.org/10.1148/radiol.222268.

[49] Kuo RYL, Harrison C, Curran TA, et al. Artificial intelligence in fracture detection: a systematic review and meta-analysis. Radiology. 2022;304(1):50–62. https://doi.org/10.1148/radiol.211785.

[50] Rajpurkar P, Lungren MP. The current and future state of AI interpretation of medical images. Drazen JM, Kohane IS, Leong TY, eds. N Engl J Med. 2023;388(21):1981–90. https://doi.org/10.1056/NEJMra2301725.

[51] Bartkowiak B, Snyder AM, Benjamin A, et al. Validating the electronic cardiac arrest risk triage (eCART) score for risk stratification of surgical inpatients in the postoperative setting: retrospective cohort study. Ann Surg. 2019;269(6):1059–63. https://doi.org/10.1097/SLA.0000000000002665.

[52] Docheva N, Ravichandran U, Adams M, et al. 769 electronic cardiac arrest risk triage score as predictor of severe maternal morbidity. Am J Obstet Gynecol. 2021;224(2): S480. https://doi.org/10.1016/j.ajog.2020.12.792.

[53] Engel KG, Heisler M, Smith DM, Robinson CH, Forman JH, Ubel PA. Patient comprehension of emergency department care and instructions: are patients aware of when they do not understand? Ann Emerg Med. 2009;53(4):454–461.e15. https://doi.org/10.1016/j.annemergmed.2008.05.016.

[54] Khoong EC, Sherwin EB, Harrison JD, et al. Impact of standardized, language-concordant hospital discharge instructions on postdischarge medication questions. J Hosp Med. 2023;18(9):822–8. https://doi.org/10.1002/jhm.13172.

[55] Powers RD. Emergency department patient literacy and the readability of patient-directed materials. Ann Emerg Med. 1988;17(2):124–6. https://doi.org/10.1016/S0196-0644(88)80295-6.

[56] Williams DM, Counselman FL, Caggiano CD. Emergency department discharge instructions and patient literacy: a problem of disparity. Am J Emerg Med. 1996;14 (1):19–22. https://doi.org/10.1016/S0735-6757(96)90006-6.

[57] Kirzinger A, Motero A, Sparks G, Valdes I, Hamel L. Public Opinion on Prescription Drugs and Their Prices. https://www.kff.org/health-costs/poll-finding/public-opinion-on-prescription-drugs-and-their-prices/; 2023.

[58] Kansagara D, Englander H, Salanitro A, et al. Risk prediction models for hospital readmission: a systematic review. JAMA. 2011;306(15):1688. https://doi.org/10.1001/jama.2011.1515.

[59] Hong WS, Haimovich AD, Taylor RA. Predicting 72-hour and 9-day return to the emergency department using machine learning. JAMIA Open. 2019;2(3):346–52. https://doi.org/10.1093/jamiaopen/ooz019.

[60] Romero-Brufau S, Wyatt KD, Boyum P, Mickelson M, Moore M, Cognetta-Rieke C. Implementation of artificial intelligence-based clinical decision support to reduce hospital readmissions at a regional hospital. Appl Clin Inform. 2020;11(04):570–7. https://doi.org/10.1055/s-0040-1715827.

[61] Mohanty SD, Lekan D, McCoy TP, Jenkins M, Manda P. Machine learning for predicting readmission risk among the frail: explainable AI for healthcare. Patterns. 2022;3(1):100395. https://doi.org/10.1016/j.patter.2021.100395.

[62] Suh HS, Tully JL, Meineke MN, Waterman RS, Gabriel RA. Identification of Preanesthetic history elements by a natural language processing engine. Anesth Analg. 2022;135(6):1162–71. https://doi.org/10.1213/ANE.0000000000006152.